Improving the Reliability and Performance of Real-Time Communications in Mobile Ad Hoc Networks

by

Tolga Numanoglu

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Wendi B. Heinzelman

Department of Electrical and Computer Engineering

Arts, Sciences and Engineering

School of Engineering and Applied Sciences

University of Rochester

Rochester, New York

2009

Curriculum Vitae

The author attended the Electrical and Electronics Engineering Department at Middle East Technical University, Ankara, Turkey from 1999 to 2003 where he received his Bachelor of Science degree in Electrical and Electronics Engineering (High Honors) in 2003. He began graduate studies at the University of Rochester in 2003. He received the Master of Science degree from the University of Rochester in 2004. He worked with Siemens Austria, Program and System Engineering in Vienna, Austria during the summer of 2006. His primary research interests include wireless communications, ad hoc and sensor networks, signal processing, and information theory.

Acknowledgements

I would like to begin by thanking Professor Wendi Heinzelman, my thesis advisor and mentor for the past 6 years. Her support, encouragement, and enthusiasm motivated me to believe in myself towards my doctorate degree. It has been a privilege to work with her as a graduate student at the University of Rochester. I thank her for the all the energy and time she has spent for me, discussing everything from research to career choices, reading my papers, and guiding my research through the obstacles and setbacks. Her professional yet caring approach towards the people she works with and her passion for living the life to the fullest have truly inspired me.

I owe my thanks to the members of my thesis committee, Gaurav Sharma, Azadeh Vosoughi, and Daniel Stefankovic, for their valuable feedback. My special thanks go to Bulent Tavli, who collaborated on the work in this dissertation and has been a great mentor to me over the years. I would like to thank all my colleagues at the University of Rochester in the Wireless Communications and Networking Group for their help. Specifically, I would like to thank Mark Perillo, Chris Merlin, Stanislava Soro, Lei Chen, and Zhao Cheng.

I would like to thank my lovely fiancee Gordana, for all her love and support; for sticking with me and putting up with my bad moods and stress; and for always being there for me.

I would be nowhere in life if I hadn't grown up in the most wonderful family one can imagine. I want to thank my parents, Numan and Nesibe, as well as my sister, Burcu, for working so hard for me, for their love and for giving me all the happiness and opportunities that most people can only dream of. I couldn't have done this without them and I dedicate this thesis to them.

This research was made possible in part by the Center for Electronic Imaging Systems (CEIS), a New York State Office of Science, Technology, and Academic Research

(NYSTAR) designated center for advanced technology, and in part by the Harris Corporation, RF Communications Division.

Abstract

Mobile ad-hoc networks (MANETs) are expected to provide certain levels of Quality of Service (QoS) under varying wireless channel capacity and noise constraints. These noisy and varying channel conditions make it difficult to provide reliable and efficient real-time communication in MANETs. In this thesis, we explore three techniques to enable better utilization of mobile ad hoc networks under varying channel conditions: (1) employing a sufficient level of coordination among the nodes, (2) using a superposed coding scheme to provide multiple data rates for users with different channel capacities through a single transmission, and (3) utilizing a mesh networking inspired multicasting approach to vary the amount of redundancy in the routing process to overcome the performance loss due to channel errors. Specifically, we explore the effects of channel noise on different types of mobile ad-hoc networking protocols when channel capacities vary dynamically due to the unpredictable nature of the wireless channel. Our work shows that utilizing coordination among the nodes in MANETs leads to better throughput and energy efficiency for the network while maintaining acceptable packet delay as imposed by the application, even in the presence of relatively high channel error rates. Furthermore, we propose a method of utilizing different channel capacities simultaneously in order to provide individual network users the ability to select the appropriate delay-throughput trade-off for multi-hop routing in MANETs. This is done by exploiting the abilities of superposed coding to provide multiple data rates to receivers simultaneously while using much less energy and bandwidth compared to traditional methods that provide multiple data rates. In addition to the multi-rate multicasting, we propose a mesh networking inspired approach that adapts the amount of redundancy according to the current link conditions. We show that this approach can achieve good QoS levels for real-time traffic scenarios while simultaneously reducing unnecessary energy dissipation. We further explore techniques for combating the prob-

lem of lossy links in mobile ad hoc networks, and we propose a qualitative model that can provide guidance as to how different approaches should be utilized together for a given scenario.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Understanding the challenges and rewards of using wireless channels for communication has been and will continue to be at the forefront of communications research for the foreseeable future. From a simple traditional single hop broadcast communication to multi-user, mobile ad-hoc networking, researchers have aimed to achieve more reliable, more robust, and faster wireless communications. The idea that drives the need for ever increasing performance of wireless networks is to be able to perform the same tasks using wireless networks with mobile users as can be performed using wired networks with stationary users. Cellular phones and wireless local area networks are two of the most significant examples that have changed the way we look at wireless networks. These type of networks require a base station, which is called infrastructure, to be able to coordinate the wireless communication between the users.

As a military driven research area, mobile ad-hoc networks (MANETs) emerged to overcome the dependence on infrastructure and provide wireless communication virtually everywhere. In addition to the unreliable characteristics of wireless channels, the lack of infrastructure increases the already dynamic and difficult to predict behavior of mobile wireless communications. There is a need for distributed coordination among the network nodes, along with dynamic adaptation of the network protocols as channel conditions change, in order to overcome the challenges presented by MANETs.

1.1 Mobile Ad-hoc Networks and Their Challenges

Physically, a MANET includes a number of geographically distributed, potentially mobile nodes connected by wireless links. Compared with other types of networks, such as cellular networks or satellite networks, the most distinctive feature of MANETs is the lack of any fixed infrastructure. The network is formed of mobile (and potentially stationary) nodes, and is created on the fly as the nodes communicate with each other. The network does not depend on a particular node, and it dynamically adjusts as some nodes join or others leave the network.

The communication network, formed by a union of the links between these nodes, is called a MANET. A MANET can either be a standalone entity or an extension of a wired network. While the military is still a major driving force behind the development of these networks, ad-hoc networks are quickly finding new applications in civilian and commercial areas. Some of the applications where MANETs can be utilized are listed as follows:

- In a hostile environment where a fixed communication infrastructure is unreliable or unavailable, such as in a battlefield.
- Support for search and rescue missions in areas with little or no wireless infrastructure support.
- Replacement for the destroyed infrastructure in disaster relief operations.
- Provide support for people and applications to exchange data in the field or in a classroom without using any network structure except that which they create by simply turning on their computers or PDAs.

As wireless communication increasingly penetrates into everyday life, new applications for MANETs will continue to emerge and become an important factor in wireless communications. Yet, MANETs pose serious challenges to designers. Driven by the number of possible applications, countless mobile ad-hoc networking protocols have been proposed to satisfy the needs of these applications. MANET protocols require major modifications on the communication stack, which traditionally operates in a layered fashion with restricted information exchange between successive layers only, in

order to optimize many contradicting parameters such as energy efficiency, Quality of Service (QoS), throughput and latency. Therefore, designing protocols for MANETs requires a more intelligent design than is required for protocols designed for wireless local area networks and cellular networks where most of the communication load is on the infrastructure that is inter-connected with wires and has a virtually infinite energy supply. In addition to the limited energy resources and the lack of infrastructure, mobility adds another level of uncertainty on the performance of a MANET.

Due to the lack of a fixed infrastructure, nodes must self-organize and reconfigure as they move, join or leave the network. All nodes are essentially the same, and there is no natural hierarchy or central controller in the network. All functions have to be distributed among the nodes. Nodes are often powered by batteries and have limited communication and computation capabilities. Also, the bandwidth of the system is usually limited. The distance between two nodes often exceeds the radio transmission range, and a transmission may have to be relayed by other nodes before reaching its destination. Consequently, a MANET network typically has a multi-hop topology, and this topology changes as the nodes move around.

The difficulties of achieving a consistent level of performance in MANETs also arise from the fact that as compared to an ordinary wired interface, wireless channels are very noisy and their bit error rates are much higher. Therefore, packet losses are more likely, and network protocols cannot be designed on the assumption of perfect transmissions and receptions. A MANET protocol should be equipped with mechanisms to recover from or prevent frequent packet losses. Note that the corrupted packets are not only the data packets but also the control packets that network protocols rely on to coordinate network operation.

The unique characteristics of mobile ad-hoc networking constantly force us to find new ways of increasing reliability and the performance of the network. However, finding a way to combine or modify existing communication tools and mechanisms also provides a means of overcoming many of the difficulties imposed by MANETs. The essence of the difference between traditional networks and a MANET can be summarized as follows: MANETs have much more dynamic channel behaviors compared with traditional wired networks.

1.2 Real-time Communications in Mobile Ad-hoc Networks

Research on real-time communications in MANETs has been focused on providing real-time support for multimedia broadcasting, multicasting and unicasting while achieving an energy-efficient protocol design. Meeting these goals requires (*i*) minimizing energy dissipation, (*ii*) providing QoS for real-time data (*e.g.,* voice, video) packets, and (*iii*) enabling bandwidth-efficient multi-hop broadcasting and multicasting. Providing QoS for multimedia traffic (*e.g.,* voice) has been a design objective for many wireless network protocols [3–9]. Most of these protocols are designed either for single-hop networks or have QoS provisions in single-hop configurations, where a certain level of infrastructure is required. There are also a few protocol architectures [10, 11] that provide QoS in multi-hop networks.

QoS routing in MANETs continues to garner interest as designers continue to create protocols driven by the need to optimize the discovery, identification and securing of resources along the route. Because of the potentially limited shared bandwidth of the network, and the lack of a central controller that can account for and control these limited resources, nodes must negotiate with each other to manage the resources required for maintaining QoS along the routes. This is further complicated by frequent topology changes. Due to these constraints, QoS routing is more demanding than best-effort or minimum-hop routing where minimizing the number of hops is the main criteria.

Motivated by the advances in technology and the understanding and maturity of multi-hop ad-hoc networks, applications that require QoS such as multimedia broadcasting and multicasting are becoming important, and protocols are needed to improve the network's reliability and performance in supporting these services. Providing QoS for data communications has been investigated extensively in the literature [11–20]. However, as MANETs continue to increase in size and complexity, further QoS functionality may be needed along with ways to provide QoS in an energy-efficient manner.

1.3 Research Contributions

This research thesis presents a performance analysis of different classes of MAC protocols in the presence of channel noise and offers new protocols that provide better QoS support, including multi-level QoS and adaptive performance to meet QoS. This work jointly addresses several open research issues in mobile ad-hoc networks, as described previously. Specific contributions of this thesis include:

- An analysis of the QoS and energy efficiency behaviors of coordinated and non-coordinated MAC protocols in the presence of channel errors. Our analysis shows that the superior properties of coordinated MAC protocols are preferable over less complicated non-coordinated MAC protocols even under noisy channel conditions.

- The first joint analysis of energy efficiency and channel errors, which provides a better understanding of how energy efficiency is affected under different levels of channel noise in mobile ad hoc networks.

- Development of a multi-rate network-wide broadcasting algorithm that utilizes the idea of superposed coding to achieve multi-rate broadcasting with a single transmission. This multi-rate protocol is the first such protocol proposed in the research community that offers multi-level QoS through a single transmission.

- Design of a multi-rate multicasting protocol that enables the multicast group members to decide on different rates for multimedia streams. Future work will explore the limits of this approach.

- Design of a mesh networking inspired approach to overcome the performance degradation caused by lossy channels. The new protocol adaptively varies the level of redundancy in the broadcast mesh to improve QoS performance. This approach achieves good performance while simultaneously reducing unnecessary energy dissipation.

- A summary of adaptive techniques for combating the problem of lossy links in mobile ad hoc networks. A qualitative model that can provide guidance as to how and which approach(es) should be utilized for a given scenario is provoded.

1.4 Thesis Structure

Related work from the current literature is first presented in Chapter 2. In Chapter 3, we demonstrate the effects of channel errors on the control traffic of coordinated MAC protocols. We also compare the effects of channel errors on the two main types of MAC protocols, namely coordinated and uncoordinated. We provide an analytical model to ensure the reliability of our simulation results. In Chapter 4, we propose a superposed coding scheme to be used in network-wide broadcasting and multicasting in MANETs. We propose techniques to provide different rates for multicast group members with different channel capacities using a single broadcast transmission. Our adaptive redundancy approach, whereby the amount of redundancy is adjusted to the current link conditions, is presented in Chapter 5. Adaptive techniques for mitigating the effects of lossy channels in mobile ad-hoc networks are investigated and discussed in Chapter 6. The thesis summary and the future research directions are provided in Chapter 7.

Chapter 2

Related Work

In the last decade, motivated by demands and funding from military based projects, the field of protocol design for wireless mobile ad-hoc networks has grown dramatically. In this section, we review some of the work from the current literature that is relevant to this thesis. We include an overview of medium access control (MAC) protocols that are employed by the work in this thesis. Specifically, the Time Reservation Using Adaptive Control for Energy Efficiency (TRACE) protocol family is summarized separately due to the fact that this thesis is shaped by these protocols and aims to improve upon them. In addition to protocol overviews, a summary of performance analysis and multi-rate broadcasting are also included due to their relevance with the work described in this thesis.

2.1 TRACE Family of Protocols

In this section, in the first part we summarize the Multi-Hop Time Reservation using Adaptive Control for Energy efficiency (MH-TRACE) protocol [21]. MH-TRACE is the MAC protocol on which Network-wide Broadcasting and Multicasting through Time Reservation using Adaptive Control for Energy efficiency (NB-TRACE and MC-TRACE) are tailored [22, 23]. The basics of the NB-TRACE and MC-TRACE protocols are provided in the second and third part of this section. All of these protocols, MH-TRACE, NB-TRACE and MC-TRACE, are used for energy-efficient real-time data communications in MANETs.

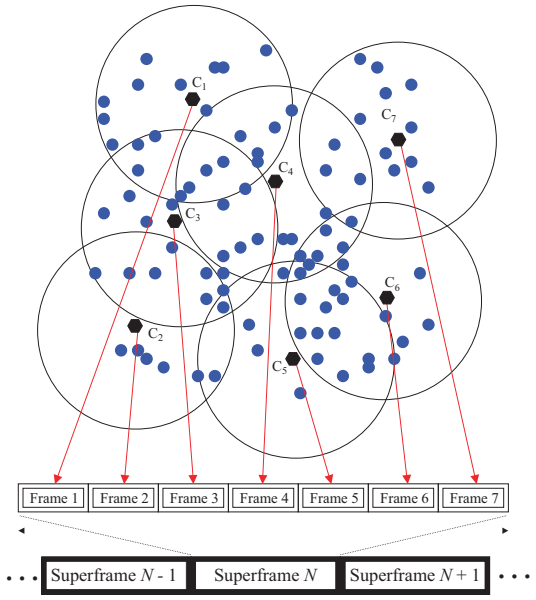


Figure 2.1: A snapshot of MH-TRACE clustering and medium access for a portion of an actual distribution of mobile nodes. Nodes C_1 - C_7 are clusterhead nodes.

2.1.1 MH-TRACE

Multi-Hop Time Reservation Using Adaptive Control for Energy Efficiency (MH-TRACE) is a MAC protocol designed for energy-efficient real-time data broadcasting [21]. Figure 2.1 shows a snapshot of MH-TRACE clustering and medium access. In MH-TRACE, the network is partitioned into overlapping clusters through a distributed algorithm. Time is organized into cyclic constant duration superframes consisting of several frames. Each clusterhead chooses the least noisy frame to operate within and dynamically changes its frame according to the interference level of the dynamic network. Nodes gain channel access through a dynamically updated and monitored transmission schedule created by the clusterheads, which eliminates packet collisions within the cluster. Collisions with the members of other clusters are also minimized by the clusterheads' selection of the minimal interference frame. Ordinary nodes are not static members of clusters, but they choose the cluster they want to join based on the spatial and temporal characteristics of the traffic, taking into account the proximity of the clusterheads and the availability of the data slots within the corresponding cluster.

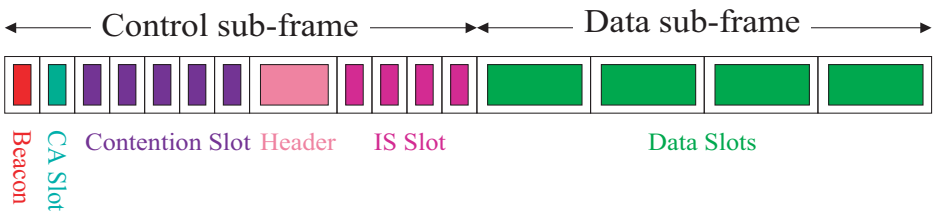


Figure 2.2: MH-TRACE frame structure.

Each frame consists of a control sub-frame for transmission of control packets and a contention-free data sub-frame for data transmission (see Figure 2.2). Beacon packets are used for the announcement of the start of a new frame; Clusterhead Announcement (CA) packets are used for reducing co-frame cluster interference; contention slots are used for initial channel access requests; the header packet is used for announcing the data transmission schedule for the current frame; and Information Summarization (IS) packets are used for announcing the upcoming data packets. IS packets are crucial in energy saving. Each scheduled node transmits its data at the reserved data slot.

In MH-TRACE, nodes switch to sleep mode whenever they are not involved in data transmission or reception, which saves the energy that would be wasted in idle mode or in carrier sensing. Ordinary nodes are in the active mode only during the beacon, header, and IS slots. Furthermore, they stay active for the data slots for which they are scheduled to transmit or receive. In addition to these slots, clusterheads stay in the active mode during the CA and contention slots. Instead of frequency division or code division, MH-TRACE clusters use the same spreading code or frequency, and inter-cluster interference is avoided by using time division among the clusters to enable each node in the network to receive all the desired data packets in its receive range, not just those from nodes in the same cluster. Thus, MH-TRACE clustering does not create hard clusters—the clusters themselves are only used for assigning time slots for nodes to transmit their data.

2.1.2 NB-TRACE

In NB-TRACE [22], the network is organized into overlapping clusters, each managed by a clusterhead (CH). Channel access is granted by the CHs through a dynamic, distributed Time Division Multiple Access (TDMA) scheme, which is organized into periodic superframes. Initial channel access is though contention; however, a node that utilizes the granted channel access automatically reserves a data slot in the subsequent superframes. The superframe length, T_{SF}, is matched to the periodic rate of voice generation, T_{PG}.

Data packets are broadcast to the entire network through flooding at the beginning of each data session. Each rebroadcasting (relay) node explicitly acknowledges (ACKs) the upstream node as part of its data transmission. Relay nodes that do not receive any ACK in T_{ACK} time cease to rebroadcast. As an exception, the CHs continue to rebroadcast regardless of any ACK, which prevents the eventual collapse of the broadcast tree. Due to node mobility, the initial tree will be broken in time, so NB-TRACE is equipped with several mechanisms to maintain the broadcast tree over time.

In NB-TRACE, a broadcast tree is formed by the initiation of a source node, however, once a tree is formed (*i.e.,* once nodes determine their roles), then other sources the existing organization to broadcast their packets.

NB-TRACE broadcasting and packet flow is illustrated in Figure 2.3. NB-TRACE is composed of five basic building blocks: (*i*) Initial Flooding (IFL), (*ii*) Pruning (PRN), (*iii*) Repair Branch (RPB), (*iv*) Create Branch (CRB), and (*v*) Activate Branch (ACB). The NB-TRACE algorithm flowchart is presented in Figure 2.4. Actually, all of theses building blocks are functioning simultaneously; however, we describe them as sequential mechanisms to make them easier to understand.

- **Initial Flooding (IFL):** A source node initiates a session by broadcasting packets to its one-hop neighbors. Nodes that receive a data packet contend for channel access, and the ones that obtain channel access retransmit the data they received. Eventually, the data packets are received by all the nodes in the network, possibly multiple times. At this point, some of the nodes have multiple upstream nodes. A node with multiple upstream nodes chooses the upstream node that has the least packet delay as its upstream node to be announced in its IS packet in order to minimize the delay.

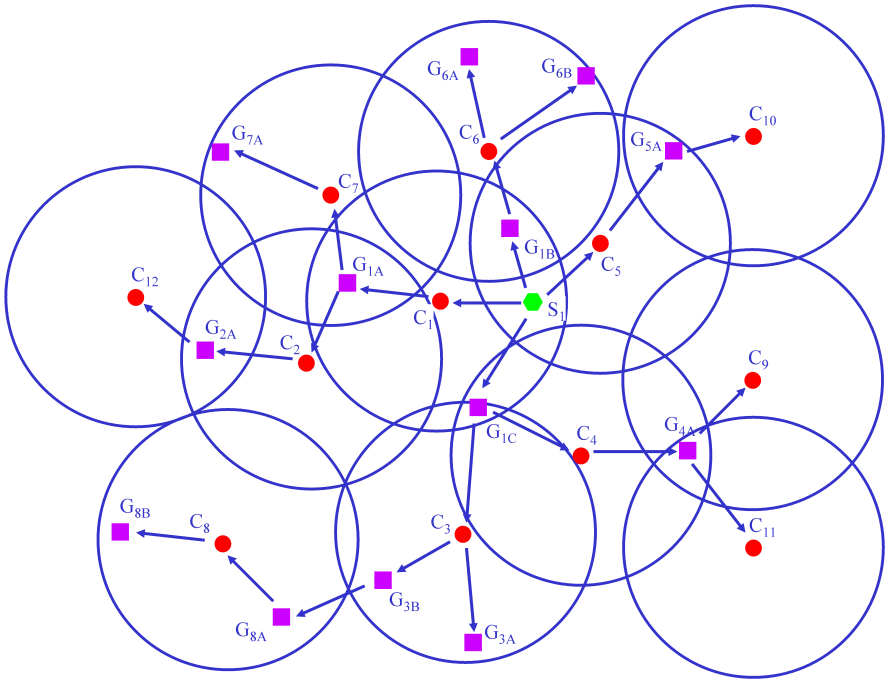


Figure 2.3: Illustration of NB-TRACE broadcasting. The hexagon represents the source node; disks are clusterheads; the large circles centered at the disks represents the transmit range of the clusterheads, squares are gateways, and the arrows represent the data transmissions.

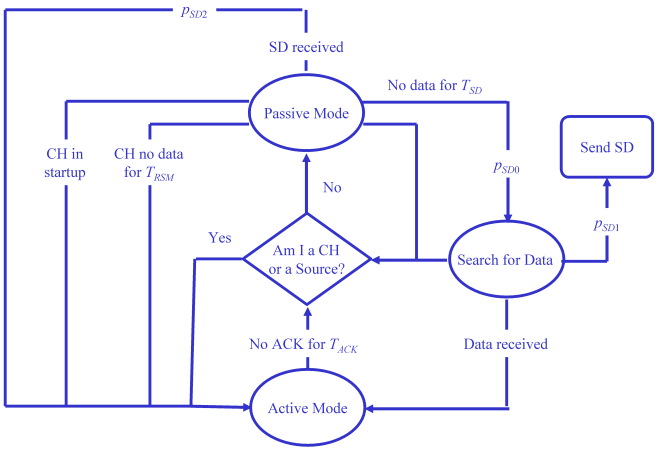


Figure 2.4: NB-TRACE flowchart.

- **Prunning (PRN):**The rebroadcasting nodes include the ID of the upstream node from which they first received the corresponding data packet in their IS packets, which provides an implicit acknowledgement for the upstream node. Relay nodes that do not receive an acknowledgement for T_{ACK} time cease rebroadcasting and return to passive mode. Nodes in passive mode do not relay packets, they just receive them, and nodes in active mode keep relaying packets. However, this algorithm has a vital shortcoming, which will eventually lead to the silencing of all relays.

  The outermost (leaf) nodes will not receive any acknowledgements, thus they will cease relaying, which also means that they cease acknowledging the upstream nodes. As such, sequentially all nodes will cease relaying and acknowledging, which will limit the traffic to the source node only.

  To solve this problem, we introduce another feature to the algorithm, which is that the CHs always retransmit, regardless of whether or not they receive an acknowledgement. Thus, the broadcast tree formed by initial flooding (IF) and pruning always ends at CHs. Note that the CHs create a non-connected dominating set. Thus, if we ensure that all the CHs relay broadcast packets, then the whole network is guaranteed to be completely covered.

- **Repair Branch (RPB):** One of the major effects of node mobility on NB-TRACE is the resignation of existing CHs and the appearance of new CHs. The appearance of a new CH generally is associated with the resignation of an existing CH. Whatever the actual situation, the nodes that receive a beacon packet from a CH in startup mode switch to active mode and rebroadcast the data packets they receive from their upstream neighbors until they cease to relay due to pruning.

  Figure 2.5 illustrates the RPB mechanism in a simple scenario. In the upper panel only node-1 is a CH and the broadcast tree consists of nodes 0 and 1. Nodes 2 and 3 receive data packets through node-1. However, due to the movement of node-3 (center panel), node-1 is out of the reach of node-3, thus, node-3 becomes a CH. Upon receiving the beacon of node-3, which indicates that it is in the startup mode, node-2, which was in the passive mode, switches to the active mode, thus, node-3 starts to receive data packets from node-2 (lower panel).

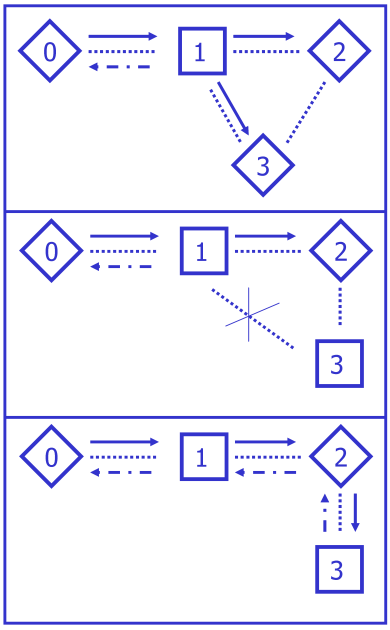


Figure 2.5: Illustration of the RPB mechanism.

- **Create Branch (CRB):** One of the basic principles of the NB-TRACE algorithm is that all the CHs should be rebroadcasting. If an ordinary node detects any of the CHs in its receive range is inactive for T_{CRB} time, then it switches to active mode and starts to rebroadcast data. As in the RPB case, redundant relays will be pruned in T_{ACK} time. The CRB mechanism is illustrated in an example scenario in Figure 2.6. Node-4, which is a CH, receives data through node-3 (upper panel). Due to mobility node-4 moves away from node-3 and the link between node-3 and node-4 is broken. However, node-4 enters into the receive range of node-2 (center panel). Upon detecting an inactive CH (node-4) in its receive range for T_{CRB} time, node-2 switches to the active mode and node-4 starts to receive data from its new upstream node, which is node-2 (lower panel).

- **Activate Branch (ACB):** An ordinary node that does not receive any data packets for T_{ACB} time switches to ACB mode, and sends an ACB packet with probability p_{ACB}. The underlying MH-TRACE MAC does not have a structure that can be used for this purpose, thus we modified MH-TRACE to be able to send ACB packets without actually affecting any major building blocks of MH-TRACE.

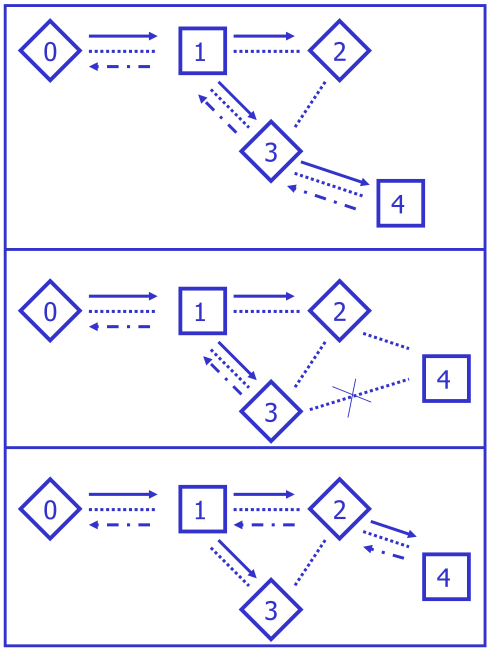


Figure 2.6: Illustration of the CRB mechanism.

ACB packets are transmitted by using the IS slots, because all the nodes will be listening to the IS slots regardless of their energy saving mode. Upon reception of an ACB packet, the receiving nodes switch to active mode, and start to relay data. If the nodes that receive ACB packets do not have data to send, they are either in ACB mode or they will switch to ACB mode. Upon receiving the first data packet, the nodes in ACB mode will switch to active mode.

2.1.3 MC-TRACE

Multicasting through Time Reservation using Adaptive Control for Energy efficiency (MC-TRACE), is an energy-efficient real-time data multicasting architecture for mobile ad hoc networks (MANETs) [23]. Being a cross-layer design, MC-TRACE utilizes a single integrated layer to perform the medium access control and network layer functionalities. The basic design philosophy behind the networking part of the architecture is to establish and maintain a multicast tree within a mobile ad-hoc network. This is achieved by using broadcasting to establish the desired tree branches and pruning the redundant branches of the multicast tree based on feedback obtained from the multicast

leaf nodes. Although these techniques have been used in many multicasting architectures in the past, the novelty in MC-TRACE is the re-engineering of these techniques in a highly energy-efficient manner for ad-hoc network multicast routing. Energy efficiency of the architecture is partially due to the medium access part, where the nodes can switch to sleep mode frequently; and partially due to the network layer part, where the number of redundant data retransmissions and receptions are mostly eliminated.

There are five basic building blocks in MC-TRACE: (i) Initial FLooding (IFL), (ii) PRuNing (PRN), (iii) MaiNtain Branch (MNB), (iv) RePair Branch (RPB), and (v) CReate Branch (CRB). MC-TRACE creates a broadcast tree through flooding (IFL) and then prunes redundant branches of the tree using receiver-based (or multicast leaf node-based) feedback (PRN). MC-TRACE ensures that every multicast node remains connected to the tree while reducing redundancy, and it uses IS slots so nodes can keep track of their role in the tree (*e.g.,* multicast relay node) as well as the roles of their neighbors. Finally, MC-TRACE contains mechanisms for allowing broken branches of the tree to be repaired locally (MNB and RPB) and globally (CRB).

The MC-TRACE architecture is designed for multiple multicast groups, and it can support multiple flows within each multicast group. Most of these mechanism were described before, however, the architecture for multicasting differs slightly from that of network-wide broadcasting. Therefore, we only provide a brief summary of these mechanism.

The initial flooding starts with the source node, which initiates a session by broadcasting packets to its one-hop neighbors. Nodes that receive a data packet contend for channel access, and the ones that obtain channel access retransmit the data they received. Eventually, the data packets are received by all the nodes in the network, possibly multiple times. Each retransmitting node acknowledges its upstream node by announcing the ID of its upstream node in its IS packet, which precedes its data packet transmission (see Figure 2.2 and Figure 2.7). The source node announces its own ID as its upstream node ID. Initially all retransmitting nodes announce a null ID as their downstream node ID. However, when an upstream node is acknowledged by a downstream node, the node updates its downstream node ID by the ID of this node. The leaf nodes (*i.e.,* nodes that do not have any downstream nodes that are acknowledging them as upstream nodes) continue to announce the null ID as their downstream node ID.

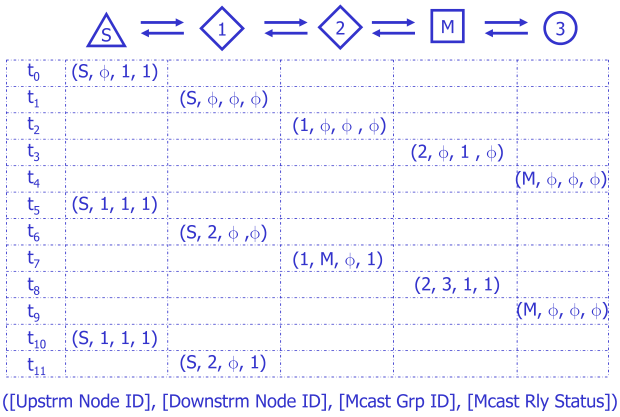


Figure 2.7: Illustration of initial flooding. Triangles, squares, diamonds, and circles represent sources, multicast group members, multicast relays, and non-relays, respectively. The entries below the nodes represent the contents of ([Upstream Node ID], [Downstream Node ID], [Multicast Group ID], [Multicast Relay Status]) fields of their IS packets (ϕ represent null IDs and t_i's represent time instants).

At this point, some of the nodes have multiple upstream nodes (*i.e.,* multiple nodes that have lower hop distance to the source than the current node) and downstream nodes (*i.e.,* multiple downstream nodes acknowledging the some upstream node as their upstream node). A node with multiple upstream nodes chooses the upstream node that has the least packet delay as its upstream node to be announced in its IS slot.

Multicast group member nodes indicate their status by announcing their multicast group ID in the IS packet (see Figure 2.7). If an upstream node receives an acknowledgment (ACK) from a downstream multicast group member, it marks itself as a multicast relay and announces its multicast relay status by setting the corresponding status (*i.e.,* multicast relay bit) in the IS packet. This mechanism continues in the same way up to the source node. Furthermore, a multicast group member that receives an ACK from an upstream multicast relay marks itself as a multicast relay also. Multicast relay status expires if no ACK is received from any downstream (for both members and non-members of the multicast group) or upstream (only for members of the multicast group) multicast relay or multicast group member for T_{RLY} time.

Initial flooding results in a highly redundant multicast tree, where most of the nodes receive the same data packet multiple times. Thus, a pruning mechanism is needed to

eliminate the redundancies of the multicast tree created by the initial flooding. Although in most cases initial flooding and pruning are capable of creating an initial efficient multicast tree, more mechanisms are needed to maintain the multicast tree in a mobile network. Maintain Branch, Repair Branch, and Create Branch mechanisms, similar to those described for NB-TRACE (see [23] for details), are utilized to maintain the multicast tree.

2.2 Broadcasting with IEEE 802.11

In broadcasting mode, IEEE 802.11 uses *p*-persistent CSMA with a constant defer window length (*i.e.,* the default minimum defer period) [8]. When a node has a packet to broadcast, it picks a random defer time and starts to sense the channel (see Figure 2.8). When the channel is sensed idle, the defer timer counts down from the initially selected defer time at the end of each time slot. When the channel is sensed busy, the defer timer is not decremented. Upon the expiration of the defer timer, the packet is broadcast.

The IEEE 802.11 standard includes an energy saving mechanism when it is utilized in the infrastructure mode [8]. A mobile node that needs to save energy informs the base station of its entry to the energy saving mode, where it cannot receive data (*i.e.,* there is no way to communicate to this node until its sleep timer expires), and switches to the sleep mode. The base station buffers the packets from the network that are destined for the sleeping node. The base station periodically transmits a beacon packet that contains information about such buffered packets. When the sleeping node wakes up, it listens for the beacon from the base station, and upon hearing the beacon responds to the base station, which then forwards the packets that arrived during the sleep period. This energy saving method results in additional delays at the mobile nodes that may affect QoS. Furthermore, this approach is not directly applicable in multi-hop networks. IEEE 802.11 also supports an energy saving mechanism in ad hoc mode called ad-hoc traffic indication message (ATIM) window, which is not an effective method for energy saving in broadcasting.

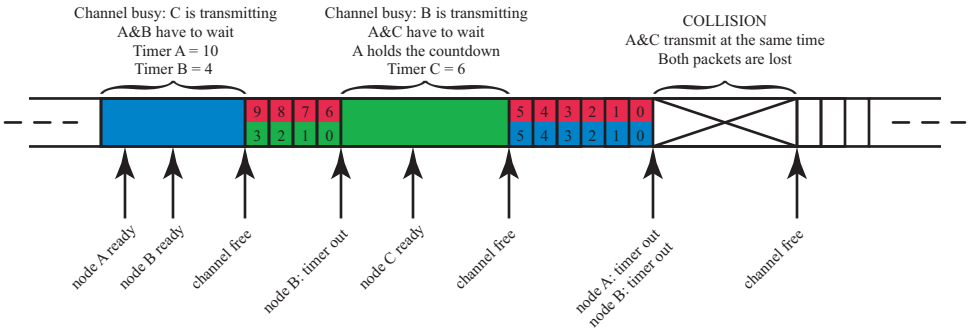


Figure 2.8: Illustration of IEEE 802.11 MAC mechanism in broadcasting.

2.3 Overview of Performance Analysis

Although performance analysis of ad-hoc networks has found some noticeable attention in the literature [24–32], there is little work done to explore the characteristics of different types of MAC protocols (*i.e.,* coordinated and non-coordinated) under varying channel noise. Our work explicitly aims to answer the question of whether a coordinated MAC protocol preserves its superior performance, or whether its higher level of vulnerability due to the dependence on the robustness of the control traffic makes it unstable under high BER levels.

A comprehensive survey of MAC approaches for wireless mobile ad-hoc networks is presented in [33]. This work individually concentrates on different MAC approaches and tries to identify their problems and discusses possible remedies. One of the important conclusions of [33] is that increased throughput results in increased energy efficiency due to a decrease in the number of retransmissions. However, it is also pointed out that one has to sacrifice some throughput in order to achieve fairness (*e.g.,* reservation based MAC protocols). Another major conclusion is that in order to achieve QoS one has to increase the persistence, which results in decreased throughput stability.

In [34] MAC protocols are compared in terms of battery power consumption in order to emphasize the characteristics of an energy-efficient MAC protocol. They concluded that reducing the number of contentions reduces the energy consumption. Moreover, reservation (*i.e.,* coordination and scheduling) is proposed as a better solution for messages with contiguous packets. However, energy efficiency under channel noise was not explored in this study.

More focused works investigating packet loss and error resilience can be found in [35] and [36]. These studies concentrated on identifying and characterizing possible sources of packet losses in ad-hoc wireless networks. Mobility and congestion are pointed out as the main reasons in mobile ad-hoc networks [35]. On the other hand, [36] takes collisions, error in radio transmission and SNR (Signal to Noise Ratio) variation into account as the main reasons for packet losses in mobile ad-hoc networks. They both provide simulation results to demonstrate the effects of each individual source of packet losses.

In [37] an adaptive frame length control approach, which is implemented at the MAC layer to compensate for the rapidly varying channel conditions of wireless networks, is presented. They showed that by adjusting the frame length, there is much to be gained in terms of throughput, effective transmission range and transmitter power for wireless links. All of their outcomes stem from the assumption that the probability of error for a longer packet is higher than the probability of error for a shorter packet. Therefore, reducing the frame length when the channel conditions are worse will improve the throughput since the effective transmission range is increased. As a result of improved throughput, less energy is consumed due to the reduced number of retransmissions as we discussed earlier. However, in their analysis they did not consider the effects of channel noise on control packets.

None of the aforementioned studies provide sufficient insight on the error resilience, in general, and the vulnerability of control traffic to channel noise, in particular, and hence the performance evaluation of MAC protocols under various BER levels. Although the impact of channel errors on the control packets is crucial to the overall performance of coordinated MAC protocols, evaluation of coordinated MAC protocols under realistic channel errors has found little attention in the literature. In this thesis, we investigate the effects of channel errors on the control traffic in a coordinated MAC protocol and determine the extent of performance deterioration. Furthermore, we present a comparative performance evaluation of a coordinated and a non-coordinated MAC protocol under a realistic error model. We believe that jointly analyzing the energy efficiency and error resilience of coordinated and non-coordinated MAC protocols and identifying the pros and cons of them will motivate future research to produce more accurate and reliable solutions to MAC related problems.

2.4 Overview of Multi-rate Broadcasting

Optimal utilization of the broadcast channel is still an active area of research. There are many studies on single hop multi-rate broadcasting, which was originally introduced by Cover [38]. The basic idea of Cover's paper is that in a broadcast scenario, where receivers have different channel capacities, a slight degradation in the rate of the worst channel can result in much higher information rates for the better channels. The limits of this new technique were explored to show that it outperforms the naive ways of broadcasting known previously [39].

Jung and Shea later called this new technique *simulcasting* [40], which is simultaneously transmitting multiple messages to different neighboring radios of a transmitter. In their work to improve the end-to-end throughput of a simple static network, they used a cross layer algorithm that employs slotted ALOHA for channel access. They investigated the original idea of Cover to show its impact on the link throughput, end-to-end throughput, and network connectivity. The conclusion of their work is that simulcasting can improve the link and end-to-end throughput in wireless ad hoc networks with only a slight degradation in other metrics, such as network connectivity. These results showed the basic capabilities and shortcomings of the multi-rate broadcasting technique. However, their work did not address issues dealing with multi-hop routing using multi-rate broadcasting. Utilizing this technique in a mobile multi-hop routing scenario has not been considered.

Before proceeding any further we want to make it clear that the term *simulcasting* is used as a portmanteau of "simultaneous broadcasting", and it usually refers to broadcasting the *same* event simultaneously by FM and AM radio or by radio and television. Jung and Shea imposed an additional meaning on this term [40], such that the term simulcasting means non-uniform modulation. Both simulcasting and non-uniform modulation describe a special case of multi-rate broadcasting, which is achieved by Cover's super-imposed coding idea and is not limited to any specific constellation [38]. In this thesis, we choose to call our proposed multicasting protocol, which utilizes Cover's superimposed coding idea, *multi-rate multicasting* to avoid any confusion and keep a consistent notation throughout the thesis.

Sun et al. [41] introduce a superposed turbo coding scheme that uses the multi-rate broadcasting idea to achieve flexible unequal error protection. Using a multi-rate turbo

coding scheme, they maintain a performance level similar to that of single level turbo trellis-coded-modulation schemes. They point out the fact that their multi-level coding scheme offers the same complexity and performance as that of typical single-rate coding, but provides the flexibility of controlling the amount of unequal error protection.

Our study is motivated by these already proven capabilities of multi-rate broadcasting. We combine this idea of multi-rate broadcasting with multi-hop routing, and we show the improvement this integration can bring to a multi-hop network. The idea of multi-rate availability is also used in a different way when bringing TV services to mobile phones, which is referred to as *mobile TV*. In mobile TV broadcasting, different rates carry different streams that can be decoded by users according to their channel quality. Broadcasters include turbo-coded streams of lower-resolution video, suitable for viewing on PDAs or cellphones, to be delivered alongside a high definition broadcast within a single TV channel. Reception of these robust low-resolution streams is possible at high mobile speeds, such as in a car or on a train, as well as urban environments, which have a lot of interference from buildings and objects [42].

2.5 Overview of Adaptive Mesh Networking

2.5.1 Multicast Routing Protocols

There are many multicast routing protocols designed for mobile ad hoc networks, and they can be categorized into two broad categories [43]: (i) tree-based approaches and (ii) mesh-based approaches. Tree-based approaches create trees originating at the source and terminating at multicast group members with an objective of minimizing a cost function. For example, the cost function to be minimized can be the distance between the source and every destination in the multicast group [44]. A multicast protocol for ad hoc wireless networks (AMRIS) [45] constructs a shared delivery tree rooted at one of the nodes with IDs increasing as they radiate from the source. Local route recovery is made possible due to this property of IDs, hence reducing the route discovery time and also confining route recovery overhead to the proximity of the link failure.

Mesh-based multicasting is better suited to highly dynamic topologies, simply due to the redundancy associated with this approach [46, 47]. In mesh-based approaches there is more than one path between the source and multicast group members (*i.e.,* a

redundant multicast-tree); thus, even if one of the paths is broken due to mobility the other paths may be available.

One such mesh-based protocol is ODMRP [46], which is based on periodic flooding of the network by the source node through control packets to create a multicast mesh. Instead of using a tree, ODMRP utilizes a mesh structure, which is robust, to compensate for the frequent route failures and trades-off bandwidth for stability, which comes with redundancy (*i.e.,* there are multiple redundant routes between the source and destinations).

2.5.2 Improving QoS Under Varying Link Conditions

We can summarize the previous work on improving QoS under varying link conditions in ad-hoc networks in three categories; (i) adapting the channel coding to overcome the varying link conditions [48], (ii) choosing routes according to link conditions in order to avoid unnecessary retransmissions and improve the delay performance [49–51], and (iii) varying the number of redundant links between the nodes of the network to increase the efficiency [52–54].

Lin et al. [48] point out that the time-varying nature of wireless channels is often ignored in ad-hoc routing scenarios. In order to overcome the performance loss due to the time-varying wireless channel, they propose a channel adaptive routing protocol that utilizes an adaptive channel coding and modulation scheme to dynamically adjust the amount of error protection. They achieve shorter delays and higher rates at the expense of a higher overhead in route set-up and maintenance. Although we plan to include this type of an approach in our future work, we leave adaptive channel coding out of the scope of this study and focus on solving the variable link problem by routing only.

Many others proposed that instead of trying to adapt the level of error protection, choosing routes that are more stable or less prone to channel errors would increase the QoS performance of the network [49–51]. They found that selecting better routes according to the predicted state of the links results in fewer unnecessary packet losses and retransmissions, increased energy efficiency, and resulted in a stable and good quality communication route. However, predicting the state of multiple routes between two nodes and choosing a route that is more stable than the others may not be favorable

for ad-hoc networks where distributed and highly scalable protocols are needed. Moreover, the time-varying network behavior makes it even harder to maintain a reliable route table.

Due to the success of mesh based routing protocols in maintaining stable communication in the face of changing link conditions [46, 55], researchers also investigated ways to reduce or adapt the number of redundant links in a route between source and destination according to the dynamic behavior of the network [52–54]. The common motivation for these works is the fact that although utilizing multiple paths from senders to receivers results in higher reliability, in less dynamic environments the additional redundancy may not be needed in terms of reliability, and increased redundancy causes increased overhead. We would like to point out that these works do not take channel errors into account in their attempts to improve the reliability and efficiency of routing in ad-hoc networking. Instead, they consider network density, mobility, and traffic load as the dynamic factors that cause the link breaks. The previous two approaches, on the other hand, considered channel errors as the source of performance loss.

Our approach differs from the literature in the way that we take channel errors and node mobility into consideration at the same time and propose an adaptive approach that aims to keep the energy consumption of the network under control while providing better QoS support. We employ a bottom-up approach by pointing out the effects of increasing redundancy and showing the trade-off between reliability and energy consumption, as opposed to the previous adaptive approaches, which start with a highly redundant routing protocol and strip off additional redundancy in less dynamic scenarios. In other words, we start with a tree-based routing protocol, which is already energy and bandwidth efficient, and improve its performance through adaptive redundancy under channel errors. Previous works in [52–54], however, focus on a highly redundant mesh routing protocol and try to optimize the efficiency of this mesh-based routing protocol under different scenarios by varying the number of redundant links between the nodes. There is no attempt to analyze the effects of channel errors. In fact, their approaches can only adapt to events that happen for a long duration of time since their adaptation heavily depends on the history of statistics gathered for an extended period of time (between 2 to 5 seconds). Therefore, their solutions do not promise any QoS guarantees against time-varying wireless channels.

We can list our contributions in this area under three categories; (i) we propose an adaptive routing redundancy algorithm that can be tailored to energy and bandwidth efficient tree-based routing protocols, (ii) our adaptive redundancy approach preserves the advantageous key features of the tree-based routing protocol and activates/deactivates itself as channel conditions get worse/better due to the time/space varying wireless channels, and (iii) we achieve routing adaptation through local adaptation rather than globally adjusting the redundancy in the network or keeping a long history of packet receptions.

Chapter 3

Performance Analysis of MAC Protocols Under Channel Errors

In wireless communications, the channel is the common interface that connects the nodes. Like every shared resource, access to the channel needs to be regulated; this resource allocation operation is performed by Medium Access Control (MAC) protocols, which are defined as the second layer of the OSI protocol stack [56]. The objective of controlling access to the channel via the MAC protocol is to avoid or minimize simultaneous transmission attempts (that will result in collisions) while maintaining a stable and efficient operating region for the whole network. Furthermore, the MAC protocol is the key element in determining many features of a wireless network, such as energy efficiency, throughput, Quality of Service (QoS), fairness, stability, and robustness [57].

MAC protocols can be classified into two categories based on the collaboration level of the network in regulating the channel access: coordinated and non-coordinated. A coordinated MAC protocol operates with explicit coordination among the nodes and is generally associated with coordinators, channel access schedules and clusters. For example, Bluetooth is a coordinated MAC protocol, where channel access within a cluster (*i.e.,* piconet) is coordinated by a coordinator (*i.e.,* piconet Master) [58]. A non-coordinated MAC protocol, on the other hand, operates without any explicit coordination among the nodes in the network. For example, IEEE 802.11 is a non-coordinated MAC protocol when operating in the broadcast mode (*i.e.,* in broadcasting mode, IEEE 802.11 becomes plain CSMA without any handshaking) [8]. Note that IEEE 802.11

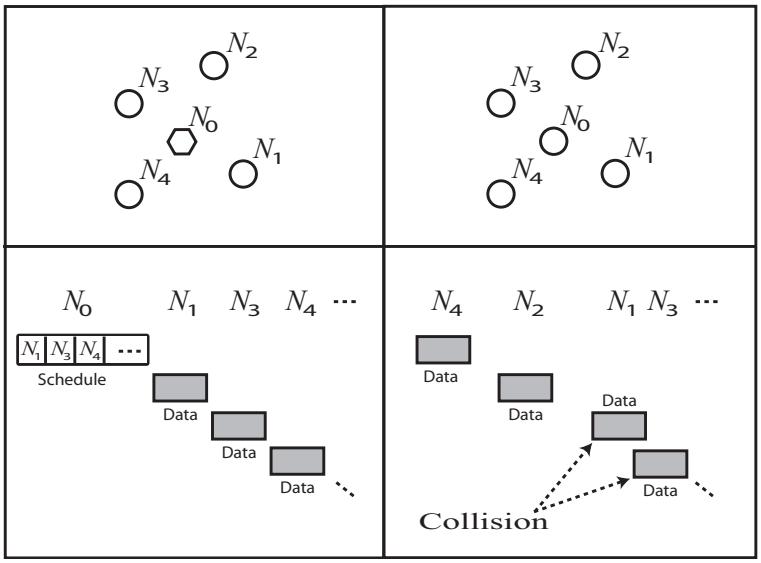


Figure 3.1: Illustration of coordinated and non-coordinated MAC protocols. The upper left and right panels show the node distributions for nodes N_0-N_4. The lower left panel shows the medium access for the coordinated scheme, where node N_0 is the coordinator and the channel access is regulated through a schedule transmitted by N_0. The lower right panel shows the channel access for the non-coordinated scheme (*e.g.,* CSMA). Overlapping data transmissions of N_1 and N_3 lead to a collision.

channel access in unicasting mode is a coordinated scheme (*i.e.,* the four way handshaking between the transmitter and receiver is a special case of a general explicit coordination scheme, such as [59, 60]).

3.1 Coordinated vs. Non-coordinated MAC Protocols

Figure 3.1 illustrates the channel access mechanism for generic coordinated and non-coordinated MAC protocols. In the coordinated MAC protocol, node N_0 is the clusterhead (coordinator) for the portion of the network consisting of five nodes. Channel access is regulated through a schedule that is broadcast by the coordinator. Upon reception of the schedule, nodes transmit their data at their allocated time, and thus collisions among nodes within the same cluster are eliminated. Furthermore, a node can switch

to a low-energy sleep mode during the slots where no transmissions are scheduled or scheduled transmissions are not of interest to a particular node. Time is organized into cyclic time frames, and the transmission schedule is dynamically updated at the beginning of each time frame. IEEE 802.15.3 is a recent example of such a coordinated MAC protocol [2]. In the non-coordinated MAC protocol, each node determines its own transmission time based on feedback obtained through carrier sensing on the channel. Thus, conflicts in data transmission attempts (*i.e.,* collisions, capture) are unavoidable in the non-coordinated scheme. In addition, none of the nodes can switch to sleep mode because future data transmissions are not known beforehand due to the lack of a scheduling mechanism.

Both coordinated and non-coordinated MAC protocols have their advantages and disadvantages.

(i) One of the most important advantages of coordinated MAC protocols is their energy efficiency due to the availability of a schedule that lets nodes enter into sleep mode without deteriorating the overall system performance. Thus, the average energy dissipation of nodes in coordinated schemes is significantly lower than in non-coordinated schemes [34].

(ii) Collisions are mostly eliminated in coordinated MAC protocols, while frequent packet collisions are unavoidable in non-coordinated protocols, especially under heavy network conditions, which may draw the network into instability in extreme conditions [61].

(iii) The average packet delay using non-coordinated MAC protocols is lower than the average packet delay using coordinated MAC protocols under mild traffic loads. However, under heavy traffic loads, packet delay in non-coordinated protocols rises to very high levels [62].

(iv) Coordinated MAC protocols are more vulnerable to packet losses than non-coordinated MAC protocols due to their dependence on the reliable exchange of control packets, such as the schedule packet. Mobility, multi-path propagation, and channel noise are the main sources of errors that cause packet losses [63].

Energy efficiency has become one of the predominant platform requirements for battery powered mobile multimedia computing devices. Therefore, the new challenge

is to provide QoS in an energy-efficient manner rather than focusing solely on QoS by ignoring the energy dissipation [64]. Consequently, there is a growing interest in energy-efficient design, mainly concentrating on MAC layer energy reduction techniques [34, 65, 66]. Most of the proposed solutions use TDMA as a MAC scheduling principle in order to utilize the benefits of having a schedule such as fairness, stability and energy efficiency by regulating the channel access, minimizing collisions and enabling power saving features, respectively.

In general, energy-efficient distributed protocol design can be described as creating an appropriate distributed coordination scheme that minimizes a radio's total energy dissipation without sacrificing its functionality, by intelligently switching between the radio's different operating modes. Actually, there are only three modes that a radio can be switched to: transmit mode, active mode (receive, carrier sense and idle modes), and sleep mode. Although further classification of the energy dissipation modes of a radio is possible (*i.e.,* deep/shallow sleep modes, transient modes, *etc.*), the aforementioned classification is detailed enough in this context. There is no way to switch between receive, idle, and carrier sense modes: when a node is in the active mode, the actual mode (receive, idle or carrier sensing) is determined by the activities of the node's neighbors, which is not a controllable design parameter. Nevertheless, the ultimate goal is to keep the radio in the low energy sleep mode as long as possible without sacrificing network performance.

The general trend in the evaluation of the performance of network protocols (*e.g.,* energy efficiency) is to ignore channel errors and assume a perfect channel [67]. Although the assumption of a perfect channel is reasonable in the initial design stage, further verification of a proposed protocol should consider error resilience. In this chapter we investigate the performance of two MAC protocols, IEEE 802.11 and MH-TRACE (Multi Hop Time Reservation using Adaptive Control for Energy Efficiency), at different Bit Error Rate (BER) levels by providing an analytical model which is well supported with ns-2 simulations. IEEE 802.11 is a well-known example of a non-coordinated MAC protocol when it is used for broadcasting. MH-TRACE is a recent example of an energy-efficient coordinated MAC protocol that relies on control packet exchanges for its operation. A comparative evaluation of IEEE 802.11 and MH-TRACE for real-time data broadcasting using a perfect channel showed that the energy

efficiency of MH-TRACE is much better than IEEE 802.11 [21]. However, due to the relatively complicated design of MH-TRACE the advantages of MH-TRACE over IEEE 802.11 are questionable under high BER levels.

3.2 Effects of Losing Control Packets

In this section we investigate the effects of control packet losses on protocol performance. Since the non-coordinated MAC protocol (IEEE 802.11) does not utilize control packets, this section focuses on the coordinated MAC protocol (MH-TRACE), exclusively. We consider a real-time voice broadcasting application, where each voice packet has a delay constraint and must be dropped after the transmission delay exceeds a certain threshold.

MH-TRACE control packets have vital importance in keeping the clustering and the scheduling mechanisms intact. However, it is not obvious what types of effects the different control packets have on protocol performance. Average network throughput, which is the total number of data packets received by all the nodes in the network, is used as the performance metric. This metric is appropriate for a single hop broadcasting scheme because errors in the control packets will directly affect the number of transmitted data packets since nodes may not be assigned channel access in a timely manner and will thus need to drop packets. As a result of dropped data packets, the number of received data packets will drop linearly. Therefore, the simulations in this section provide a better understanding of the vulnerability of the medium access protocol to control packet losses.

For the simulations in this section we use a six node fully connected static network to clearly observe the effects of packet losses. When there are no channel errors, all nodes should be able to transmit and receive without any packet drops or collisions. There will be only one clusterhead in the network due to the fact that there cannot be two clusterheads that can hear each other directly. We used the *ns*-2 simulator to evaluate the system performance. Simulation parameters are given in Table 3.3. The channel rate is set to 2 Mbps, and we simulated conversational voice coded at 32 Kbps, which corresponds to one voice packet per superframe. The simulations are run for 1000 s and repeated with the same parameters five times.

Beacon, header and contention packets form the backbone of the protocol and enable MH-TRACE to operate efficiently. Thus, we investigate the effects of loosing each of these control packets on the performance of MH-TRACE.

3.2.1 Beacon

Beacon packets are used to announce the existence and continuation of the clusterheads to the nodes in the transmit range of the clusterheads. Since the beacon packet is the main control packet for clusterheads to inform the other nodes about their existence, the stability of a cluster depends on the successful transmission and reception of beacon packets.

When an ordinary node cannot receive beacon packets from any of the clusterheads in its receive range, it continues to operate normally for the next superframe until it fails to receive a beacon packet for a second time, sequentially. At the beginning of the next superframe after missing two beacon packets, the node goes into the startup state that will lead to the formation of a new cluster. At this point the node picks a random time to transmit its own beacon packet to contend to become a clusterhead, and it begins to listen to the channel. If another node's beacon is heard in this period, then the node just stops its timer and starts normal operation. Otherwise, when the timer expires, the node sends a beacon and assumes the clusterhead position. Furthermore, when two clusterheads enter in each other's receive range, the one that receives the other's beacon first resigns. These mechanisms ensure the continuity of the protocol in the face of clusterhead failure due to node failure or the clusterhead moving out of range of other nodes in the cluster. However, the probability for a node to miss three consecutive beacons due to channel errors and overtake the clusterhead position when the clusterhead is actually still available is very low.

Other than clusterhead stability, missing beacon packets also slightly increases the average packet delay when two successive beacons are missed. Increase in delay is mainly due to the fact that the node loses time when it enters into startup mode. On the other hand, missing the beacon once does not affect the node at all, provided that the header is successfully received. Since the header contains all the required information about the data transmission schedule, losing beacon packets actually has very little effect on throughput.

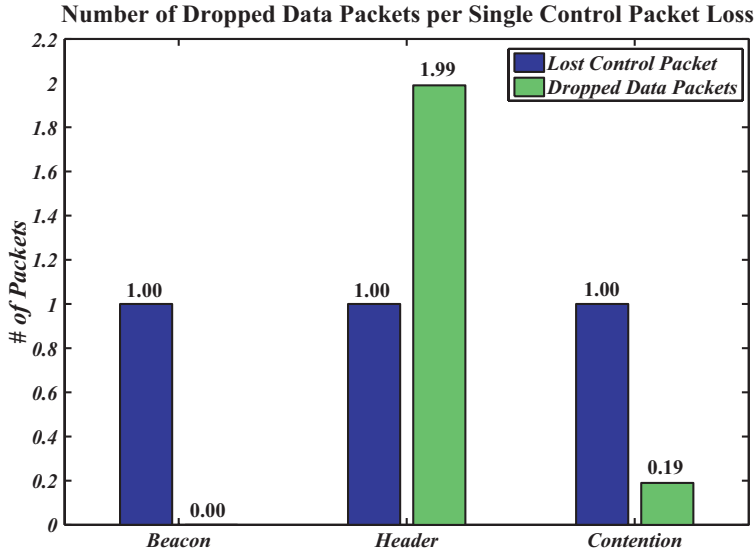


Figure 3.2: MH-TRACE performance degradation in terms of dropped data packets for beacon, header, and contention packet losses.

We ran simulations using uniform packet error probabilities for beacon reception (*i.e.,* only beacon packets experience channel errors and all the other packets are error free). Although this error model is rather simplistic, it is sufficient for our purposes in this section. We utilized 1.0%, 3.0%, and 5.0% packet error probabilities. Note that a 5.0% packet error probability represents a harsh environment [63]. However, the simulation results reveal that the performance of MH-TRACE does not deteriorate significantly. Indeed, loosing 10,000 beacon packets resulted in only 15 dropped data packets. In Figure 3.2 the effects of losing beacon packets are presented in normalized form. For example, for each dropped beacon packet, there are only 0.0015 dropped data packets, on average. Note that the data packet drop rate is marked as 0.00 in Figure 3.2 for beacon packet error since the actual value, 0.0015, is so small.

3.2.2 Header

Header packets are sent by the clusterheads to announce the data transmission schedule of the current frame. The transmission schedule is a list of nodes that have data slots reserved in the current frame, along with their data slot numbers. Therefore, missing a

header packet puts a node on hold if it has data to send. Since in voice communications delay bounds are stringent, if the waiting time for a voice packet exceeds a certain threshold it is dropped (*i.e.,* 50.0 ms in this study).

In order to support QoS, once a node obtains a slot its reservation is renewed automatically by the clusterhead as long as the node continues to transmit data. However, when a particular node misses a header packet, it cannot transmit during its reserved data slot and its reservation is cancelled by the clusterhead. Furthermore, a node that misses its reserved data slot still thinks that it has a reserved data slot. Thus, during the next frame when the node does not hear its ID in the schedule, it understands that it needs to contend again in the following superframe. This chain reaction increases the packet delay as much as three superframe times. Since time spent to get a data slot results in increased delay, data packets are dropped when they exceed the packet drop threshold, which is approximately twice the superframe time. Therefore, missing a single header results in at least one, and most probably two, dropped data packets.

In Figure 3.2 the simulation results obtained with the header packet error model are presented. As expected, the average number of dropped data packets per missed header packet is close to two.

3.2.3 Contention

Each node contends for channel access when it has data to send but did not reserve a data slot in the previous cyclic superframe. A node randomly chooses a sub-slot to transmit its request. If the contention is successful (*i.e.,* no collisions or error occurred), the clusterhead grants a data slot to the contending node if there are available data slots. If the node ID is not in the schedule, which is embedded in the header packet, the node understands that its contention was unsuccessful and waits for the next contention period.

We utilized CBR traffic for the evaluation of the impact of channel errors on the beacon and header packets. However, we use a statistical model of a voice activity detection based voice codec to evaluate the impact of channel errors on the contention packets due to the fact that the contention packets are not like beacon or header packets (*i.e.,* contention packets are event triggered rather than periodic). Once a node gets channel access, it will not loose it and will not be using contention packets for the rest

of the simulation time. Thus, to prevent this situation, which does not let us monitor the effects of packet losses on bursty traffic, we employed a statistical voice source model. According to the voice source model, speech is classified into "spurts" and "gaps" (*i.e.,* gaps are the silent moments during a conversation). During gaps, no data packets are generated, and during spurts, data packets are generated at 32 Kbps data rate. Both spurts and gaps are exponentially distributed statistically independent random variables, with means 1.0 s and 1.35 s, respectively [59].

Losing contention packets introduces additional delays into the network and causes data packets with critical delay values to be dropped. However the impact in this case is not as large as for errors in the header packets. Depending on traffic load and node density of the network, the results provided in Figure 3.2 are subject to change. As the traffic load and node density increase, average delay in the system also increases. Therefore, delay caused by losing contention packets results in a higher probability of dropped data packet. The next section addresses the design of a realistic wireless channel model, which we will be employing in our detailed simulations.

3.3 A Realistic Wireless Channel Model

Modeling channel errors in wireless communications has been extensively studied in the literature [24] [32]. Modeling the channel error as a Markov Process (MP) is a popular and realistic approach in wireless network simulations and has been verified by experimental data. We want to evaluate the performance of the MAC protocols; thus, the scenario we employ is single hop data broadcasting, which does not require a routing protocol on top of the MAC protocol. Furthermore, in single-hop broadcasting the overall performance (*e.g.,* QoS, energy dissipation) is directly determined by the performance of the MAC protocol.

In this study, we model the wireless channel as a first order Markov process with two states [68]. This model has been frequently employed in the literature to model fading errors at higher layers to calculate the average block error rate (packet error rate) [24–29, 31, 32] . In the literature, most of the channel error models are based on the assumption that data packet transmissions are independent and identically distributed (i.i.d.) In addition, many coding schemes and protocols were initially designed for i.i.d.

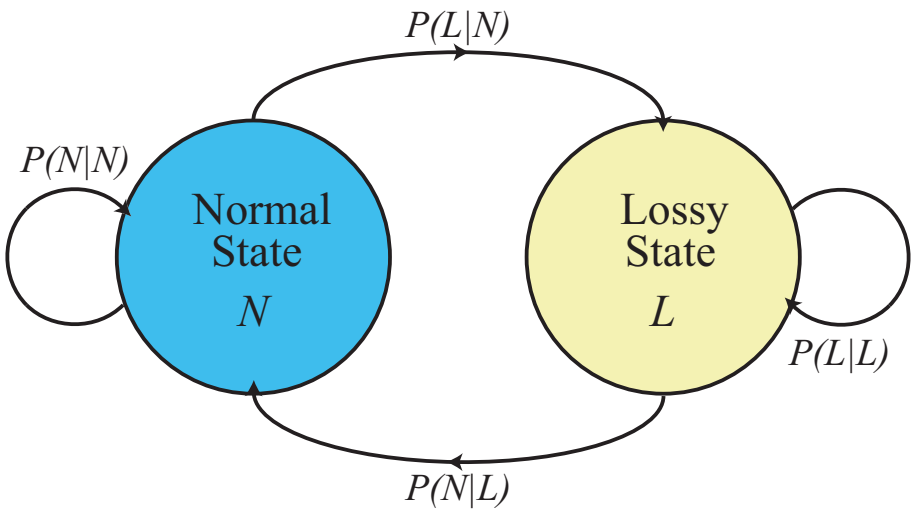


Figure 3.3: Gilbert-Elliot channel model.

channels. It has been shown that the special structure of Markov approximation makes it naturally useful and tractable for this purpose.

The Markov chain assumes that an adequate description of a system is given by a finite number of states. Each state is assigned a probability of the system being in that state. For example, the typical movement of the stock market could be considered as a simple two-state model in terms of up and down movement of the index. The study of Markov approximation for fading channels dates back to the early work of Gilbert [24] and Elliott [25], who built a two-state Markov channel known as the Gilbert-Elliott channel. In a simplified Gilbert model [26], the error probabilities in "bad" and "good" states are 1 and 0, respectively. Assuming 1 and 0 denote successful and erroneous transmission in a given slot, the state transition diagram is shown in Figure 3.3. The Normal state represents a perfect channel in which there is no error present (*i.e.*, probability of error is zero), whereas the Lossy state represents a wireless channel in which no packet can be delivered without error (*i.e.*, probability of error is one). The channel statistics are controlled by a set of transition probabilities that determine the individual probabilities P(N) and P(L).

The entropy rate of discrete finite random variables $(X_1, X_2, ..., X_N)$ is defined as:

$$H_0 = \lim_{N \to \infty} \frac{H(X_1, X_2, ..., X_N)}{N}. \tag{3.1}$$

If the random variables are stationary, we have

$$\lim_{N\to\infty} \frac{H(X_1, X_2, ..., X_N)}{N} = \lim_{N\to\infty} H(X_N|X_1, ..., X_{N-1}), \tag{3.2}$$

and in case of stationary Markov sequence we have

$$H_0 = \lim_{N\to\infty} H(X_N|X_{N-1}) = H(X_2|X_1). \tag{3.3}$$

For a stationary Markov sequence [68], the net probability flow between the two states is zero once the stationary distribution has been reached (*i.e.,* entropy of a state is constant at equilibrium). In our case we have four equations, Equations (3.4)-(3.7) with six unknowns, and therefore, it is possible to assign the desired stationary probabilities to both states $\{P(N), P(L)\}$ and calculate the transition probabilities $\{P(N|L), P(L|N)\}$ accordingly.

$$P(N) \cdot P(L|N) = P(L) \cdot P(N|L), \tag{3.4}$$

$$P(N) + P(L) = 1, \tag{3.5}$$

$$P(N|N) + P(N|L) = 1, \tag{3.6}$$

$$P(L|L) + P(L|N) = 1. \tag{3.7}$$

Using the two state model, one can generate a simulated wireless channel behavior for the protocol under study and perform realistic simulations. In Figure 3.4, the state transition behavior of the Gilbert model is illustrated. The channel spends some percentage of the total simulation time, determined by , in the normal state and some time determined by in the lossy state . Moreover, errors occur in bursts due to the fact that the channel spends portions of time in both states. When the channel is in the lossy state, errors are introduced according to the length of the packet: the probability of error for a longer packet is higher than the probability of error for a shorter packet [69] [37]. In other words, it is more likely for a data packet to be in error than for a beacon packet, which is the shortest packet in MH-TRACE, to be in error.

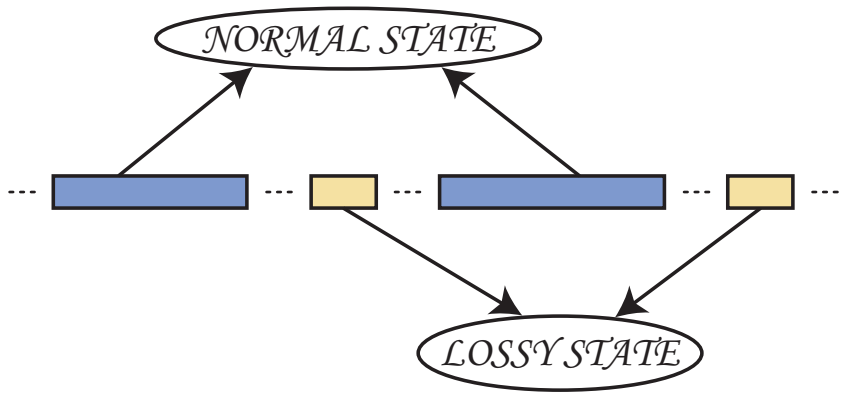


Figure 3.4: State transition with the Gilbert-Elliot channel model.

For the rest of the simulations the Gilbert model is employed with transition probabilities and state probability distribution shown in Table 3.1. The state probabilities P(N) and P(L) can be set according to the requirements by using Equations (3.4)-(3.7). We employ the model proposed according to the empirical data, which is obtained by observing the data communications in a Global System for Mobile communications (GSM) network [30].

3.4 Evaluation of IEEE 802.11 and MH-TRACE Using the Gilbert Channel

In this section we present the simulation results and analysis of IEEE 802.11 and MH-TRACE for a 1 km by 1 km area network with 80-200 nodes using the parameters in Table I. In the first set of simulations, we evaluate the protocols' performance using a static network, while in the second set of simulations, we use a Random Way-Point mobility model with node speeds chosen from a uniform distribution between 0.0 m/s

Table 3.1: Gilbert-Elliott channel model statistics.

State	**P(State)**	**P(L|State)**	**P(N|State)**
N (Normal)	0.9375	0.01	0.99
L (Lossy)	0.0625	0.85	0.15

Table 3.2: Simulation Setup

PARAMETER	SET 1	SET 2
Protocol	MH-TRACE/ IEEE 802.11	MH-TRACE/ IEEE 802.11
Number of Nodes	80-200	80-200
Simulation Time	200s	200s
Number of Repetition	5	5
Error Model	None/Gilbert	None/Gilbert
Node Mobility	Stationary	Mobile

and 5.0 m/s (the average pace of a marathon runner) with zero pause time. We used the statistical voice source model described in Section 3.2.3. The simulations are repeated with the same parameters five times, and the data points in the figures are the average of the ensemble and the error bars are the standard deviation of the ensemble. Table 3.2 presents the key simulation settings for the two sets of simulations.

Beacon, CA, contention, and IS packets are all 4 bytes. The header packet has a variable length of 4-18 bytes, consisting of 4 bytes of packet header and 2 bytes of data for each node to be scheduled. Data packets are 104 bytes long, consisting of 4 bytes of packet header and 100 bytes of data. Each slot or sub-slot includes 16 sec of guard band (*IFS*) to account for switching and round-trip time.

3.4.1 SET1: Network with Stationary Nodes

In this section the throughput of the two protocols (IEEE 802.11 and MH-TRACE) are compared for a network of stationary nodes. The protocols are simulated using both a perfect and a lossy channel whose statistics are given in Table 3.1 (*i.e.,* the channel spends 6.25% of the time in the lossy state and rest of the time, 93.75%, the channel is in the normal state). The parameters of the channel model are taken from the empirical data presented in [30]. The ratio of the time spent in the lossy and normal states affects the severity of the channel. For example, in a two node scenario, where one node continuously transmits data packets and the other listens, the packet loss ratio is directly determined by the percentage of time spent in the lossy state (*e.g.,* 6.25% of the packets

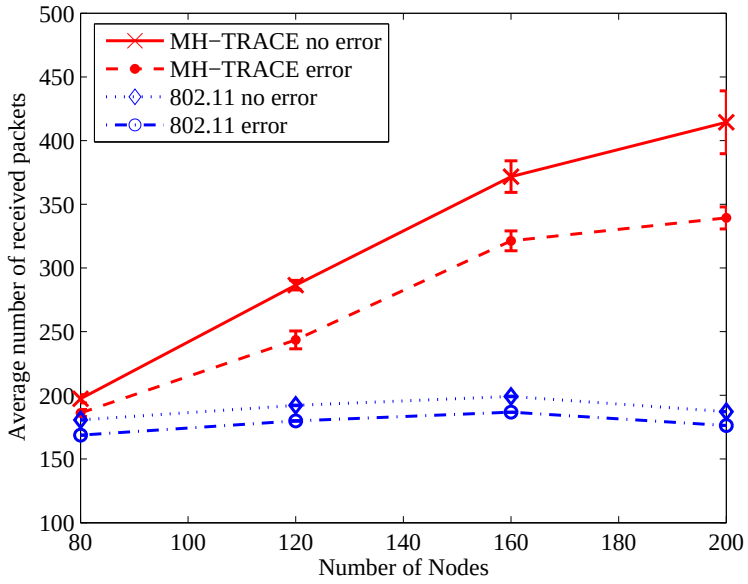


Figure 3.5: SET 1 (stationary nodes): Average number of received packets per node per second versus number of nodes.

are lost with a channel that spends 6.25% of the time in the lossy mode, given that the probability of data packet loss in the lossy state is close to unity). Moreover, the probability that a packet is in error also depends on the length of the packet (*i.e.,* short packets such as beacon and contention packets have smaller probabilities of being in error than longer packets such as data packets). In the lossy state, the probability of dropping a data packet is unity, the probability of dropping a header packet is 0.18, and the rest of the control packets have a 0.04 probability of error.

Figure 3.5 presents the average number of received packets per node per second versus the number of nodes. The curves labeled as “no error” are associated with the perfect channel scenario, whereas the ones labeled as “error” are obtained using the lossy channel model. When there are only 80 nodes with a perfect channel, the throughput of MH-TRACE (197.26 3.26 packets/node/s) is 9.16% more than that of 802.11 (180.71 0.01 packets/node/s). As the number of nodes increases, the difference between the two protocols increases. For 200 nodes MH-TRACE outperforms 802.11 by a factor of 2.2 in terms of received packets per node per second. IEEE 802.11 throughput is lower than MH-TRACE throughput due to excessive collisions experienced by IEEE 802.11.

For the lossy channel, the throughput of both protocols deteriorates in comparison to the performance with a lossless channel, but the loss in MH-TRACE is larger than the loss in 802.11. When the number of nodes is 80, both protocols loose approximately 6.3% of their performance under perfect channel conditions. Note that the channel stays in the lossy state approximately 6.3% of the total time, and all the data packets are dropped during the lossy state. On the other hand, control packets are dropped with less probability since packet loss probability is directly related with the packet length. With an increase in the number of nodes, both control and data traffic of MH-TRACE become heavier and the number of dropped packets increase. Therefore, channel errors degrade the performance by nearly 18.1%, whereas throughput loss in 802.11 remains at the same level (6.3%). Thus, approximately 12% of the packet losses in MH-TRACE are due to the loss of control packets. The dominant factor is the header packet losses. Despite the fact that MH-TRACE throughput is reduced from 414 to 339 packets/node/s, it still performs 92.5% better than 802.11.

3.4.2 SET2: Network with Mobile Nodes

In this set of simulations we evaluate the protocols' performance in terms of throughput, stability, data packet delay, and energy dissipation for a network of mobile nodes. Figure 3.6 shows the throughput of 802.11 and MH-TRACE as a function of the number of nodes in the network. Due to node mobility, the throughput of both protocols is reduced slightly, both in the lossless and lossy channel conditions. However, the general trend seen in the static case is preserved.

Figure 3.7 shows the average clusterhead lifetime (a) and the average number of clusterheads (b) over all simulation time, respectively. These metrics are related to the stability of the clustering algorithm of MH-TRACE. Although there are small differences between the lossless and lossy cases in the behavior of the clustering structure, these differences are not significant. Thus, channel errors do not affect the clustering algorithm of MH-TRACE significantly; rather the channel access mechanism is most affected from the channel errors, which manifests itself by a slight reduction in the throughput in excess of the IEEE 802.11 throughput reduction. The discussion on the effects of control packet losses in Section 3.2, in particular, the impact of beacon and header packet losses, supports this argument. Header packet losses, which are related

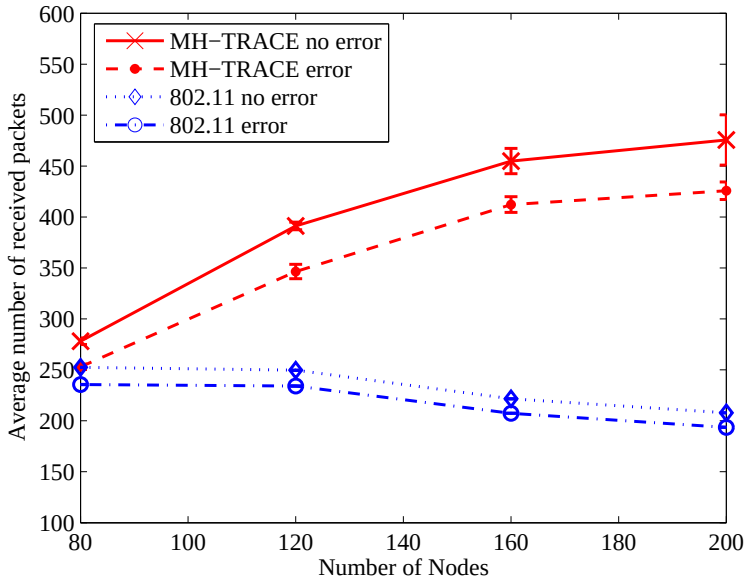
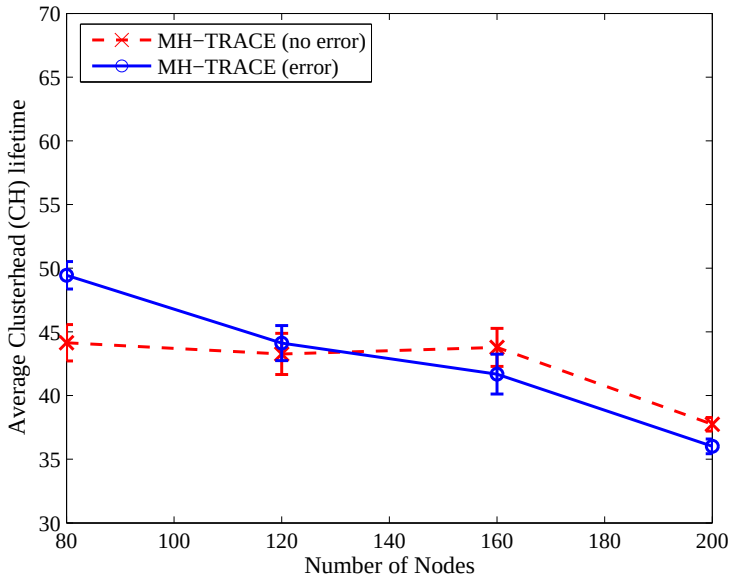


Figure 3.6: SET 2 (mobile nodes): Average number of received packets per node per second versus number of nodes.

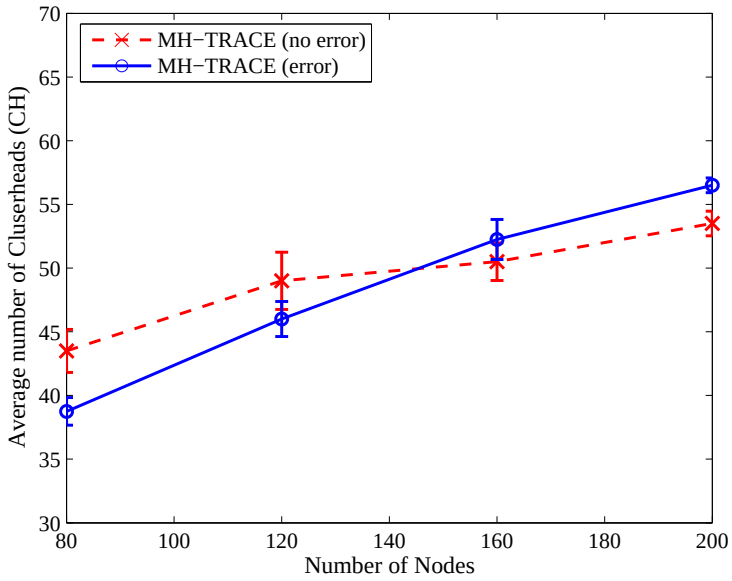
with channel access, cause more data packet losses than beacon packet losses, which are related with the clustering mechanism.

One of the most important advantages of MH-TRACE over IEEE 802.11 is its better energy efficiency. The average energy dissipation of MH-TRACE and IEEE 802.11 with lossless and lossy channel conditions as a function of node density is presented in Figure 3.8. Energy dissipations of both protocols are insensitive to the channel conditions in the application scenario we considered. Nevertheless, MH-TRACE energy dissipation stays less than 40% of the energy dissipation of IEEE 802.11 energy dissipation.

Figure 3.9 shows the average data packet delay for MH-TRACE and IEEE 802.11 as a function of node density. Note that the maximum packet delay is 50 ms, which is dictated by the application layer. MH-TRACE packet delay is higher than IEEE 802.11 packet delay in both lossless and lossy channel conditions due to the fact that in MH-TRACE nodes can have channel access only once in a superframe time, whereas in IEEE 802.11 channel access is not restricted. Both MH-TRACE and IEEE 802.11 have comparatively higher packet delays in a lossy channel. The increase in the packet delay



(a) SET 2 (mobile nodes): Average life time of clusterheads versus number of nodes.



(b) SET 2 (mobile nodes): Average number of clusterheads versus number of nodes.

Figure 3.7: Clusterhead stability versus number of nodes.

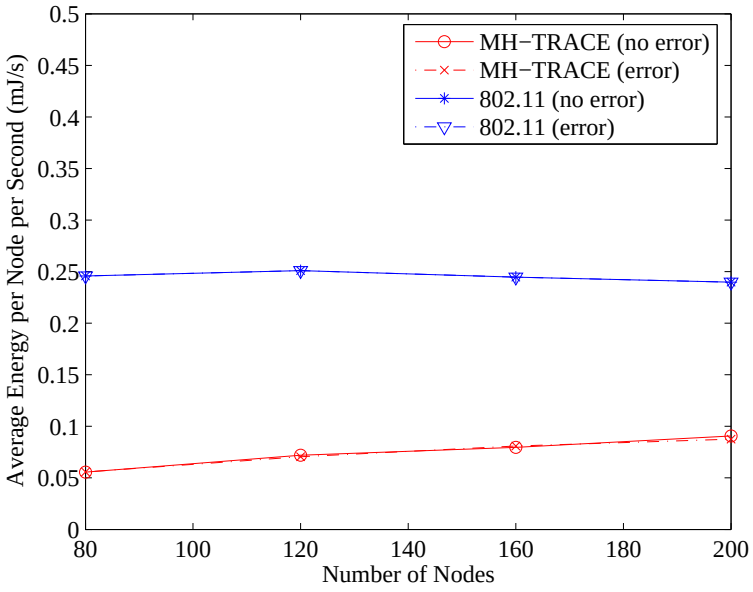


Figure 3.8: SET 2 (mobile nodes): Average energy consumption per node per second versus number of nodes.

in MH-TRACE is mainly due to the header packet losses, as once a node loses a header packet, it loses several frame times before regaining channel access.

3.5 Analytical Model

In this section, we develop an analytical model to estimate the performance of MH-TRACE as a function of BER. However, our model essentially models a generic coordinated MAC protocol, thus, it is not necessarily specific to MH-TRACE and it can easily be extended to analyze any coordinated MAC protocol with little modification (*e.g.,* IEEE 802.15.3 [2] and EC-MAC [70]). For example, IEEE 802.15.3 has a similar channel access mechanism to MH-TRACE, where time is organized into cyclic superframes and channel access is granted through a control packet that includes the schedule (*i.e.,* a beacon packet). Therefore, modeling the performance of IEEE 802.15.3 will be essentially the same as our modeling of MH-TRACE. In our analysis we do not consider any error correction scheme, thus, if there is at least one bit error within a packet,

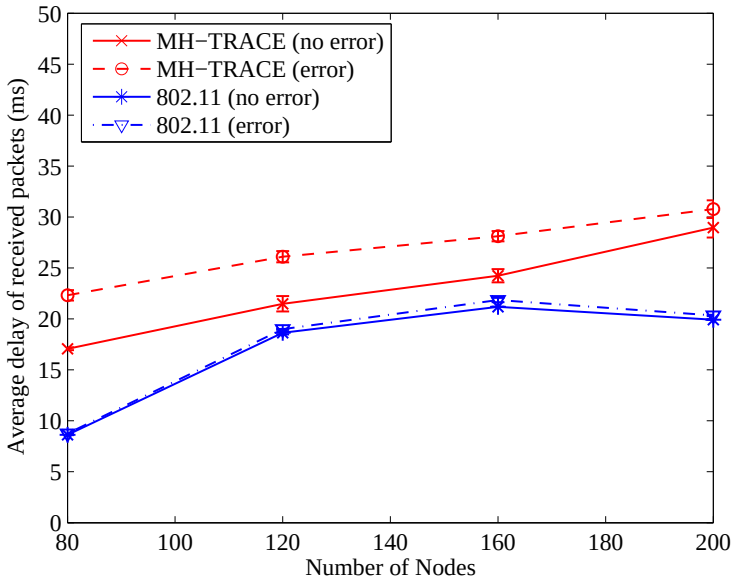


Figure 3.9: SET 2 (mobile nodes): Average data packet delay versus number of nodes.

then that packet is discarded. Random packet errors are independently introduced at the receivers.

If a protocol cannot maintain the desired level of performance, then its energy efficiency becomes meaningless. Thus, in order to achieve meaningful energy efficiency, it is absolutely necessary to make sure that a protocol does not deteriorate system performance while saving energy. MH-TRACE is a protocol designed primarily for energy efficiency, and it is obvious that under ideal channel conditions its energy efficiency will be superior to any non-coordinated protocol. However, the question is whether MH-TRACE preserves its performance in the face of channel errors. In this section we seek the answer to this question through analysis supported with simulations.

3.5.1 Basic Model

To demonstrate our approach clearly and with a simple example, first we consider a fully-connected network with a small number of static nodes. The number of data slots in one superframe is high enough to support all of the nodes in the network (see Table 3.3).

Table 3.3: Simulation parameters

Acronym	Description	Value
T_{SF}	Superframe duration	25.172 ms
T_F	Frame duration	3.596 ms
N_F	Number of frames	7
N_{DS}	Number of data slots per frame	7
N_C	Number of contention slots per frame	6
T_B	Beacon slot duration	32 μs
T_{CA}	CA slot duration	32 μs
T_C	Contention sub-slot duration	32 μs
T_H	Header slot duration	92 μs
T_{IS}	IS sub-slot duration	32 μs
T_D	Data slot duration	432 μs
N/A	Data packet size	104 B
N/A	Header packet size	4-18 B
N/A	All other control packet size	4 B
IFS	Inter-frame space	16 s
T_{drop}	Packet drop threshold	50 ms
T_{VF}	Voice packet generation period	25.172 ms
P_T	Transmit power	0.6 W
P_R	Receive power	0.3 W
P_I	Idle power	0.1 W
P_S	Sleep power	0.0 W
D_{Tr}	Transmission range	250 m
D_{CS}	Carrier Sense range	507 m

When there are no channel errors, all nodes should be able to transmit and receive without any packet drops or collisions. There will be only one clusterhead in the network due to the fact that there cannot be two clusterheads that can hear each other directly.

The number of data packets generated per node per second, (DP_{node}), is equal to the packet rate (R_{packet}) of MH-TRACE (*i.e.,* one packet per superframe time ($1/T_{sf}$)).

$$DP_{node} = R_{packet} = \frac{1}{T_{sf}} \tag{3.8}$$

DP_{node} represents the number of data packets generated by a single node in the network and can be regarded as the maximum number of packets a node can transmit given that it has full access to a perfect channel whenever it needs. However, a lossy channel will cause packet drops and therefore the throughput of the network will drop accordingly.

In Figure 3.2, the corresponding throughput losses due to dropped beacon, header and contention packets are given to illustrate the impact of the particular control packet on overall protocol performance. In these results only the specified control packets are lost due to channel errors and all the other packets are not affected [71].

These results are from Section 3.2 where we simulated a six node fully connected static network to clearly observe the effects of packet losses. When there are no channel errors, all nodes should be able to transmit and receive without any packet drops or collisions. There will be only one clusterhead in the network due to the fact that there cannot be two clusterheads that can hear each other directly. We utilized 1.0%, 3.0%, and 5.0% packet error probabilities. Note that a 5.0% packet error probability represents a harsh environment [63]. We used the *ns*-2 simulator to evaluate the system performance.

As can be seen from the figure, a lost header packet has the most impact on the performance of MH-TRACE. Loss of contention packets cause 10 times less impact on throughput than loss of header packets (0.19). Finally, for each beacon packet dropped, only 0.0015 data packets are dropped. Like beacon packet losses, losing other control packets (*e.g.,* IS, CA) do not significantly affect the throughput of the network. Thus, we conclude that the header and contention packets are the only control packets whose loss due to channel noise affect the network performance.

Therefore, we can write the equation for the *transmit* throughput of a single node (*i.e.,* transmit throughput per node per second T_{node}) in terms of the data packets dropped

before transmission due to lost header packets (DPL_H) and contention (DPL_C) packets:

$$T_{node} = DP_{node} - DPL_H - DPL_C \tag{3.9}$$

Both (DPL_H) and (DPL_C) can be written as the product of three parts. (*i*) Number of data packets dropped per header or contention packet loss (DPL_{perH} or DPL_{perC}). (*ii*) Number of header or contention packets sent to a node or clusterhead per second (HP_{node} or CP_{node}). (*iii*) Probability of dropping a header or contention packet ($P_H or P_C$).

As contention packets are relatively short (4 bytes), they are less likely to be dropped than header packets (16 bytes for 6 broadcasting nodes). Furthermore, since the sources are constant bit rate (CBR) and MH-TRACE utilizes automatic channel access renewal, once a node gets channel access, it will not loose it and, thus, will not need to transmit contention packets for the rest of the simulation time. Moreover, the number of dropped data packets per lost header packet is 10 times larger than the number of dropped data packets per lost contention packet, as shown in Figure 3.2. Therefore, it is reasonable to assume that the effect of losing contention packets can be neglected. Furthermore, by ignoring the control packets other than the header packet, we focus on a more general model rather than an MH-TRACE specific model. Based on these assumptions, the transmit throughput per node per second becomes:

$$T_{node} = DP_{node} - DPL_H \tag{3.10}$$

$$T_{node} = \frac{1}{T_{sf}} - DPL_{perH} \times HP_{node} \times P_H. \tag{3.11}$$

In Equation (3.11), DPL_{perH} is a constant (1.99) and HP_{node} is equal to DP_{node} since each node receives one header per super frame from its clusterhead. Finally P_H depends on the length of the header packet L_H and is calculated from the Bit Error Rate (BER) of the channel.

$$P_H = \{1 - (1 - BER)^{L_H}\}. \tag{3.12}$$

Therefore,

$$T_{node} = \frac{1}{T_{sf}} - 1.99 \times \frac{1}{T_{sf}} \times \{1 - (1 - BER)^{L_H}\} \tag{3.13}$$

$$T_{node} = \frac{1}{T_{sf}} \times [1 - 1.99 \times \{1 - (1 - BER)^{L_H}\}]. \tag{3.14}$$

In order to get the number of received packets per second in the network, we need to multiply the transmit throughput per node per second with the number of neighboring nodes $N-1$ (note that all the nodes can hear each other in this network). Moreover, each data packet is received with a probability P_D, which is the probability that a data packet (with length $L_D = 104$ bytes) goes through the channel with no error at a given BER. Accordingly, the receive throughput per node per second (T) becomes:

$$T = (N-1) \times T_{node} \times (1 - BER)^{L_D}. \tag{3.15}$$

Note that the receive throughput per node per second of IEEE 802.11 is simply equal to $\frac{N-1}{T_{sf}} \times (1 - BER)^{L_D}$ since in CSMA-type protocols such IEEE 802.11 in broadcasting mode, only data packets are sent through the lossy channel and the throughput is determined by the BER of the channel and length of a data packet.

We used the ns-2 simulator to validate the analytical model. The channel rate is set to 2 Mbps, and all nodes have a CBR (Constant Bit Rate) data source with 32 Kbps data rate, which corresponds to one voice packet per superframe. The simulations are run for 1000 s and repeated with the same parameters five times.

In Figure 3.10, the analytical model for MH-TRACE and IEEE 802.11 are plotted against increasing BER. Also, the simulation results are included for both protocols to demonstrate the accuracy of the models. The throughput of MH-TRACE drops by almost 50% at a BER around 7×10^{-4}. On the other hand, IEEE 802.11 retains almost 55% of its initial throughput at the same BER (note that the initial throughputs of both protocols are the same). This difference can be translated into the fact that IEEE 802.11 performs 10% better than MH-TRACE, which experiences a worse performance degradation due to lost coordination packets [71].

These results show that the analytical model proposed to estimate the throughput of MH-TRACE is quite accurate. The model captures the fact that coordinated MAC protocols are more vulnerable than non-coordinated MAC protocols to channel noise due to their dependence on the robustness of the control traffic. In Figure 3.10, MH-TRACE experiences a steeper loss than IEEE 802.11 for BER values greater than 10^{-4}, which is the point where header packet losses become the dominant factor in performance degradation. Our model captures this unique behavior of MH-TRACE. Therefore, our first conclusion is that the increased throughput loss occurs when a coordinated MAC

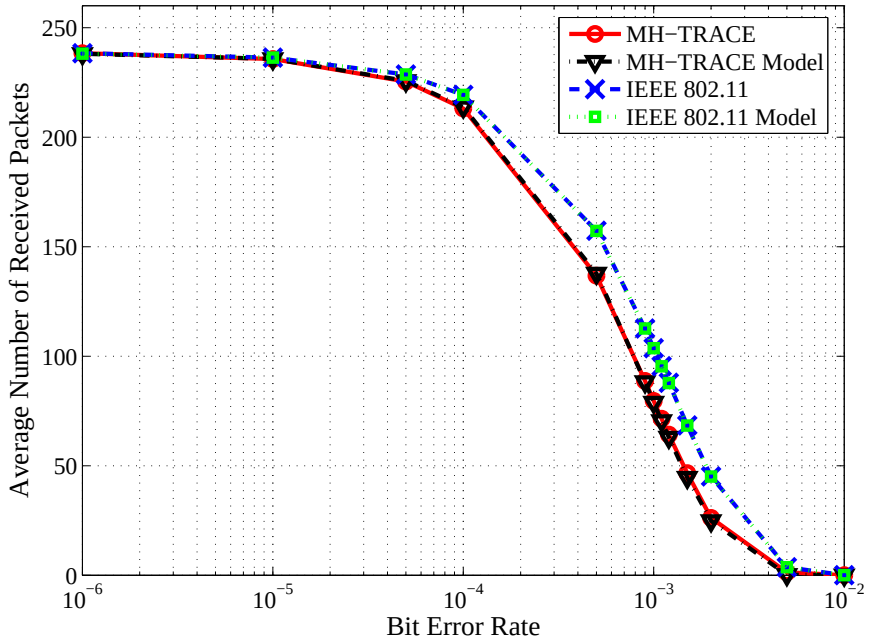


Figure 3.10: Average number of received packets per node per second versus bit error rate (BER).

protocol starts to lose its control packets. In our case, header packets are lost first since they are the longest control packet in MH-TRACE (see Table 3.3).

Before starting to derive a general model for MH-TRACE throughput, we want to mention that in our model, we treated the clusterhead as a regular node inside the network, but in reality, a clusterhead would not drop any data packets due to lost header packets since the clusterhead is the one generating the header packets. Therefore, our model slightly underestimates the throughput of MH-TRACE by treating the clusterhead as an ordinary node.

3.5.2 General Model

In this section, we consider a rectangular field ($L \times H$) in which a certain number of nodes (N), which have a communication radius (r), are randomly deployed. We use a statistical model of Voice Activity Detector (VAD) equipped voice source model that classifies speech into *spurts* and *gaps* (*i.e.,* gaps are the silent moments during

a conversation). During gaps, no data packets are generated, and during spurts, data packets are generated at 32 Kbps data rate. Both spurts and gaps are exponentially distributed statistically independent random variables, with means $\eta_s = 1.0s$ and $\eta_g = 1.35s$, respectively [59]. The reason for using such a statistical voice source is that the transmission schedule will change frequently (*i.e.,* at the end of spurts nodes cease transmitting and their granted data slot will be taken away from them, and they will need to contend for channel access at the beginning of the next spurt), even in the absence of the channel errors, which is necessary to asses the system performance for a coordinated MAC protocol in the face of a dynamically changing transmission schedule.

Our approach to this more complex model will be basically the same as before. We begin by calculating the transmit throughput per node per second (T_{node}) when the channel is perfect. In addition to Equation (3.14), we need a term that captures the effect of the voice source model. This term can easily be represented with the ratio of spurts to the whole conversation (η). Therefore, we can write T_{node} as in Equation (3.16).

$$\begin{aligned} T_{node} &= \frac{1}{T_{sf}}[1 - 1.99\{1 - (1 - BER)^{L_H}\}][\eta] \\ &= \frac{1}{T_{sf}}[1 - 1.99\{1 - (1 - BER)^{L_H}\}][\frac{\eta_s}{\eta_s + \eta_g}] \end{aligned} \tag{3.16}$$

After obtaining the expression for the transmit throughput per node per second, we have to find an expression for the average number of nodes within the communication range of a given node (*i.e.,* the average number of neighbors for a given node). In Figure 3.11, the rectangular field is partitioned into three different regions according to the coverage characteristic of a node in a particular region. For example, a node inside region 1 (*e.g.,* n_2) has its full coverage within the boundaries of the field. Therefore, any node inside region 1 utilizes 100% of its total coverage. Whereas nodes inside regions 2 and 3 (*e.g.,* n_1 and n_3) have a part of their coverage outside the field of interest and consequently the average percentage coverage for these nodes is less than 100%. Finding the percentage coverage for each region will lead us to the average number of neighbors.

We start the derivation of the percentage with region 2. In Figure 3.12 the approach we used for obtaining the percentage is given. The area of the piece of circle shaded in Figure 3.12 can be expressed as follows:

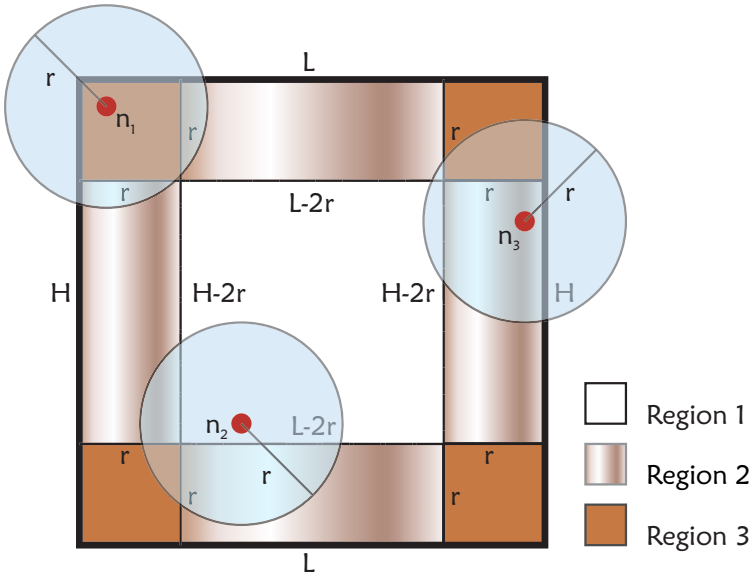


Figure 3.11: Rectangular field partitioned into three different regions.

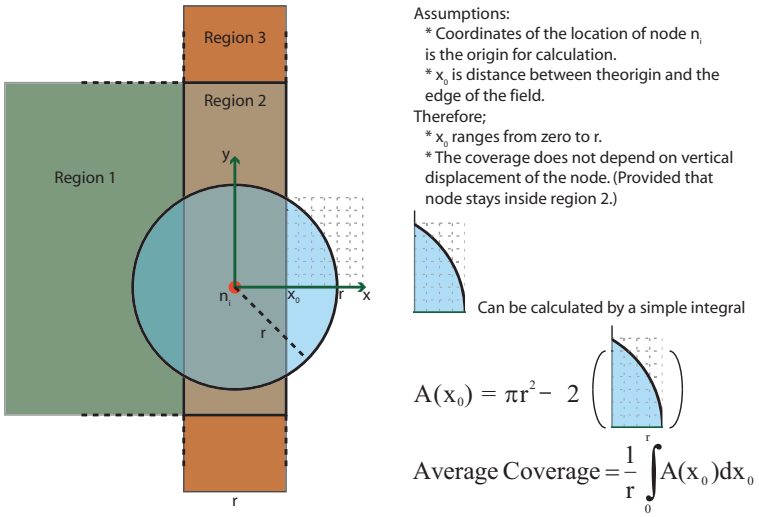


Figure 3.12: Calculation of the percentage coverage of a node inside region 2.

$$
\begin{aligned}
I &= \int_{x_0}^{r} \sqrt{r^2 - x^2}dx \\
&= \frac{\pi}{4}r^2 - \frac{x_0}{2}\sqrt{r^2 - x_0^2} - \frac{r^2}{2}\arcsin(\frac{x_0}{r}).
\end{aligned}
\tag{3.17}
$$

Thus, the average coverage for region 2 (α_2) becomes,

$$
\begin{aligned}
\alpha_2 &= \frac{1}{r}\int_0^r A(x_0)dx_0 \\
&= \frac{1}{r}\int_0^r \left(\pi r^2 - 2I(x_0)\right)dx_0 \\
&= \pi r^2 - \frac{2}{3}r^2.
\end{aligned}
\tag{3.18}
$$

After obtaining the average coverage as in Equation (3.18), we can easily calculate the percentage coverage of region 2 (σ_2).

$$
\sigma_2 = \frac{\alpha_2}{\pi r^2} \quad = 1 - \frac{2}{3\pi} \tag{3.19}
$$

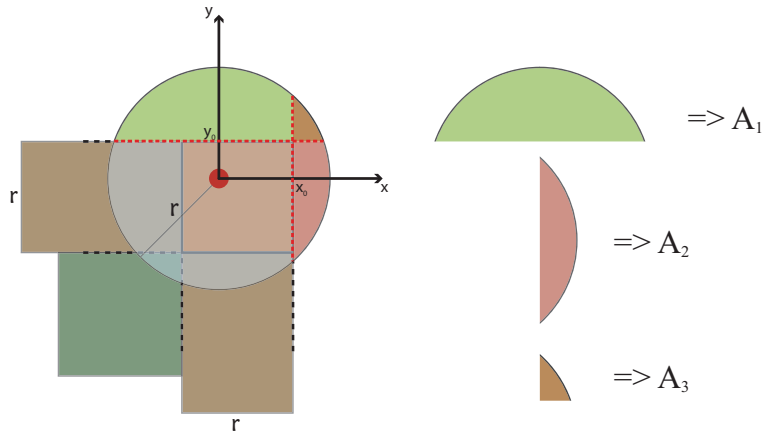


Figure 3.13: Calculation of the percentage coverage of a node inside region 3.

Next we derive the average coverage for region 3 (σ_3). The area in question is divided into three parts (see Figure 3.13). According to this partitioning we have $A = \pi r^2 - (A_1 + A_2 - A_3)$, which is the coverage for a node inside region 3. The integrals for A_1 and A_2 are the same as I given in Equation (3.17) and can be expressed as $2I(x_0)$ and $2I(y_0)$, respectively.

$$
\begin{aligned}
A_3 &= \int_{x_0}^{\sqrt{r^2-y_0^2}} \left(\sqrt{r^2-x^2} - y_0\right) dx \\
&= -\frac{y_0\sqrt{r^2-y_0^2}}{2} + \frac{r^2 \arcsin(\frac{\sqrt{r^2-y_0^2}}{r})}{2} \\
&\quad - \frac{x_0\sqrt{r^2-x_0^2}}{2} - \frac{r^2 \arcsin(\frac{x_0}{r})}{2} + y_0 x_0.
\end{aligned} \tag{3.20}
$$

After obtaining A_3, we can calculate the average coverage α_3 by taking the average of A.

$$
\begin{aligned}
\alpha_3 &= \frac{1}{r^2}\int_0^r \int_0^r A\left(x_0, y_0\right) dx_0 dy_0 \\
&= \pi r^2 - \frac{29}{24} r^2.
\end{aligned} \tag{3.21}
$$

Thus, σ_3 becomes:

$$
\sigma_3 = \frac{\alpha_3}{\pi r^2} \quad = 1 - \frac{29}{24\pi} \tag{3.22}
$$

This is the last percentage coverage we needed to calculate the overall percentage coverage (σ), or the average number of nodes within the range of a given node inside the rectangular field. Below we give the resulting σ in terms of the communication radius r, the length of the field L and the height of the field H.

$$
\sigma = \frac{\sigma_1(L-2r)(H-2r) + 2\sigma_2(H+L-4r)r + 4\sigma_3 r^2}{LH} \tag{3.23}
$$

This expression can be used to calculate the average number of neighboring nodes (N_N) for a node inside of a rectangular field by multiplying σ with πr^2 (*i.e.,* the coverage of a node with communication radius r) and the node density ($\frac{(N-1)}{LH}$). Note that there are $N-1$ nodes remaining that can be neighbors.

$$
N_N = \frac{(N-1)\sigma\pi r^2}{LH} \tag{3.24}
$$

Now, we can combine Equation (3.24) with Equation (3.16) to get the receive throughput per node per second, T.

$$
T = N_N \times T_{node} \times (1 - BER)^{L_D} \tag{3.25}
$$

According to our model, given that we have a constant simulation area and the same traffic model, throughput increases as the number of nodes in the network increases. In

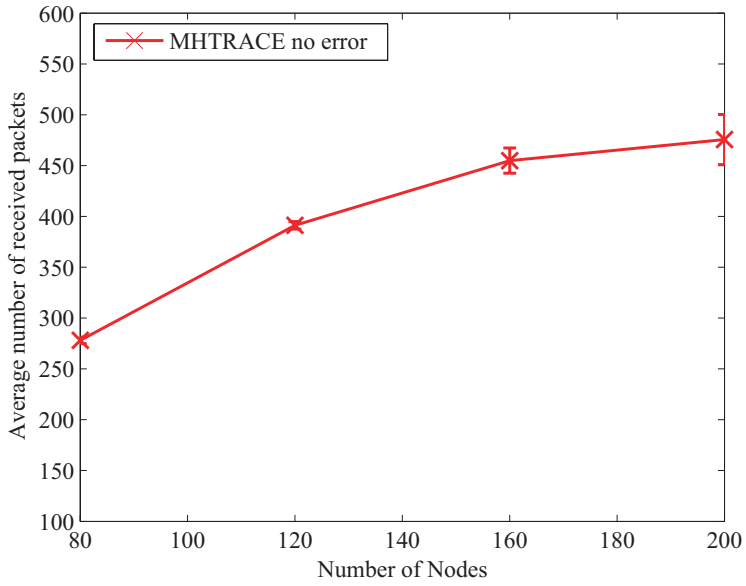


Figure 3.14: Average number of received packets per node per second versus number of nodes (mobile).

other words, the model suggests that throughput increases linearly with increasing node density. However, our previous work showed that throughput per node per second goes into saturation as the number of nodes in the network increases (see Figure 3.14). This trend is a result of packet collisions and drops emerging from mobility and increased contention for channel access [61]. According to this fact, we have to modify our initial throughput value (throughput when there is a perfect channel) in order to get a more accurate model for throughput. Since it is extremely challenging to model the dynamical behavior in Figure 3.14 analytically, the initial throughput values are calibrated according to feedback from simulation results.

3.6 Simulation Environment

To test the performance of MH-TRACE and IEEE 802.11 with increasing BER levels and to test the validity of our model, we ran simulations using the ns-2 network simulator [72]. We simulated conversational voice coded at 32 Kbps with VAD (see

Section 3.5.2), which corresponds to one voice packet per superframe. The channel rate is set to 2 Mbps and the standard IEEE 802.11 physical layer is employed for both protocols. All the simulations are run with 100 or 200 nodes, moving within a 1 km by 1 km area for 100 seconds according to the Random Way-Point (RWP) mobility model with node speeds chosen from a uniform distribution between 0.0 m/s and 5.0 m/s. In this work we use 5.0 m/s, which is the average pace of a marathon runner, as our upper limit; however, we have observed that the performance of single-hop broadcasting in MH-TRACE does not change with increased mobility. Pause time is set to zero to avoid any non-moving nodes throughout the simulations. The transport agent used in the simulations is User Datagram Protocol (UDP), which is a best effort service. The simulations are repeated with the same parameters six times, and the data points in the figures are the average of the ensemble. Acronyms, descriptions and values of the parameters used in the simulations are presented in Table 3.3 and Table 3.4.

Beacon, CA, contention, and IS packets are all 4 bytes. The header packet has a variable length of 4-18 bytes, consisting of 4 bytes of packet header and 2 bytes of data for each node to be scheduled. Data packets are 104 bytes long, consisting of 4 bytes of packet header and 100 bytes of data. Each slot or sub-slot includes 16 μs of guard band (IFS) to account for switching and round-trip time.

We used the standard energy and propagation models of ns-2 [72] without any modifications. Transmit power, P_T, consists of a constant transmit electronics part and a variable power amplifier part. The propagation model is a hybrid propagation model, which assumes Free-Space propagation for short distances and Two-Ray Ground propagation for long distances. In the simulations we used a constant transmit power, which results in a constant transmission range, D_{Tr}, of 250 m and constant carrier sense range, D_{CS}, of 507 m. Receive power, P_R, is dissipated entirely on receiver electronics. Idle power, P_I, is the power needed to run the electronic circuitry without any actual packet reception. In sleep mode, the radio is shut down so sleep mode power, P_S, is very low.

In this study, we want to evaluate the performance of the MAC protocols; thus, the scenario we employ is single-hop data broadcasting, which does not require a routing protocol on top of the MAC protocol. Furthermore, in single-hop broadcasting, the overall performance (*e.g.,* energy dissipation, QoS *etc.*) is directly determined by the performance of the MAC protocol.

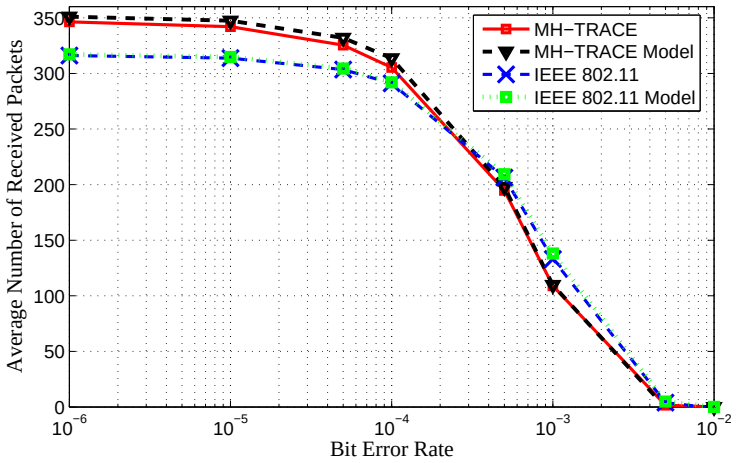
Table 3.4: Simulation Setup

PARAMETER	VALUE
Number of Nodes	100 & 200
Simulation Area	1000m x 1000m
Simulation Time	200s
Number of Repetitions	6

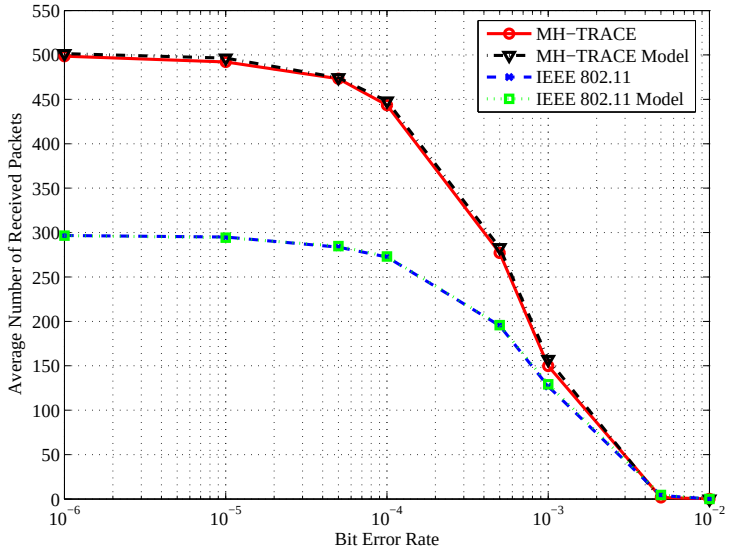
3.6.1 Throughput

Figure 3.15 (a)and (b) present the throughput of MH-TRACE and IEEE 802.11 obtained from analytical models and simulations as functions of BER with 100 nodes and 200 nodes, respectively. Throughput is defined as the average number of received bit error free data packets per node per second. The analytical model developed in Section 3.5.2 (Equation (3.25)) for MH-TRACE is in very good agreement with the simulation results presented in the figures. The model for IEEE 802.11 is obtained by using the probability of successful data packet transmission ($(1 - BER)^{L_D}$) and the initial throughput value, which also closely follows the simulation results for IEEE 802.11.

The difference between the initial throughput values of MH-TRACE, where the BER rate is too low to affect the throughput (*i.e.,* BER $< 10^{-4}$), for the 100-node network (see Figure 3.15 (a)) and the 200-node network (see Figure 3.15 (b)) is due to the fact that both the number of transmissions and receptions are directly proportional to the total number of nodes in the network; thus, when the number of nodes is doubled, in ideal conditions, total throughput should be quadrupled. Hence, the throughput per node should be doubled. However, non-idealities, such as packet drops, keeps the throughput less than the ideal value. IEEE 802.11 throughput for the 200-node network is lower than the 100-node network throughput because of a very high collision rate. Note that while IEEE 802.11 collision resolution mechanism (*i.e., p*-persistent CSMA in broadcasting) has a similar performance with MH-TRACE in the 100-node network, it becomes increasingly ineffective with the increasing node density (*i.e.,* IEEE 802.11 throughput is 60 % of MH-TRACE throughput in the 200-node network).



(a) (100 nodes): Average number of received packets per node per second versus bit error rate (BER).



(b) (200 nodes): Average number of received packets per node per second versus bit error rate (BER).

Figure 3.15: Average number of received packets versus BER.

There are two mechanisms that decrease the throughput of MH-TRACE with increasing BER: (i) with the increasing BER, more data packets are corrupted, which is also true for IEEE 802.11. Thus, the throughput decreases with increasing BER and (ii) the increase of the corrupted header packets results in unutilized data slots for MH-TRACE, whereas in IEEE 802.11 this is not a problem due to the lack of header packets. This situation creates an interesting tradeoff: while scheduling through header packets results in very high channel utilization in congested networks, it prevents nodes from channel utilization in high BER levels. However, when we examine the figures we see that MH-TRACE throughput is lower than IEEE 802.11 throughput only in low node density networks and only for extremely high BER levels (*i.e.,* the 100-node network and BER $\geq 10^{-3}$). Note that at BER $= 10^{-3}$ only 45 % of the data packets are non-corrupted, which is not an acceptable operating condition. For all other situations, MH-TRACE throughput performance is better than IEEE 802.11. Furthermore, in the 200-node network MH-TRACE throughput never drops below that of IEEE 802.11 throughput at any BER level. Thus, coordination through header packets is preferable over non-coordination regardless of the BER level of the network, especially in high congestion networks, from a throughput performance point of view.

3.6.2 Stability

Figure 3.16, presents the average clusterhead lifetime for the 100-node network and the 200-node network as a function of BER. Only the clusterheads that have a minimum lifetime of $10T_{SF}$ are counted in order to filter frequent clusterhead changes due to mobility and collisions so that only stable clusterheads are taken into consideration. Average clusterhead lifetime in the 100-node network is higher than the clusterhead lifetime in the 200-node network due the fact that the average number of clusterheads in a denser network is higher than the average number of clusterheads in a sparser network. This is because in sparse networks some areas are not covered by any clusterhead, and in fact, these areas are unpopulated by any node. On the other hand, in dense networks there are barely any uncovered areas. Thus, the total coverage of dense networks is higher, which can be made possible by higher number of clusterheads. A higher number of clusterheads in the same network topology (*i.e.,* 1 km by 1 km network) results in less inter-clusterhead separation on the average, which increases the chance of one cluster-

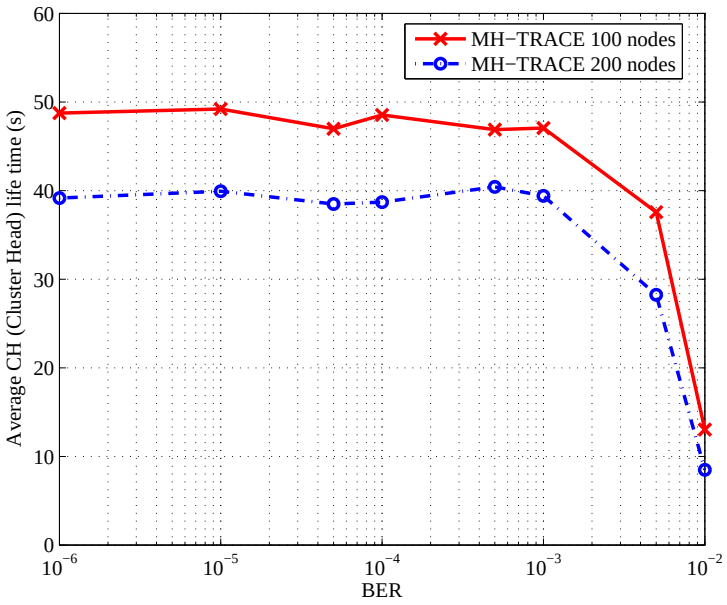


Figure 3.16: Average CH lifetime versus bit error rate (BER).

head moving into another's transmission range and resigning (*i.e.,* there cannot be any other clusterhead in the receive range of a clusterhead). Therefore, the average clusterhead lifetime in the 200-node network is lower than the average clusterhead lifetime in the 100-node network.

Clusterhead stability is not significantly affected by the BER level of the network for relatively low BER levels (*i.e.,* BER $\leq 10^{-3}$). This is because only 4 % of beacon packets are corrupted at BER $= 10^{-3}$, on the average. However, at BER $= 10^{-2}$, more than a quarter of the beacon packets are corrupted, which results in significantly shorter average clusterhead lifetime. Note that a node starts to contend for being a clusterhead if it does not receive a beacon packet for two consecutive superframes. Nevertheless, at BER $= 10^{-2}$, 99.98 % of the data packets are corrupted. Thus, maintaining a cluster structure is not a meaningful consideration at such high BER levels.

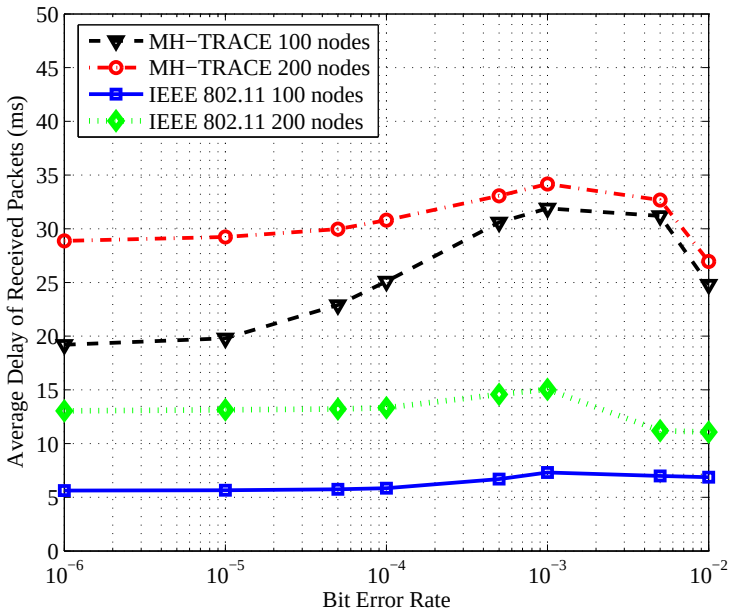


Figure 3.17: Average data packet delay versus bit error rate (BER).

3.6.3 Packet Delay

Figure 3.17 presents the average data packet delay for MH-TRACE and IEEE 802.11 as a function of BER. Data packets are dropped at the MAC layer if the data packet delay exceeds T_{drop}, which is 50 ms. MH-TRACE packet delay is higher than IEEE 802.11 packet delay at all BER levels in both the 100-node network and the 200-node network due to the fact that in MH-TRACE nodes can have channel access only once in a superframe time, whereas in IEEE 802.11 channel access is not restricted. MH-TRACE has comparatively higher packet delays as the BER level increases towards 10^{-3}. The increase in the packet delay in MH-TRACE is mainly due to the header packet losses, as once a node loses a header packet, it loses several frame times before regaining channel access. IEEE 802.11 packet delay is almost constant. Packet delay is not very informative for BER levels higher than 10^{-3}, because the throughput decreases to unacceptably low values. Average packet delay is higher in denser networks for both MH-TRACE and IEEE 802.11 due to the fact that higher channel utilization brings longer delays at the queue. The delay of MH-TRACE for both node densities tends to

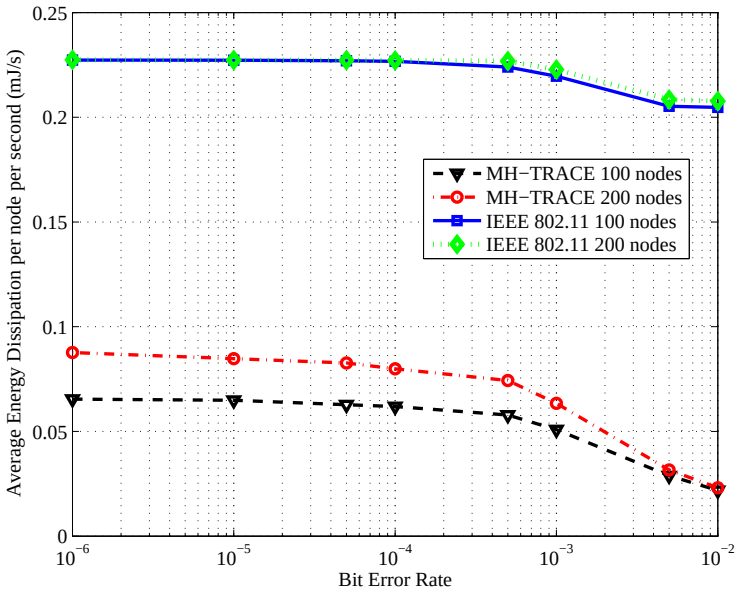


Figure 3.18: Average energy consumption per node per second versus versus bit error rate (BER).

converge to similar values when BER $> 10^{-3}$. Although it is less obvious, the same behavior is also present in IEEE 802.11 case.

3.6.4 Energy Dissipation

One of the most important advantages of MH-TRACE over IEEE 802.11 is its better energy efficiency. Average energy dissipation per node per second for MH-TRACE and IEEE 802.11 with 100 and 200 nodes as a function of BER are presented in Figure 5.2. MH-TRACE energy dissipation under all BER levels and node densities is less than 40 % of the energy dissipation of IEEE 802.11.

IEEE 802.11 energy dissipation does not show any significant change with increasing BER due to the fact that the dominant energy dissipation terms in IEEE 802.11 are receive and carrier sensing and they are not significantly affected by the BER. This is because the energy dissipated for receiving a non-corrupted packet and a corrupted packet is the same. Furthermore, in carrier sensing only the presence of the carrier is important, which is not affected by the BER level of the network. Packet transmissions are

also not related with BER level (*i.e.,* data packets are coming from the application layer and they are not routed). IEEE 802.11 energy dissipations in the 100-node network and the 200-node network are very close because the network is already in saturation conditions in the 100-node network (*i.e.,* full channel utilization) and this situation does not change in the 200-node network (*i.e.,* energy dissipated for a successful reception is the same with energy dissipated on a completely overlapping collision) from an energy dissipation point of view.

MH-TRACE energy dissipation is higher for the 200-node network than the 100-node network, because of the increase in channel utilization. Note that MH-TRACE is not utilizing all of the available bandwidth in the 100-node network (*i.e.,* a significant portion of the data slots are unused). The reason for the sharp decrease in energy dissipation of MH-TRACE for both node densities for BER $> 10^{-3}$ is that the nodes spend most of their time in sleep mode. Since a large portion of the header packets are corrupted, nodes cannot have channel access. Note that in MH-TRACE, a node is only awake if there is a scheduled data transmission. If the header packet is not received, then the corresponding node stays in the sleep mode for the whole frame time.

3.7 Summary

In this chapter, we investigated the impact of channel errors on the performance of MH-TRACE and IEEE 802.11, which are examples of coordinated and non-coordinated MAC protocols, respectively, through *ns*-2 simulations using the Gilbert-Elliot channel model. As expected, the impact of channel errors is more severe on MH-TRACE than IEEE 802.11 due to the dependence of MH-TRACE on robust control packet traffic. Nevertheless, the performance of MH-TRACE remains superior to that of IEEE 802.11, even in the presence of large channel errors. Hence, the major conclusion of this study is that coordinated MAC protocols are preferable over non-coordinated MAC protocols even under noisy channel conditions [71].

We developed an analytical model for the performance of MH-TRACE as a function of network area, number of nodes and BER of the channel. We presented ns-2 simulations both to demonstrate the validity of the analytical model and to show the degradation in the MAC protocols' (*i.e.,* IEEE 802.11 and MH-TRACE) performance

with increasing BER. As expected, the impact of channel errors is more severe on MH-TRACE than IEEE 802.11 at extremely high BER levels due to the dependence of MH-TRACE on robust control packet traffic. Nevertheless, as the node density increases, MH-TRACE performs better than IEEE 802.11 (in terms of throughput and energy efficiency) even under very high BER levels due to its coordinated channel access mechanism [73, 74].

The major conclusion of this chapter is that the energy efficiency and QoS performance of coordinated MAC protocols are superior to those of non-coordinated MAC protocols. The relatively better QoS performance of non-coordinated MAC protocols at extremely high BER levels is actually deceiving due to the fact that such a low level of QoS is not beneficial to the application layer. Finally, we point out that for higher data rates or node densities coordinated protocols are expected to perform better in terms of initial throughput due to their controlled access mechanisms.

While in this chapter we investigated the effects of channel errors using a fixed channel bit error rate model for all transmitter-receiver pairs, in the next chapter we extend this model to links with varying channel capacities. Specifically, we explore ways in which varying channel capacities can be utilized to provide multi-rate support for network-wide broadcasting and multicasting, thereby further improving the network performance.

Chapter 4

Multi-rate Broadcasting in MANETs

Previous work extended the MH-TRACE MAC protocol to provide support for network-wide broadcasting of data, creating a new protocol called NB-TRACE, and to provide support for multicast transmission of data, through a protocol called MC-TRACE (see Chapter 2 for details of these protocols). Extensive simulations have shown that NB-TRACE is much more efficient than existing network-wide broadcasting techniques and MC-TRACE can create energy-efficient, reliable multicast groups.

In this chapter, we further extended this TRACE framework to allow users to select the appropriate trade-off among energy dissipation, delay and data quality, enabling TRACE to be much more responsive to end-user requirements and adaptive as these requirements change over time. Coupled with scalable video or audio coding, this multi-rate approach enables users to receive high quality video/audio with a potential delay increase.

4.1 Motivation

The wireless medium and its unique characteristics have been continuously examined to achieve many of the greatest advancements in communication technology. In mobile ad-hoc networks this medium is often called the *broadcast channel*, where transmissions occur from one user to many others simultaneously. Although broadcasting makes it easier to reach many receivers through a single transmission, it is difficult to satisfy the individual needs of different receivers simultaneously. In a broadcast channel, the

capacities of the communication links from the source to the intended recipients vary greatly due to differences in communication range, fading, and interference on these links.

In such a scenario, the rate of transmission was bounded by the receiver with the worst channel capacity before Cover and Bergmanns overcame this problem and came up with a revolutionary approach to utilizing the broadcast channel [38, 39]. Their conclusion was that it is possible to provide multiple rates to different users through a single transmission while causing little degradation for the receiver with the worst channel capacity. In this chapter, we propose a new way of exploiting the multi-rate broadcasting idea, which was initially proposed to improve the single hop throughput in broadcasting scenarios.

We concentrate on multicasting, where source node(s) convey information to the members of a multicast group, possibly through the use of non-multicast group member nodes within the network, in a multi-hop transmission fashion. Multicasting requires efficient techniques to route packets originated from different sources to different destinations with a required degree of Quality of Service (QoS) and reliability. We believe that the importance of multicasting has been and will be increasing as the number of wireless capable devices and the demand for them continue to increase. Therefore, it is crucial to design a multicasting protocol that incorporates the potential increase in the diversity and number of wireless networking devices. In this chapter, our aim is to meet this goal by providing a framework for multi-rate multicasting. In order to accomplish this, we start with the physical layer and work our way up through the communications stack.

We take the idea of multi-rate broadcasting one step further than the original goal of single-hop broadcasting, and we utilize multi-rate broadcasting to provide a flexible throughput-delay trade-off. This trade-off between throughput and delay can be exploited within the network according to the different requirements of the different members of a multicast group. Multi-rate availability through a single transmission can be used to facilitate the co-existence of streams with different importance. Using Cover's idea of superimposing information, an additional (probably less important) stream can be transmitted simultaneously with the essential stream, which carries crucial information needed by all of the members of a multicast group.

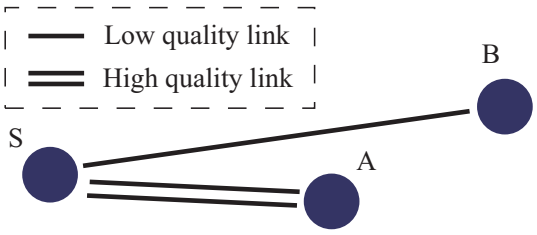


Figure 4.1: Illustration of different link qualities in broadcasting mode.

We can further explain the idea of multiple channels for a broadcast transmission with the following example. In Figure 4.1, the source node (S) is broadcasting and two other nodes (A and B) are receiving. Node A is closer to the source than node B. Thus, when node S broadcasts a packet, the received energy per bit at node A is higher than that at node B because of the decrease in the received signal strength due to the extended propagation distance. Therefore, node A has the "high quality" link in this case.

Due to differences in qualities of the links between a broadcasting source node and the intended recipients, it is advantageous to adjust the transmission scheme for broadcasting data so that a single transmission can be best received at all receivers (*i.e.,* with a single transmission, the nodes with "high quality" links receive "high rate" information whereas the nodes with "low quality" links receive "low rate" information). This can be achieved by using Cover's theory of superposed information [38]. In this technique, whereby superposed information provides multi-rate transmission, nodes with good channels will receive higher data rates while nodes with poor channels will receive at a rate close to the original (*i.e.,* the single level transmission rate). Note that, throughout this chapter having a good/bad channel implies the presence of a high/low quality link between the transmitter receiver pair.

We propose the new idea of multi-rate multicast routing, and we incorporate this idea into an existing multicasting protocol, MC-TRACE [23]. We name this new version of MC-TRACE as Multi-rate MC-TRACE (MMC-TRACE). We combine the idea of multi-rate coding with scalable source coding [75] to provide multi-level resolution of conversational voice or video, and through simulations we show the benefits and drawbacks of multi-rate multicasting. We also provide an in-depth analysis of the optimal selection of transmission radii for multi-rate multicasting.

4.2 Superposed Coding

According to the channel coding theorem, any transmission rate R can be achieved by using a standard $(2^{n(\mathrm{R}-\epsilon)}, n, \lambda_n)$ code [68]. The theorem states that all rates below capacity C are achievable, that is, for every $\epsilon > 0$ and rate $\mathrm{R} < C$, there exists a sequence of $(2^{n\mathrm{R}}, n)$ codes with maximum probability of error $\lambda_n \leq \epsilon$, for n sufficiently large. This coding makes sure that, with high probability, channel noise will not cause errors in the decoding process. The idea of *superposed coding* is to add additional coding on top of the first coding in such a way that already separated codewords are displaced again according to new information. The second displacement is smaller than the first one, and it is unlikely to be decoded by any receiver with a poor channel. However, any receiver with a good channel will be able to decode both the first and the second displacements of the codewords. Therefore, nodes with poor channels decode only the lower rate information (*i.e.,* the first displacement of the codewords) while nodes with good channels decode both the lower rate and the additional information. This process is visualized in Figure 4.2.

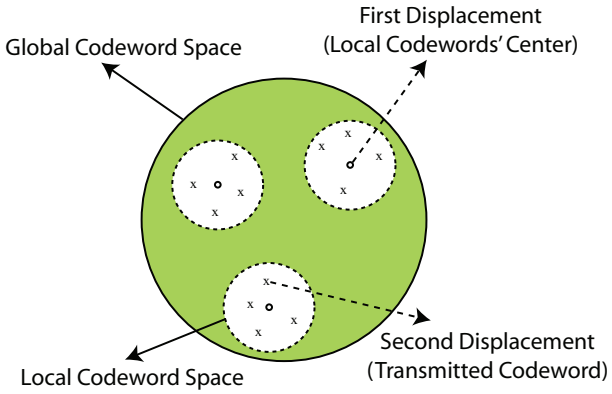


Figure 4.2: Codeword distribution after two levels of displacements.

Note that introducing additional information to the codewords will make them more vulnerable to channel errors, since the new distance between the codewords is smaller than it would be if we coded only the lower rate information. On the other hand, Cover proved that a degradation in the rate for the poor channel will allow a more rewarding increase in the rate for the good channel [38]. Bergmans proved that the set of rates achieved by this technique defines the upper limit for a broadcast channel [39].

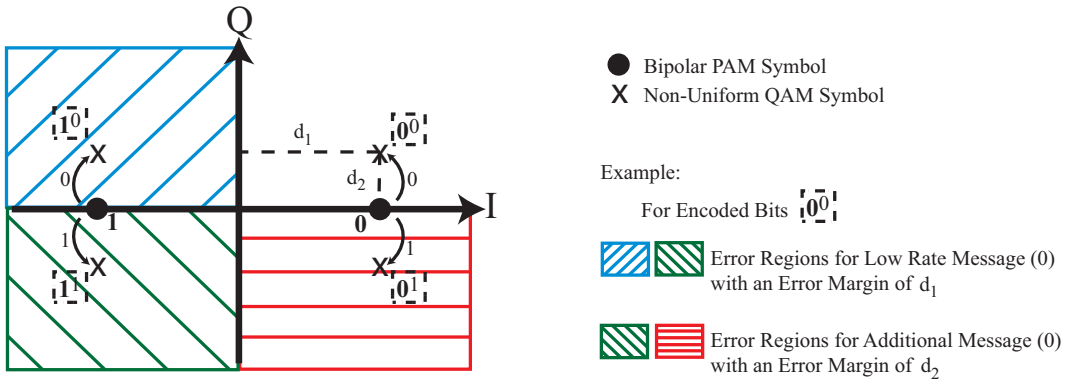


Figure 4.3: Constellation diagram for non-uniform quadrature amplitude modulation (QAM). The noise margins and error regions for the selected symbol are also illustrated.

We demonstrate the idea of superposed coding by using a simple binary pulse amplitude modulation (BPAM) constellation to represent the low rate information. In order to superpose the high rate information we transform the binary PAM constellation into nonuniform quadrature amplitude modulation (QAM). Figure 4.3 illustrates the error regions and different noise margins, d_1 and d_2, for the low rate and additional information, respectively. Note that these regions and margins correspond to the selected symbol. The nonuniform spacing makes it much easier for a receiver to correctly recover the low rate information (first bit) than the additional information (second bit). Therefore, the first bit can be decoded by all receivers within the transmission range of the source, while the second bit conveys the additional information that can only be recovered by receivers with good channels. In this way, we simultaneously send two packets using a single transmission and offer two different rates of broadcasting, which potentially may lead to doubled link throughput.

Using non-uniform QAM as shown in Figure 4.3 increases the transmit energy per symbol (ε_s) from d_1^2 units to;

$$\varepsilon_s = d_1^2 + d_2^2, \tag{4.1}$$

units. On the other hand, if we assume that the energy per symbol is kept constant at ε_s units we get the following relationship between d_1, d_2 and ε_s;

$$d_1 = \sqrt{\varepsilon_s - d_2^2} \tag{4.2}$$

Equation 4.2 tells us that as d_2 increases d_1 has to decrease to compensate for the extra energy needed for increasing additional information coverage. This will result in a reduced range of transmission and hence, reduced connectivity of the network. Note that this example can be easily extended to provide more than two levels of superposed information.

We have determined the amount of increase in the transmit energy in terms of noise margins in the constellation diagram. However, in order to determine the amount of transmit energy needed to provide multi-rate availability at a certain distance we have to look at the propagation model. The propagation model determines the path loss along a link and also the effective coverage area of a transmitter.

In MANET's, the complexity of signal propagation makes it difficult to obtain a single model that characterizes path loss accurately for different environments. When tight system specifications must be met, accurate path loss models can be obtained from complex analytical models or empirical measurements [1]. However, for our general tradeoff analysis, we use a simple model that captures the essence of signal propagation without resorting to complicated path loss models, which are only approximations to the real channel anyway. We use a hybrid model that consists of the free space and the two-ray ground reflection models. Friis presented the following equation to calculate the propagation loss in free space at a distance r from the transmitter [76].

$$\varepsilon(r) = \frac{\varepsilon_s G_t G_r \lambda^2}{(4\pi r)^2 L}, \tag{4.3}$$

where ε_s is the transmit energy per symbol, G_t and G_r are the antenna gains of the transmitter and the receiver, respectively, L ($L \geq 1$) is the system loss, and λ is the wavelength. In our simulations we set $G_t = G_r = 1$ and $L = 1$.

This model considers the communication range as a circle around the transmitter. At long distances, the Friis model provides less accurate prediction of the received signal strength compared with the two-ray ground reflection model, which considers both the direct path and a ground reflection path [77].

$$\varepsilon(r) = \frac{\varepsilon_s G_t G_r (h_t)^2 (h_r)^2}{r^4 L}, \tag{4.4}$$

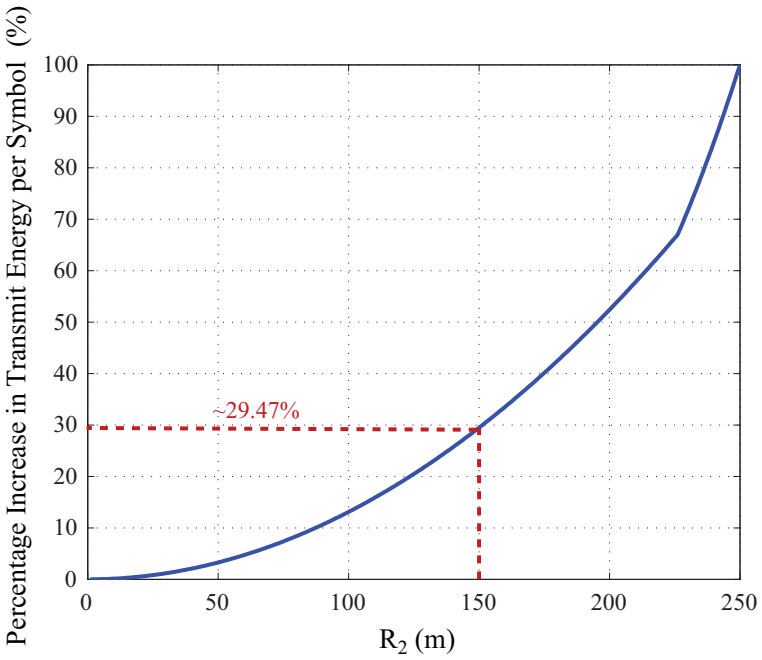


Figure 4.4: Increase in the transmit energy per symbol with increasing R_2 (with fixed $R_1 = 250m$).

where h_t and h_r are the heights of the transmit and receive antennas, respectively. To be consistent with the free space model, we again set $G_t = G_r = 1$ and $L = 1$. Equation 4.4 shows a faster propagation loss than Equation 4.3 as distance increases. However, the two-ray model does not provide an accurate estimation of received energy per symbol for short distances due to the oscillation caused by the constructive and destructive combination of the two rays. Instead, the free space model should be used when d is small. Therefore, a cross-over distance r_c is calculated in the hybrid model [77]. When $r < r_c$, Equation 4.3 is used to predict receive signal strength, and when $r > r_c$, Equation 4.4 is used. This cross-over distance, r_c, calculated as,

$$r_c = \frac{4\pi h_t h_r}{\lambda} \tag{4.5}$$

Using Equations 4.3-4.5 and the modulation scheme shown in Figure 4.3, for a fixed single rate transmission radius $R_1 = 250m$ we can calculate the amount of additional energy per symbol needed to introduce the aforementioned multi-rate modulation scheme. With R_1 constant and utilizing the relationship in Equation 4.1, we plot the

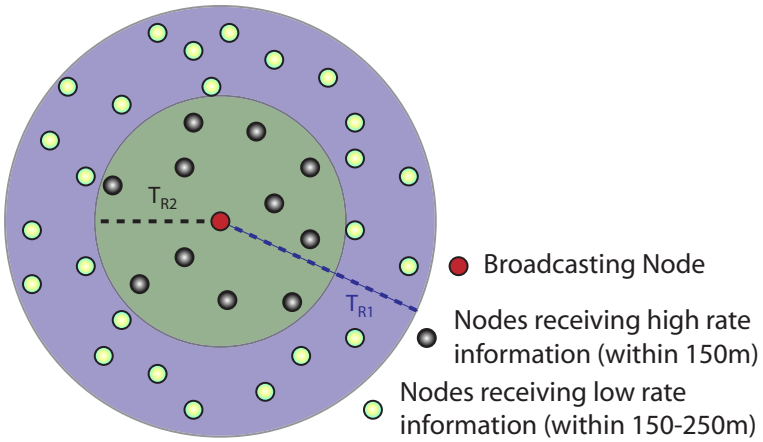


Figure 4.5: Overlapping communication regions for the low rate and the additional information.

percentage increase in the transmit energy per symbol needed to provide multi-rate coverage R_2 (see Figure 4.4). This figure illustrates how much more energy per symbol we require to cover a larger region where nodes receive both the low and additional rate information (*i.e.,* the *high-rate* information as we call in this thesis). Figure 4.5 illustrates the two regions where different rates are available. These figures can be utilized to determine the amount of extra energy per symbol needed for a multi-rate broadcasting system, where the multi-rate coverage area is of the utmost importance.

On the other hand, if we need to keep the initial energy per symbol constant, we have to determine the amount of energy to dedicate to each rate (*i.e.,* low-rate and additional rate). Figure 4.6 illustrates the energy per symbol distribution between the two rates achieved through non-uniform QAM modulation. We assume that ε_s is the initial transmit energy per symbol that is needed to provide a $250m$ transmission radius for the traditional single-rate broadcasting scenario that employs binary PAM. The value of ε_s is kept constant, and at the intersection point of the curves, where $R_1 = R_2$, this energy per symbol ε_s is shared between the two rates of the transmission. At this point the modulation scheme becomes regular QAM and is no longer non-uniform as depicted in Figure 4.3.

It is also interesting to see how the fixed transmit energy per symbol is shared between the two rates as we increase the coverage of the additional rate transmission (*i.e.,*

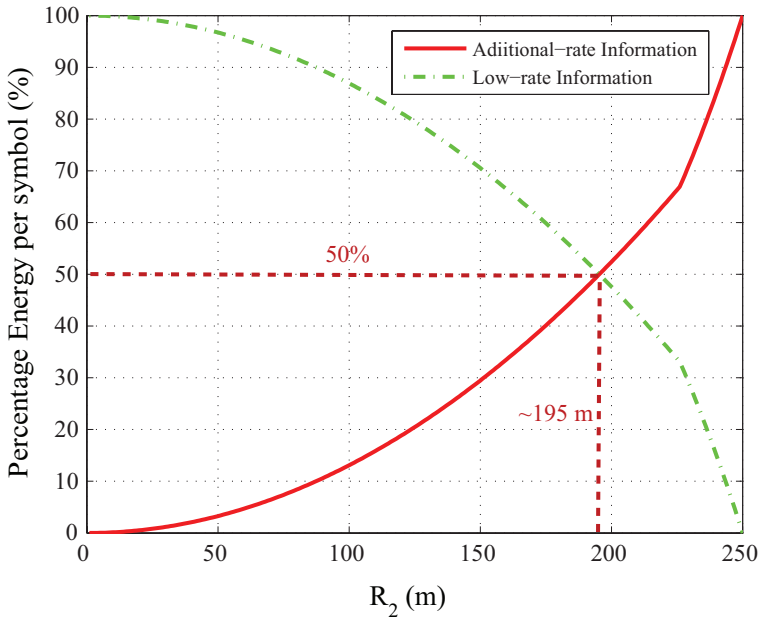


Figure 4.6: Percentage energy per symbol, when compared to initial transmit energy per symbol ε_s (fixed), used for modulating the low-rate and the additional-rate information. As R_2 increases, the coverage for additional rate information increases, while the amount of energy used for low-rate information decreases, and hence R_1 decreases.

increasing R_2). As Cover showed to the world [38], we have confirmed that with a little sacrifice from the low-rate coverage, we will be able to have a drastic increase in the high-rate transmission radius R_2. At the point where half of the initial transmit energy per symbol has been sacrificed, we have equal coverage of $\sim 195m$ for both rates. This is a direct result of the quadratic relationship between the energy per symbol and the noise margins (Equation 4.1). Note that Figure 4.6 was generated using the hybrid propagation model and Equation 4.2. However, this relationship can be easily derived for any channel propagation model.

Our study is motivated by these already proven capabilities of multi-rate broadcasting and also by the fact that these ideas have not been employed in a multicasting scenario before. In the remainder of this chapter, we discuss our idea of combining multi-rate broadcasting with multi-hop routing, providing multi-rate network-wide broadcasting and multi-rate multicasting. We show the improvement this integration can bring to a multi-hop network.

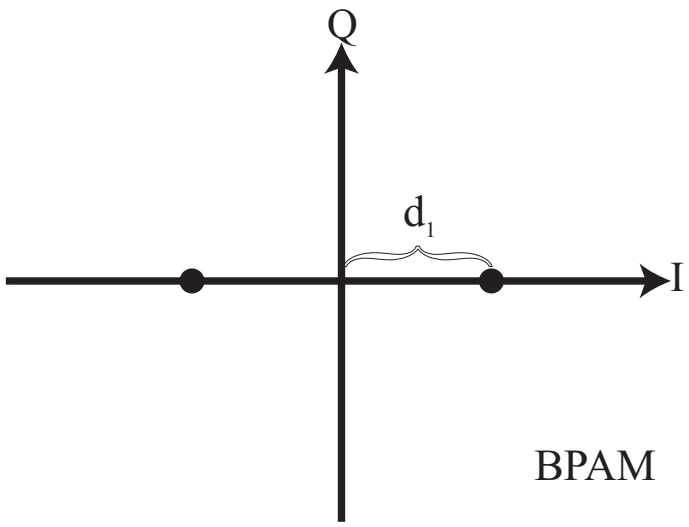


Figure 4.7: The Binary Pulse Amplitude Modulation (BPAM) constellation.

4.3 Constellation Diagram Design in Multi-rate Multicasting

In this study, we employ a two-level multi-rate multicasting scenario. However, the main idea of multi-rate multicasting need not be limited to two levels. Various non-uniform constellations can be designed to extend this idea to more levels, which, coupled with scalable coding, provides even more flexibility and choices to the multicast members.

In this section we provide further insight into superposed coding. We start with a Binary Pulse Amplitude Modulation (BPAM) constellation and modify this according to Cover's pointers presented in his paper about broadcast channels [38]. Figure 4.7 shows the simple constellation diagram of BPAM. Figure 4.8 shows the non-uniform Quadrature Amplitude Modulation (QAM) constellation which is obtained by assigning a smaller displacement to the quadrature component than the in-phase component in the BPAM constellation. This can be thought of as both non-uniform QAM or non-uniform QPSK [40].

As we discussed earlier in Section 4.2, uneven distribution of constellation points results in different noise margins for the bits encoded in a transmitted symbol. We can

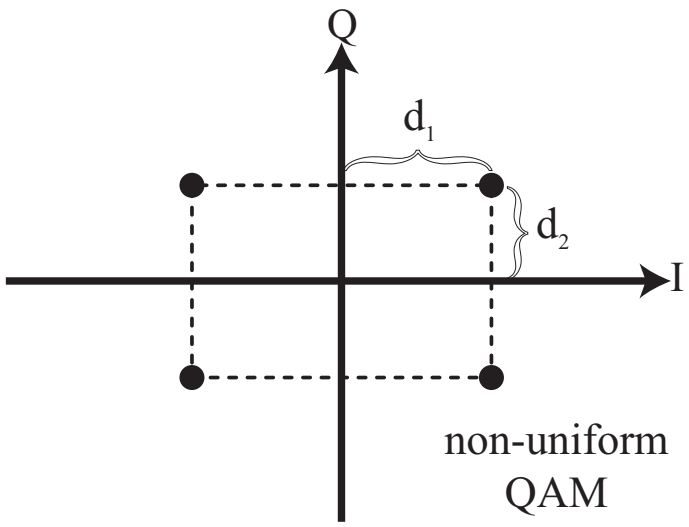


Figure 4.8: Non-uniform Quadrature Amplitude Modulation (QAM).

further use this idea to introduce many levels of noise margins. Figure 4.9 shows a non-uniform 16-ary QAM constellation, which provides four different noise margins for the encoded four bits of information per symbol. If we assume again that the propagation loss mainly depends on distance, we have four different penetration distances for these bits. In other words, there are four regions in which the transmitted symbol can be decoded at four different rates. Sun et al. [41] utilize a special case of the non-uniform 16-ary QAM constellation in achieving a flexible unequal error protection.

Figure 4.10 illustrates a 64 point non-uniform constellation that utilizes more decoding regions when different rates are needed. Adjusting the distribution of the symbols in a non-uniform M-ary constellation can be determined according to several design parameters. Density of the network, network boundaries, power limitations, mobility patterns, antenna type, and requirements of the application can be listed as some of the many factors that may affect the constellation diagram design. In this thesis, we utilize the non-uniform QAM (non-uniform QPSK) constellation to achieve two different rates at two different distances to achieve multi-rate multicasting. However, any of the constellations above can be used with slight modifications of the multicasting protocol.

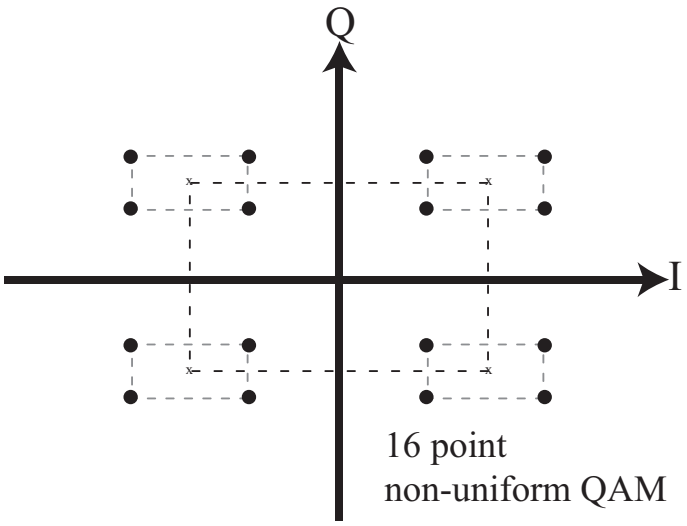


Figure 4.9: 16 point non-uniform Quadrature Amplitude Modulation (QAM).

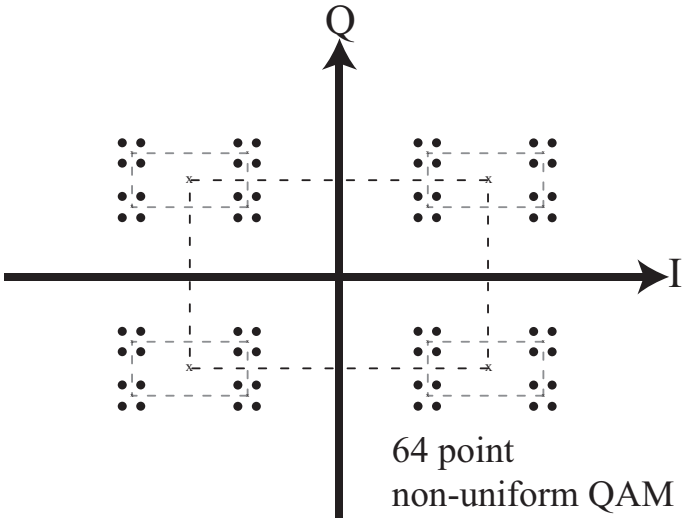


Figure 4.10: 64 point non-uniform Quadrature Amplitude Modulation (QAM).

4.4 Multi-rate Network-wide Broadcasting

After explaining both the idea of superposed coding and the NB-TRACE protocol (see Section 2.1.2), we now combine these techniques to achieve multi-rate network-wide broadcasting (*i.e.,* multicasting to the entire network). Nodes that have *high quality* links (*i.e.,* nodes that can decode both of the superposed information, which we name *low rate* and *additional* information) with the source node can rebroadcast the superposed low rate and additional information pair. On the other hand, any node with a *low quality* link with the source node cannot decode the additional information, therefore, it has to rebroadcast only the low rate information. There is also a third case in which nodes receive both low rate and additional information through different paths with different delays. This availability allows the nodes to decide whether they want to receive packets with low rate and low delay (low rate data comes through longer hops and thus fewer hops, which results in lower delay) or high rate with higher delay (high rate data comes through shorter hops and thus more hops, which results in higher delay). Note that, for simplicity we assume there are only two states for a given link. However, as one increases the number of levels in superposed coding, the number of link quality levels will increase accordingly.

We propose that one of the most effective ways of exploiting the multi-rate information availability is to employ scalable source coding and make full use of all the available rate at every time instant. Scalable source coding is basically a hierarchical coding scheme, where coarser representations are embedded into finer ones, thereby allowing access to the source at a variety of resolutions. Rimoldi [78] generalized the scalable source coding problem and discovered necessary and sufficient conditions for the achievability of any sequence of rates and distortions. Scalable source coding is often used in wireless communication applications where the available communication rate is time-varying [75]. However, at a given time during a transmission there is only a certain rate of information available, which is determined by the available communication rate. In other words, although the broadcast information rate is variable when we employ scalable source coding, there is only a single transmission rate at a given time. On the other hand, when we combine multi-rate transmission with scalable coding in a MANET, nodes with different link qualities will have different rates available to them instead of a single rate that varies according to source coding.

Table 4.1: Simulation Setup

PARAMETER	VALUE
Number of Nodes	256
Simulation Area	1000m x 1000m
Simulation Time	100s
Transmission Ranges (Low-High)	(150m-250m)
Number of Repetitions	5
Node Mobility	Random Way-Point

4.5 Multi-rate Network-wide Broadcasting Simulations

In this section, we want to investigate both ends of the delay vs. throughput (quality) trade-off by using NB-TRACE and Flooding with IEEE 802.11. Table 4.1 summarizes the simulation setup we used to investigate these architectures. We performed two sets of simulations where each set has a different priority. First, we prioritize the reception of the packets with lowest delay. This leads to a quicker network-wide broadcasting and mainly the low rate traffic is forwarded by the nodes. In the second set of simulations, throughput is the priority for the nodes, which thus have to forward the high rate traffic. All the nodes receive both the low rate and the additional information while compromising the delay since the high rate information can only be recovered by the nodes within R_2 as shown in Figure 4.5. These two extreme priorities helps us to demonstrate the limits and capabilities of superposed coding on network-wide broadcasting. Note that in both scenarios, the source node broadcasting is the same, however, which nodes rebroadcast and the rates they use for rebroadcasting change depending on the priorities.

We simulated conversational voice coded at two different rates, namely the low rate ($13Kbps$) and the high rate ($26Kbps$). The high rate corresponds to two (superposed) voice packets per superframe and can be decoded by nodes within $R_1 = 150m$, whereas the low rate corresponds to one voice packet per superframe and can be decoded by nodes within $R_2 = 250m$. The channel rate is set to 2 Mbps and the standard IEEE 802.11 physical layer is employed for both architectures. All the simulations are

Table 4.2: Simulation Parameters

Acronym	Description	Value
T_{SF}	Superframe duration	61.5ms
N_F	Number of frames	7
N_{DS}	Number of data slots per frame	14
N_C	Number of cont. slots per frame	15
N/A	Data packet size	110B
N/A	Header packet size	36B
N/A	All other control packet size	10B
T_{ds}	Packet drop threshold (@ source node)	61.5ms
T_{di}	(@ intermediate nodes)	250.0ms
T_{VF}	Voice packet generation period	61.5ms
D_{CS}	Carrier Sense range	507m

run with 256 nodes, moving within a 1 km by 1 km area for 100 seconds according to the Random Way-Point (RWP) mobility model with node speeds chosen from a uniform distribution between 0.0 m/s and 5.0 m/s. The source node is stationary and located in the middle of the network. Acronyms, descriptions and values of the parameters used in the simulations are presented in Table 4.2.

4.5.1 Throughput and Packet Delay

The throughput results are given in terms of average Packet Delivery Ratios (PDRs) of both architectures in Table 4.3. Table 4.4 presents the average packet delay and delay jitter values. The number of data packets received by any node other than the source node are averaged and divided by the total number of data packets broadcasted by the source node to obtained the average PDRs. PDRs of both NB-TRACE and Flooding with IEEE 802.11 are above 99% when the priority is low delay. However, when the priority is throughput IEEE 802.11 PDR decreases (less than 90%) due to the increase in the traffic. Note that even though the PDR decreases, the total throughput increases (88% of 26 Kbps is more than 99% of 13 Kbps).

Table 4.3: Packet Delivery Ratios (PDRs) according to the two priorities: Delay and Throughput

Architecture	**Packet Delivery Ratio (Delay)**	**Packet Delivery Ratio (Throughput)**
NB-TRACE	99.6% 99.2% (min)	99.4% 97.7% (min)
Flooding (IEEE 802.11)	99.5% 99.4% min)	88.1% 84.3% (min)

In the high throughput case, nodes are forced to receive and rebroadcast high rate information in order to achieve high throughput. High rate information has a range of $150m$, $100m$ less than low rate information. This leads to a smaller communication radius and an increased number of rebroadcasting nodes. Therefore, the amount of traffic needed to achieve network-wide broadcasting increases, which causes the performance drop in Flooding with IEEE 802.11. On the other hand, increased traffic brings the availability of high rate information for all the nodes in the network. Actually, the throughput of the network is almost doubled for NB-TRACE even though the PDR value stays almost the same. This is simply because the high rate packets have twice as much information as the low rate packets.

Table 4.4: Packet Delay and Delay Jitter according to the two priorities: Delay and Throughput

Architecture	**Packet Delay & Delay Jitter (Delay)**	**Packet Delay & Delay Jitter (Throughput)**
NB-TRACE	61.0 ms 10.6 ms (jitter)	192.9 ms 10.9 ms (jitter)
Flooding (IEEE 802.11)	12.3 ms 63.7 ms (jitter)	41.1 ms 69.6ms (jitter)

When we impose delay as a constraint on packet forwarding, only a fraction of nodes in the network receive high rate packets that are broadcasted by the source node.

Other nodes are forced to receive low rate packets, which have lower delay values, since they propagate through the network faster (*i.e.,* instead of waiting to recover the high rate information, nodes have to settle with lower delay, low rate information). On the other hand, having throughput as the priority forces nodes to wait for multi-rate broadcast by intermediate nodes that are close enough to the source node located in the middle of the simulation area. This increases the packet delay since the packets are propagated throughout the network with a smaller effective radius of communication, which is $150m$ instead of $250m$ in this case. In NB-TRACE, data packet transmissions are coordinated by clusterheads and data slots become available with the period T_{SF}. This cyclic frame structure results in higher delay values for NB-TRACE. Flooding with IEEE 802.11 has lower delay values since IEEE 802.11 allows nodes to transmit whenever the channel is available. On the other hand, NB-TRACE jitter is less than 16% of IEEE 802.11 jitter in both cases because of the automatic renewal of the channel access, which reduces the variation in the interarrival times of data packets. In fact, low jitter is the most important QoS parameter in multimedia communications.

4.5.2 Energy Consumption

One of the most important advantages of NB-TRACE over Flooding with IEEE 802.11 is its better energy efficiency. Average energy dissipation per second for NB-TRACE and Flooding with IEEE 802.11 with 256 nodes as a function of packet forwarding priority are presented in Table 4.5.

NB-TRACE energy dissipations both under lower delay and high throughput priorities are less than 20% of the energy dissipation of Flooding with IEEE 802.11. Yet, the energy dissipation of NB-TRACE is higher when the priority is higher throughput, which is expected because the amount of traffic is doubled. Furthermore, more hops are needed to reach outer nodes of the network and consequently more nodes are involved in the rebroadcasting process, hence fewer nodes can stay in the sleep mode.

Flooding with IEEE 802.11 energy dissipation does not show any significant change when we change the packet forwarding priority due to the fact that flooding already engages all nodes in the network to rebroadcast all data packets. The only difference is in transmission where almost 14% more power is needed when forwarding the high rate information than is needed for flooding just the low rate information.

These results help us understand what we are trading off in order to double our throughput (by receiving both the low rate and the additional information). The two extreme constraints on packet forwarding presented in this section should bound any result that is obtained by using this combination of delay and throughput constraints.

Table 4.5: Energy Consumption according to the two priorities: Delay and Throughput

Architecture	**Energy Consumption (Delay)**	**Energy Consumption (Throughput)**
NB-TRACE	35.2 mJ/s	48.5 mJ/s
Flooding (IEEE 802.11)	237.3 mJ/s	240.1 mJ/s

4.6 Multi-rate Multicasting

In a multicast scenario, members of the multicast group are often left without any control on the rate of broadcasting of the source node. This is due to the fact that it is neither efficient nor desired to adjust the parameters of the multicast source according to the resources and/or requirements of each individual multicast group member. In fact, trying to do so turns multicasting into unicasting, where multiple streams are transmitted, with each stream adjusted according to the individual destination node. Therefore, multicast parameters such as degree of forward error correction and type of source and channel coding and therefore, the resulting bit rate of the transmission are often controlled by the source node, and the only choice that a multicast group member usually has is whether or not to listen to the multicast stream.

4.6.1 Overview

The goal of our work is to provide more choices for the individual multicast members through a single transmission by the multicast source node. To achieve this, we combine the multi-rate broadcasting technique described in section 4.2 with the previously introduced multicasting protocol MC-TRACE to achieve multi-rate multicasting. Nodes that have *high quality* links with the source node (*i.e.,* nodes that can decode

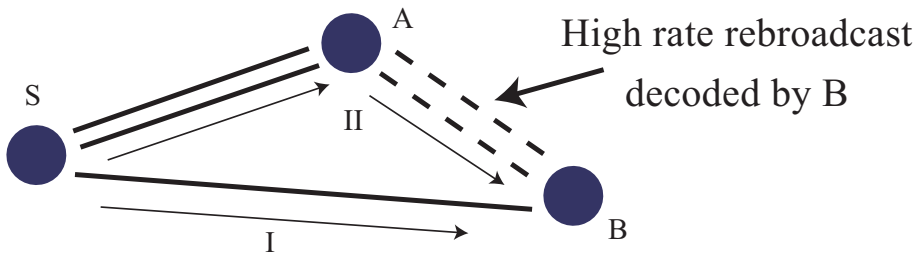


Figure 4.11: Two different rates of information available at node B. Flow I, low rate information directly from the source and flow II, high rate information through node A.

both of the superposed information, which we name *low rate* and *additional* information) can rebroadcast the superposed low rate and additional information pair. On the other hand, any node with a *low quality* link with the source node cannot decode the additional information, therefore, it has to rebroadcast only the low rate information. There is also a third case in which nodes receive both low rate and additional information through different paths with different delays. This availability allows the nodes to decide whether they want to receive packets with low rate and low delay (low rate data comes through longer hops and thus fewer hops, which results in lower delay) or high rate with higher delay (high rate data comes through shorter hops and thus more hops, which results in higher delay). Note that, for simplicity we assume there are only two states for a given link. However, as one increases the number of levels in superposed coding, as show in Section 4.3, the number of link quality levels will increase accordingly.

Figure 4.11 shows an example where node B can receive low rate information with low delay directly from the source node S (flow *I*) or high rate information with high delay rebroadcast from node A (flow *II*). Using conventional broadcasting, where only one rate is available, information flow *I* would be the only choice for node B since the rebroadcast from node A would be the same as the source's broadcast. However, using superposed coding enables multi-rate broadcasting, providing options for node B to receive either low rate or high rate information. This idea is illustrated in Figure 4.12 as well. At time T_1 the source node transmits the first superposed high rate packet $P_{L_1}|P_{A_1}$ consisting of the low rate information packet P_{L_1} superposed with the additional rate information packet P_{A_1}. The row starting with R_1 shows the packets received (*i.e.,*

	S	A	B
T_1	$P_{L_1}\|P_{A_1}$	—	—
R_1	—	$P_{L_1}\|P_{A_1}$	P_{L_1}
T_2	$P_{L_2}\|P_{A_2}$	$P_{L_1}\|P_{A_1}$	P_{L_1}
R_2	—	$P_{L_2}\|P_{A_2}$	$P_{L_1}\|P_{A_1}$ & P_{L_2}
⋮			

Figure 4.12: Packet Flow.

decoded) by nodes A and B. Node A decodes both the low rate and the additional information, while node B, having a bad link with S, decodes only the low rate part of the superposed packet. The next set of transmissions takes place at time T_2. At this time S broadcasts the next superposed high rate packet $P_{L_2}|P_{A_2}$. At the same time, nodes A and B rebroadcast their previously received packets $P_{L_1}|P_{A_1}$ and P_{L_1}, respectively. As can be see from the next row of Figure 4.12, node B has both rates of information available and can choose either of the flows (I or II in Figure 4.11) according to its delay-throughput requirements.

We propose that one of the most effective ways of exploiting the multi-rate information availability and the cross-layer protocol design is to have more than one rate available at every time instant. Scalable source coding is basically a hierarchical coding scheme, where coarser representations are embedded into finer ones, thereby allowing access to the source at a variety of resolutions. Rimoldi [78] generalized the scalable source coding problem and discovered necessary and sufficient conditions for the achievability of any sequence of rates and distortions. Scalable source coding is often used in wireless communication applications where the available communication rate is time-varying [75]. However, at a given time during a transmission there is only a certain rate of information available, which is determined by the available communication rate. In other words, although the broadcast information rate is variable when we

employ scalable source coding, there is only a single transmission rate at a given time. On the other hand, when we combine multi-rate transmission with scalable coding in a MANET, nodes with different link qualities will have different rates available to them instead of a single rate that varies according to source coding. This diversity in the rates results in the availability of multiple flows at a given node (as we have shown in Figure 4.12). This enables the node to select the appropriate delay-quality trade-off.

4.6.2 The Approach

In MC-TRACE, the basic design philosophy behind the networking part of the architecture is to establish and maintain a multicast tree within a mobile ad hoc network using broadcasting to establish the desired tree branches and pruning the redundant branches of the multicast tree based on feedback obtained from the multicast members. We aim to utilize the already existing feedback traffic of MC-TRACE to implement the multi-rate broadcasting. The energy efficient behavior of MC-TRACE should help us compensate for the energy consumption due to the extra traffic generated by multi-rate broadcasting.

Before we can incorporate multi-rate broadcasting into multicast routing, we need to modify the way multicast trees are formed. In MC-TRACE, the branch formation and pruning starts with the source node, which initiates the flooding by broadcasting packets to its one-hop neighbors. Nodes that receive a data packet contend for channel access, and the ones that obtain channel access retransmit the data they received. Eventually, the data packets are received by all the nodes in the network, possibly multiple times. Each retransmitting node acknowledges (ACKs) its upstream node by announcing the ID of its upstream node in its IS packet, which precedes its data packet transmission.

As a first step towards implementing multi-rate MC-TRACE (MMC-TRACE), two independent mechanisms are required to assure the high-rate packet delivery. The first mechanism is a flag in the IS packet, which is set by the transmitting node to declare that the succeeding transmission will carry high-rate information. If the flag is not set, this means that the succeeding transmission will only carry low-rate information. This mechanism is called *high-rate availability* (see Figure 4.13). In addition to the availability of the high-rate information, we also need to ensure the decodability of it. This is achieved by superposing training data with the payload of the original IS

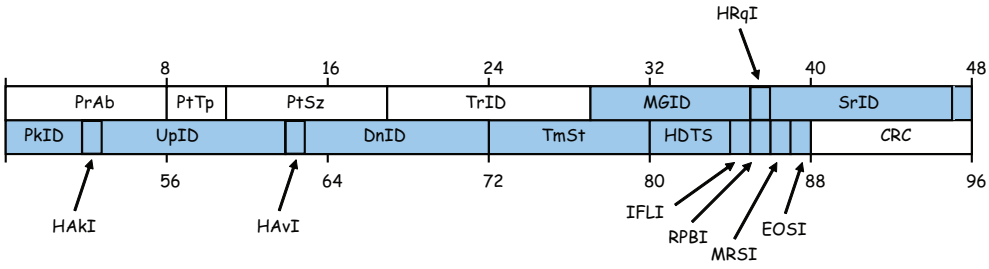


Figure 4.13: MC-TRACE Information Summarization (IS) packet format and fields. PrAb (8 bits, Preamble), PtTp (3 bits, Packet Type), PtSz (8 bits, Packet Size), TrID (10 bits, Transmitter Node ID), MGID (8 bits, Multicast Group ID), HRqI (1 bit, High-rate Request Indicator), SrID (9 bits, Source Node ID), PkID (5 bits, Packet ID), HAkI (1 bit, High-rate ACK Indicator), UpID (9 bits, Upstream Node ID), HAvI (1 bit, High-rate Availability Indicator), DnID (9 bits, Downstream Node ID), TmSt (8 bits, Timestamp), HDTS (4 bits, Hop Distance To Source), IFLI (1 bit, Initial Flooding Indicator), RPBI (1 bit, Repair Branch Indicator), MRSI (1 bit, Multicast Relay Status Indicator), EOSI (1 bit, End-of-Stream Indicator), CRC (8 bits, Cyclic Redundancy Check). Packet fields shown with light background are mandatory fields of an IS packet. The dark background fields, which are superposed with the training data, are the payload of the IS packet.

packet shown in Figure 4.13. This data must be known by all the nodes in the network. Nodes attempt to decode the superposed training data in order to decide whether or not they will be able to decode the succeeding high-rate packet transmission. We name this mechanism *high-rate decodability*. Note that we assume that the channel does not vary significantly during the time between the successive transmissions of IS and corresponding data packets (less than nine milliseconds). As a result, any receiver can assure the decodability of any high-rate data by making use of these two mechanisms. Although the availability of high-rate information may vary at a given time, we employ the second mechanism, high-rate decodability, regardless of the availability of high-rate information in order to ensure a high-rate capable branch formation.

After the initial flooding, nodes have multiple upstream nodes (*i.e.,* multiple nodes that have lower hop distance to the source than the current node) and downstream nodes (*i.e.,* multiple downstream nodes acknowledging the some upstream node as their up-

stream node). A node with multiple upstream nodes chooses two upstream nodes that have the least packet delay and can provide low and high-rate data as its upstream nodes (N_H and N_L) to be announced in its IS slot.

Multicast group member nodes indicate their status by announcing whether or not there is a need for high-rate packets in the IS packet (see the High-rate Request Indicator (HRqI) in Figure 4.13). If an upstream node receives an acknowledgment, H-ACK or L-ACK from a downstream multicast group member, it marks itself as a multicast relay and announces its status by forwarding the need for high-rate packets to one of its previously determined upstream nodes with the help of the IS packet. This mechanism continues in the same way up to the source node. Multicast relay status expires if no ACK is received from any downstream (for both members and non-members of the multicast group) or upstream (only for members of the multicast group) multicast relay or multicast group member for T_{RLY} time [23].

In Figure 4.14, we illustrate the branch formation algorithm from the point of view of nodes that participate in forming a multicast branch. This algorithm dictates that a branch formation starts with a multicast group member node and works its way up to the source node as we mentioned before. The difference between the two branch formations emerges when a multicast group member decides to request high-rate information. During the initial flooding all nodes in the network determine their possible lowest delay and high-rate capable upstream nodes, N_L and N_H, respectively. After a group member decides that high-rate information is needed, the member node must acknowledge (H-ACK) its upstream high-rate neighbor (N_H), informing that high-rate information is needed. This branch is maintained and repaired (if necessary) by the acknowledgement process shown in Figure 4.14.

Using this branch formation algorithm, each multicast group member can decide whether or not to acknowledge an upstream neighbor that provides low-rate or high-rate, based on the node's requirements. The high-rate decodability mechanism plays an important role in our branch forming algorithm. Therefore, every transmitting node uses this mechanism to let the nodes within their single-hop transmission range determine their possible high-rate upstream nodes. If the multicast member does not choose to require high-rate from its upstream node, branch formation occurs with L-ACKs as described in [23].

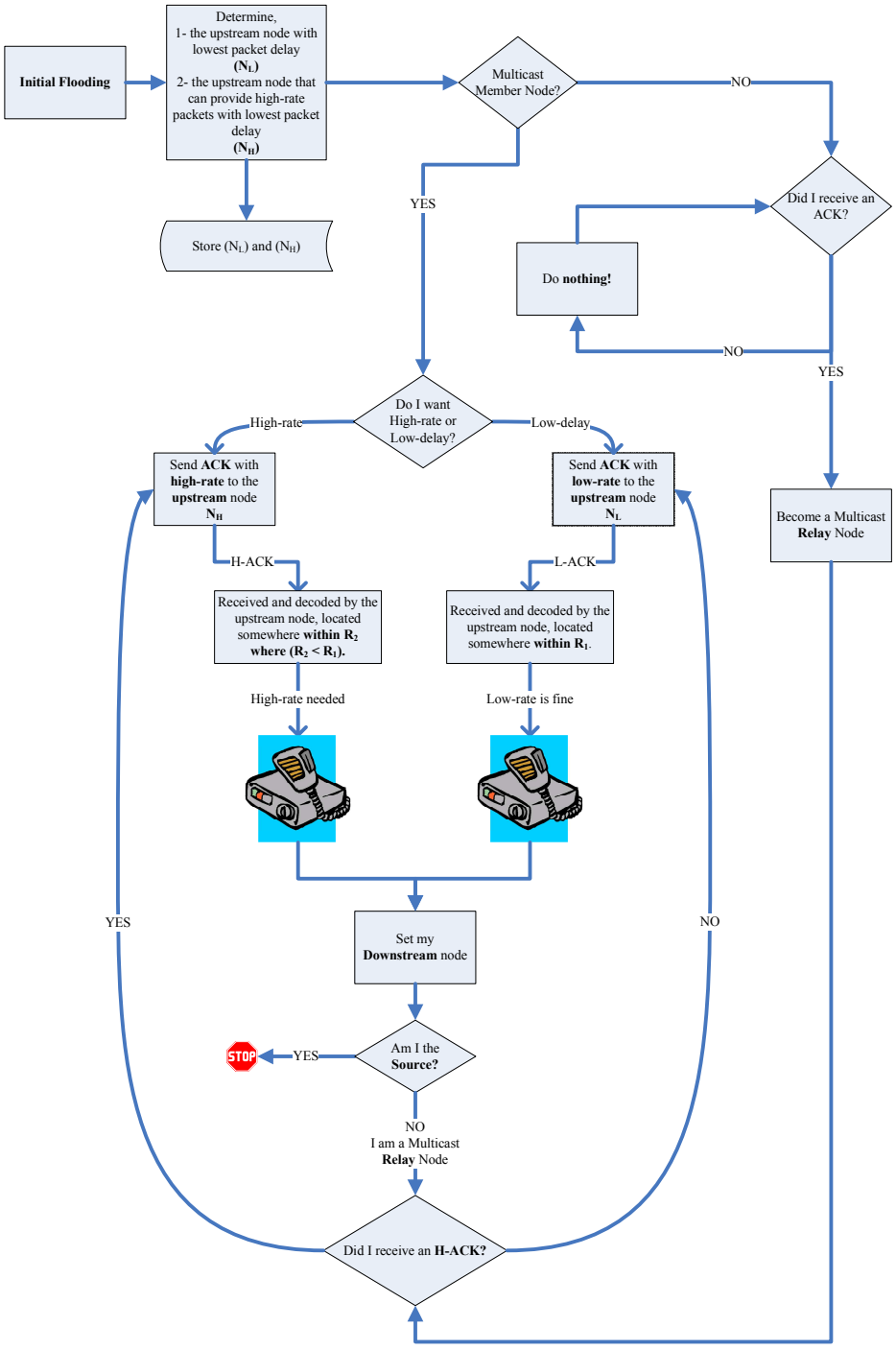


Figure 4.14: Flow chart demonstrating the branch formation algorithm for multi-rate multicasting.

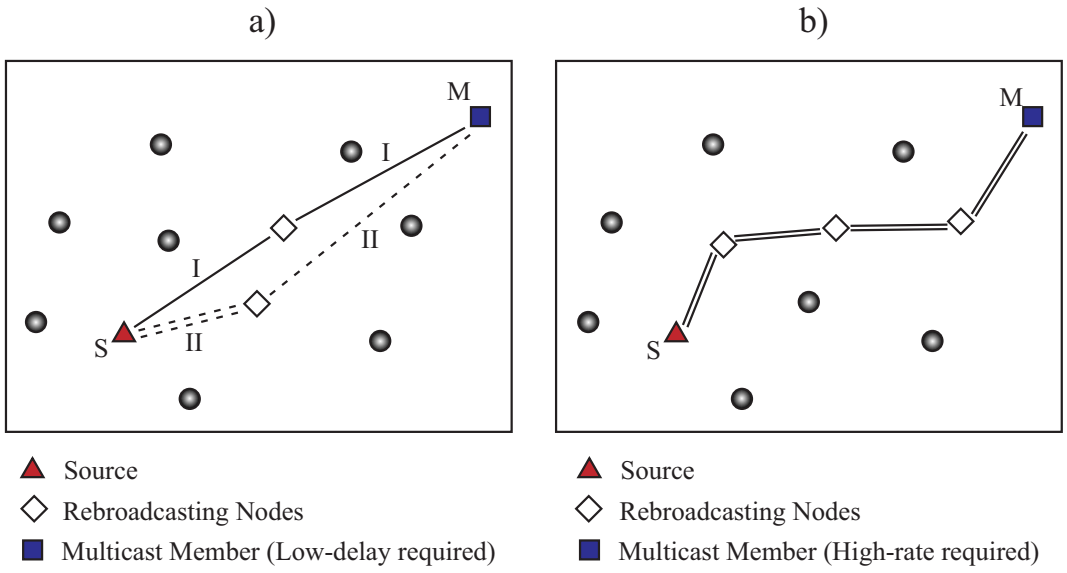


Figure 4.15: **a)** Low-delay priority forwarding. **b)** High-rate priority forwarding.

In Figure 4.15, we demonstrate two possible scenarios where the multicast member has two different priorities, low-delay and high-rate, respectively. The low-delay priority is often the goal for many routing protocols, and this might be achieved by using different priorities such as minimum number of hops routing and shortest distance routing (if location information is available). We demonstrate two examples of multicast branches that can be formed between the source (S) and the multicast member (M) after the initial flooding and pruning. Branch I is a regular branch that could have been formed by MC-TRACE. However, branch II includes a high-rate link between the source (the triangle) and the rebroadcasting node (the diamond) and can only be formed if there is multi-rate availability. Although this high-rate connectivity is not utilized by the next hop of branch II, both branches are two hops long and multicast branch candidates for the multicast member (M). In such a case where M has more than one upstream node (*i.e.,* nodes with lower hop distance to the source than M), the node that has the least packet delay is chosen as the upstream node and announced in M's IS slot.

In Figure 4.15-b, the branches are formed according to the branch formation algorithm described in Figure 4.14. The reason for the difference between the different branch formations obtained by utilizing MMC-TRACE is the fact that the multicast

group members have two different priorities, namely high-rate and low-delay. For these two different requests, different multicast branches are formed between the members. When the multicast group members request different rates simultaneously, only MMC-TRACE can support all these requests without having to send two different transmissions. This is the main advantage of MMC-TRACE over single-rate multicasting.

The performance of multi-rate multicasting can be calibrated according to the specific needs of different scenarios by modifying the multi-rate broadcast parameters such as the number of rates available, type of modulation, and power allocation between the rates. Utilizing a non-uniform M-ary constellation (see Section 4.3) for multi-rate multicasting requires some straight-forward modifications of the MMC-TRACE protocol described here to account for $log_2 M$ different potential paths.

4.7 Multi-rate Multicasting Simulations

In this section, we provide simulation results to obtain a better understanding of the trade-offs multi-rate multicasting introduces. As described previously in Section 4.6, we modified the MC-TRACE protocol so that it includes support for multi-rate broadcasting. This modified version is named MMC-TRACE, standing for Multi-rate Multicasting through Time Reservation using Adaptive Control for Energy efficiency. In all our simulations MMC-TRACE is used with the parameters shown in Table 4.6.

We performed extensive simulations using the network simulator-2 (ns-2) in order to show the benefits of multi-rate multicasting over regular multicasting. We simulated conversational voice coded at two different rates, namely the low rate ($13Kbps$) and the high rate ($26Kbps$). The high rate corresponds to two (superposed) voice packets per superframe and can be decoded by nodes within $R_1 = 150m$, whereas the low rate corresponds to one voice packet per superframe and can be decoded by nodes within $R_2 = 250m$. The channel rate is set to 2 Mbps and the standard IEEE 802.11 physical layer is employed for MC-TRACE. MMC-TRACE utilizes modified versions of the standard IEEE 802.11 physical layer with a channel rate of 2M symbols/sec in order to support different modulation techniques. The network has 256 nodes, moving within a 1 km by 1 km area for 100 seconds according to the Random Way-Point (RWP) mobility model with node speeds chosen from a uniform distribution between

Table 4.6: Simulation Setup

Acronym	Description	Value
–	Simulation Area	1000m x 1000m
T_S	Simulation Time	100s
–	Number of Repetitions	10
–	Node Mobility	Random Way-Point
T_{SF}	Superframe duration	61.5ms
N_F	Number of frames	7
N_{DS}	Number of data slots per frame	14
N_C	Number of cont. slots per frame	15
–	Data packet size	110B
–	Header packet size	36B
–	All other control packet size	10B
T_{ds}	Packet drop threshold @ source node	61.5ms
T_{di}	Packet drop threshold @ intermediate nodes	250.0ms
T_{VF}	Voice packet generation period	61.5ms
D_{CS}	Carrier Sense range	507m

0.0 m/s and 5.0 m/s. The source node is stationary and located in the middle of the network. simulations are repeated 10 times to obtain more statistically accurate results. In our plots, error bars, which show the standard deviation of each data point, are also provided.

4.7.1 Basics of Muti-rate Multicasting

In this section we perform simulations to illustrate the capabilities of multi-rate multicasting in an end-node driven branch formation scenario. Multicast group members can have two different requests, which lead to two different throughput-delay performances. First, all the members request high-rate data packets to be forwarded to them.

Second, low-delay packets are prioritized by all group members. According to these two different requests from the end nodes, two possibly different branches are formed between each member and the source, much like the ones illustrated in Figures 4.15-a and b. We compared the results obtained by MMC-TRACE to that of MC-TRACE with different modulation schemes while keeping the energy per symbol (ε_s) constant.

First, we use binary PAM (BPAM), shown in Figure 4.3, as the initial constellation employed by MC-TRACE. In this case, according to Equation 4.1, the same energy per symbol ε_s results in a larger noise margin than both the non-uniform QAM margins. Therefore, Equation 4.4 tells us that MC-TRACE with BPAM results in a larger transmission range than *250m*.

Second, we double the rate of transmission of MC-TRACE by using 4-ary PAM and regular QAM constellations. These constellations provide high-rate transmission with a single transmission range that is smaller than *250m*. Again, we reach this conclusion by using Equations 4.1-4.5. As a result of the reduced transmission range, network connectivity is relatively reduced in the case of 4-ary PAM and QAM when compared to that of BPAM.

The first set of simulations deal with the throughput characteristics of multi-rate multicasting. Figure 4.16-a shows the throughput comparison between MMC-TRACE and regular MC-TRACE employing BPAM, 4-ary PAM and uniform QAM. Throughput is doubled for MMC-TRACE when multicast member nodes request high-rate packets. On the other hand, in order to double the throughput of MC-TRACE, the modulation scheme has to be changed.

In addition to these results, Figure 4.16-b illustrates the trade-off presented by multi-rate multicasting. Through a single transmission, we are able to achieve two different throughput-delay characteristics, which could only be achieved by two transmissions with different modulation schemes using the conventional MC-TRACE protocol. During these simulations we kept the transmit energy per symbol (ε_s) constant in order to show the capabilities of multi-rate broadcasting without sacrificing much extra energy consumption (see Figure 4.16-c).

The number of packets transmitted from the MAC layer in these simulations has increased in the multi-rate MC-TRACE case when nodes prioritized high-rate packet reception. Figure 4.16-d shows the variation in the number of packets rebroadcasted

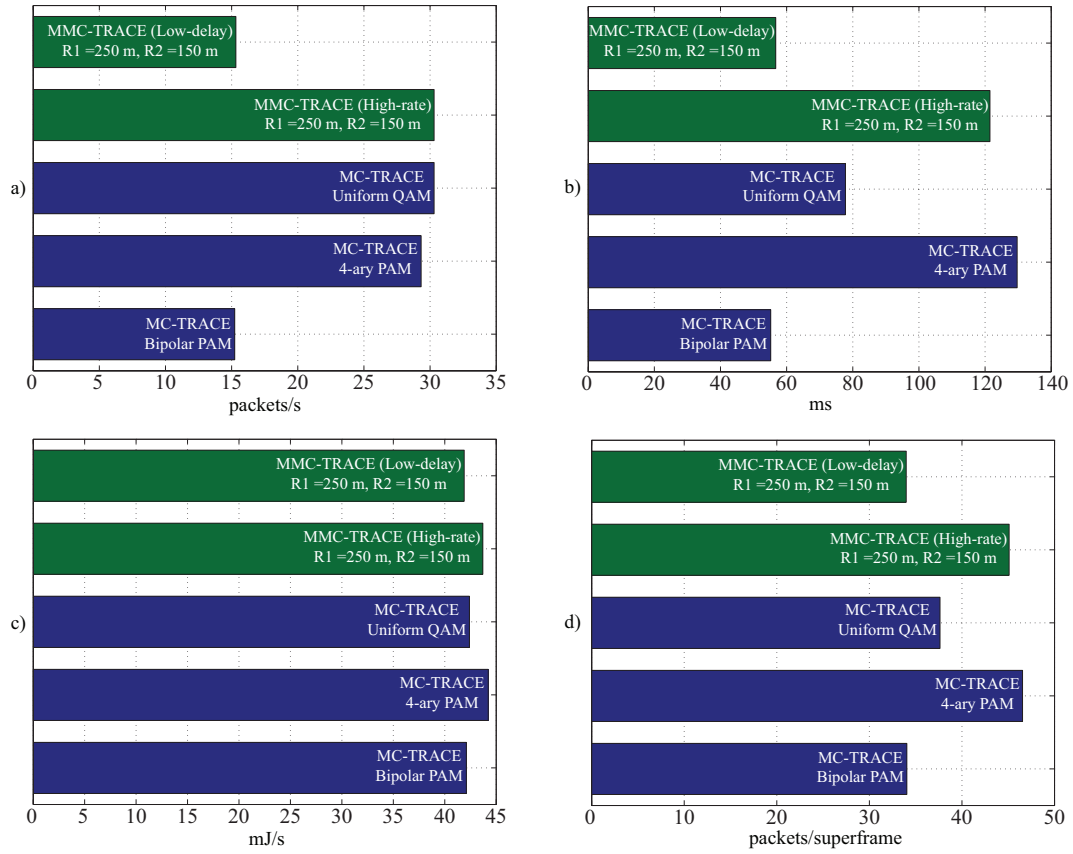


Figure 4.16: **a)** Number of packets received per multicast member per second. **b)** Average multicast packet delay. **c)** Average energy dissipation per node per second. **d)** Number of packets rebroadcasted per super frame (MAC layer).

as the protocol, modulation and priorities change. In MC-TRACE with QAM, 4-ary PAM and MMC-TRACE with high-rate priority, the number of packets rebroadcasted is larger than that of MC-TRACE with BPAM and MMC-TRACE with low-delay priority due to the reduction in the effective transmission range as the noise margin provided by the corresponding constellation is reduced. This forces data packets to travel through multicast branches that involve more hops, hence more rebroadcasts.

These figures illustrate MMC-TRACE's ability to provide a trade-off between delay and throughput through a single transmission without having to change the constellation diagram. For two different requests, different multicast branches are formed between the members and the source. These branches can only be supported by MMC-TRACE

without having to send two different transmissions. This is the main advantage of MMC-TRACE over single-rate multicasting.

4.7.2 Delay-Throughput Trade-off in MMC-TRACE

In order to clearly show this unique ability of MMC-TRACE, we performed a set of simulations where half of the multicast members request high-rate whereas the other half demand low delay packets. In Figure 4.17-a, the distribution of high and low rate and total packets are given. The delay values for these two groups of nodes are shown in Figure 4.17-b. We can conclude that multicast member nodes that request high-rate packets sacrifice packet latency in order to double the delivered throughput. On the other hand, nodes requesting low delay packets have a reduced latency but half the throughput of the high-rate requesting nodes.

Note that there are some cases where a multicast member node, due to mobility, gets disconnected from its branch and attached to a branch through an upstream node that has to rebroadcast high-rate data. This might happen when the upstream node in question has more than one multicast member node as its downstream nodes and has both requests (high-rate and low delay) imposed on it. In our simulation we favor the high-rate request over the low delay request when there is such a conflict. Therefore, in these cases a multicast member node might receive high-rate packets even if it has requested low-delay packets.

MC-TRACE is one of a number of cross-layer energy efficient routing protocols that provide better throughput, energy efficiency and latency when compared to simple flooding based protocols [23]. The results presented in this section show that highly coordinated multicasting protocols can be modified to provide multi-rate services without sacrificing the protocol integrity. This conclusion is based on the fact that MMC-TRACE has almost identical energy dissipation behavior as MC-TRACE, and the fact that MMC-TRACE is able to reproduce the throughput-delay characteristic of regular MC-TRACE [23].

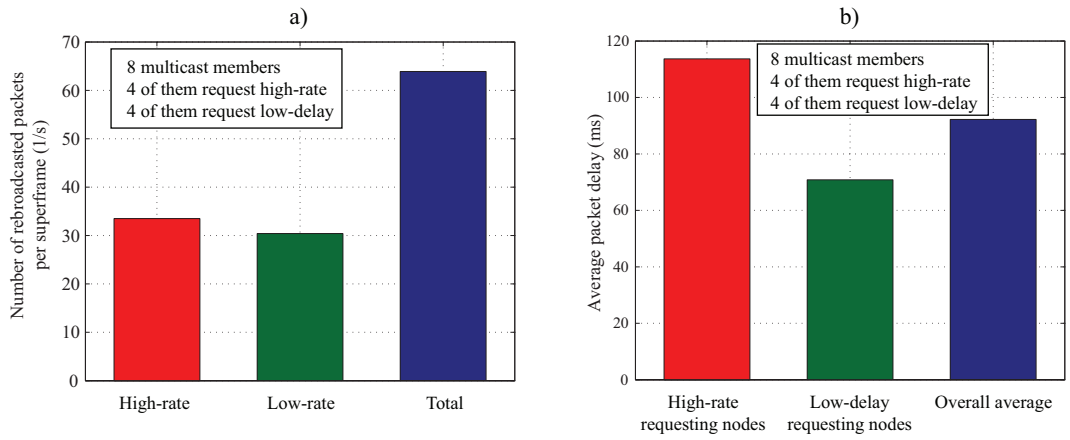


Figure 4.17: **a)** Number of packets rebroadcasted per superframe. Summation of the high-rate and low-rate packets gives the total number of packets transmitted from the MAC layer. **b)** Average packet delay values for both high-rate and low delay requesting nodes along with the overall packet delay.

4.7.3 Effects of Multicast Member Position

In this set of simulations, we aim to show the burden imposed on the network as a function of the distance between the source and a member node when the multicast member node demands high-rate data. There are again 256 nodes, five of which are multicast member nodes. Among these five, only one member requests high-rate packets at a time while we observe the change in the traffic, delay and energy consumption of the network. Moreover, we repeat this simulation as we vary the distance between the source and the member node that requests the high-rate data. Since all of the nodes except the source node are mobile, we focus on the average distance between the source and member node.

Figure 4.18-a shows the types and number of packets transmitted by all the nodes in the network versus the distance between the source and the multicast member requesting the high-rate information. The overall packet traffic increases as the member node moves further away from the source. This trend is dictated by the high-rate packet traffic caused by the high-rate requesting member node. The number of low-rate packets transmitted per superframe decreases because of the fact that the average distance between the source and the remaining four low-delay requesting nodes also decreases

as we keep the overall average distance between the source and the multicast members constant.

Figures 4.18-b and 4.18-c illustrate the delay and energy consumption as we change the average distance between the source and multicast member node. The average packet delay values are a direct result of the increased packet delay of the high-rate requesting node. We have also observed that the high-rate requesting node may affect the packet delay of the low-delay requesting nodes. This happens if a low-delay requesting node attaches itself to a high-rate multicast branch after a mobility induced tree breakage. The energy consumption, however, is not affected by these changes and stays almost constant as the multicast member node, requiring the high-rate data, moves farther away from the source node.

These results imply that multi-rate multicasting can support any high-rate requesting multicast member regardless of its position and without too much impact on the rest of the network. Moreover, this support causes only a very slight increase in the energy consumption due to the occasional increased traffic, which is caused by the high-rate request from the multicast member.

4.7.4 Effects of Node Density

In these simulations we investigate the effect of node density on the performance of multi-rate multicasting. We keep the simulation area fixed at 1 km by 1 km and vary the number of nodes in the network. The number of multicast members is kept constant at 4, and all the nodes are mobile except the source node, which is located in the middle of the simulation area. We provide the results obtained by using the regular MC-TRACE with binary PAM to be able to observe the effects of introducing multi-rate support on the characteristics of the MC-TRACE protocol. Figures 4.19-a and 4.19-b illustrate the packet delivery ratio (PDR) and the packet delay of the multicast network as the number of nodes in the network changes from 50 to 300. All four multicast members request high-rate packets, and we forced all the nodes in the network to forward high-rate packets in order to capture the deliverability of high-rate packets for different node densities.

Note that packet reception radii, R_1 and R_2, are kept constant at $250m$ and $150m$, respectively. Therefore, when the number of nodes in the network decreases, the net-

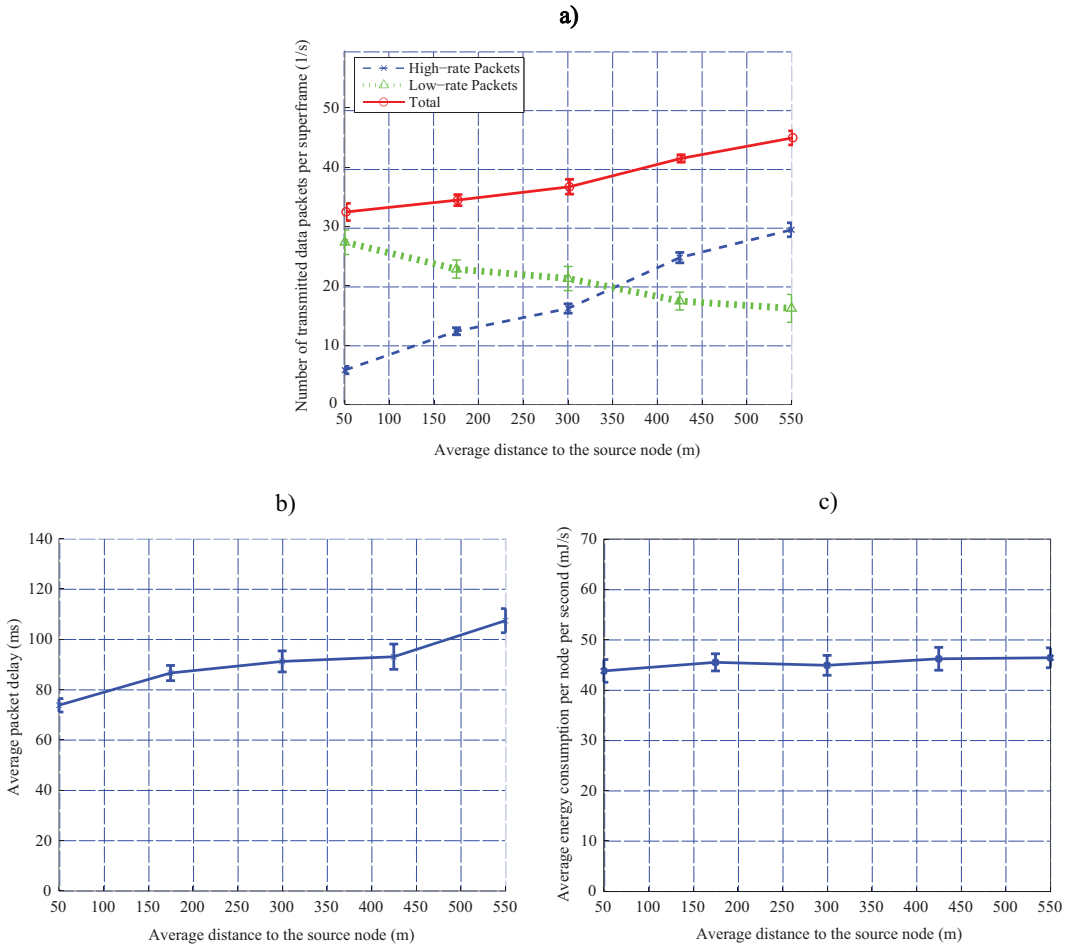


Figure 4.18: **a)** Number of packets transmitted per superframe versus the distance between the source and the multicast member requesting high data rate. **b)** Average packet delay versus the distance between the source and the multicast member requesting high data rate. **c)** Average energy dissipation per node per second versus the distance between the source and multicast member requesting high data rate.

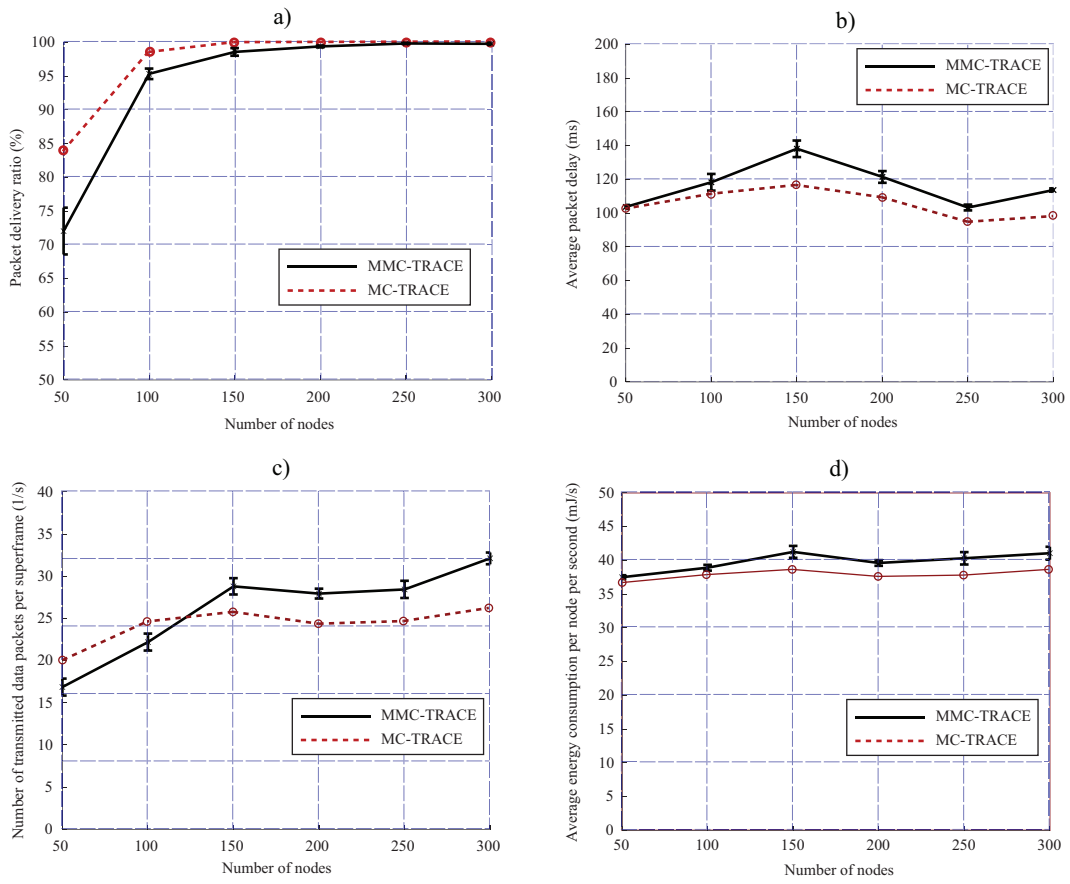


Figure 4.19: **a)** Packet delivery ratio versus the number of nodes in the network. The number of multicast members is 4. **b)** Average packet delay at the MAC layer versus the number of nodes in the network. The packet drop threshold at intermediate nodes is $T_{di} = 250ms$. **c)** Number of transmitted packets per superframe versus the number of nodes in the network. The number of multicast members is 4. **d)** Average energy consumption per node per second versus the number of nodes in the network. The number of multicast members is 4.

work connectivity gradually goes down for a given fixed transmission range. As the connectivity of the network decreases, data packets either travel longer routes that involve a larger number of hops or are dropped because some part of the network is disconnected from the rest. Moreover, increasing packet delay causes the intermediate nodes to drop packets with delay values higher than the packet drop threshold T_{di}. This results in decreased delay values, since only packets with delay values below the limit are being delivered. As another consequence, the throughput of the network starts to decrease as a result of preemptively dropping packets with high delays. Figure 4.19-c, which shows the number of packets transmitted per superframe, illustrates this fact.

The energy consumption values are presented in Figure 4.19-d. The energy efficiency of the TRACE protocol keeps the energy consumption almost the same for all node densities. The slight variation in the curve follows the transmission traffic presented in Figure 4.19-c.

Before we compare these results to that of regular MC-TRACE, it is important to note that for a fixed energy per symbol the transmission range for MC-TRACE with BPAM is larger than both transmission ranges of MMC-TRACE. As a result of this, MC-TRACE will be able to deliver the data packets with longer and fewer numbers of hops. This implies a better connectivity for the network and consequently, all the characteristics such as packet delay, throughput and number of transmitted data packets will improve. Therefore, we used $250m$ as the transmission range for our simulations with MC-TRACE to provide a meaningful comparison between MC-TRACE and MMC-TRACE. This results in smaller energy per transmitted symbol in MC-TRACE compared with MMC-TRACE.

We can observe that the throughput, packet delay, and energy consumption characteristics for both protocols follow the same trend with a little deviation between them. This shows that introducing multi-rate availability does not effect the general behavior of the MC-TRACE protocol under varying node densities. Therefore, the observed limitations, other than the ones caused by reduced transmission range, are mainly due to the original characteristics of MC-TRACE. Note that, for all the data points in the results, presented throughout the Section 4.7, standard deviation values stays within $\sim 3\%$ of the mean values.

4.8 Transmission Power Considerations in Multi-rate Broadcasting

Now that we have seen the advantages multi-rate multicasting can provide using a fixed set of radii R_1 and R_2, we turn our attention here to exploring the most energy-efficient means of exploiting this multi-rate capability. In this section we discover the bounds on the radii R_1 and R_2 as well as the optimal transmission radii R_1 and R_2 that minimize overall energy dissipation in a multi-rate multicasting scenario.

In ad-hoc wireless networks, each node should transmit with just enough power to guarantee the connectivity of the network and avoid causing unnecessary interference. This approach makes sure that the information can be conveyed between every pair of nodes within the network while causing minimal disruption to other traffic sharing the same channel. Gupta and Kumar showed [79] that the critical transmission range for a wireless network to achieve asymptotic connectivity is $O\left(\sqrt{(\frac{\log n}{n})}\right)$ if nodes are uniformly and independently distributed in a disk of unit area. This expression can be used to obtain the absolute lower limit for the transmission powers in multi-rate broadcasting and can be used to determine the lower bound for the transmission radius R_2, since we do not want to compromise the high-rate connectivity of the network. To determine the upper bound, we use the basics of superposed coding, which states that the power spent on sending the additional-rate information should be comparatively smaller than that of sending the low-rate information (*i.e.,* $R_1 > R_2$).

Given these bounds on R_1 and R_2, in this section we determine the transmission radii R_1 and R_2 for multi-rate multicasting that minimize the overall energy consumption in the network while ensuring that all nodes can obtain high-rate data traffic. We consider the non-uniform QAM constellation given in Section 4.2.

4.8.1 Optimization Process

For any constellation with fixed geometry, when the energy per symbol increases, the distance between the constellation points also increases. This results in a longer transmission radius and thus the number of retransmissions goes down. There exists an optimal tradeoff between the expected number of retransmissions and the transmit power to minimize the total energy dissipated to receive the data. At the physical layer, there

are two main components that contribute to energy loss in a wireless transmission, the loss due to the channel and the fixed energy cost to run the transmission and reception circuitry [80].

We begin by expressing the total energy dissipation in terms of the single-hop energy dissipation and the average number of hops between the source-destination pair. This type of investigation to find an energy-efficient optimal hop distance has been covered by Y. Chen *et al.* [81], where they analytically derive optimal hop distance for a particular radio energy dissipation model.

$$E_{total} = \lceil \frac{D}{d_{hop}} \rceil \cdot E_{hop}, \tag{4.6}$$

where D is the average distance and $\lceil \frac{D}{d_{hop}} \rceil$ is the average number of hops between each multicast source-member pair, and E_{hop} is the per-hop energy to transmit a packet. We will rewrite this equation without the ceiling function to achieve a continuous representation of E_{total}. However, at the end of this section we will come back to this representation and compare it with the continuous version.

E_{hop} can be written as follows:

$$E_{hop} = E_{send} + E_{rec}, \tag{4.7}$$

where E_{send} is the energy to send each packet and E_{rec} is the energy to receive each packet.

We can explicitly express these individual components in terms of the fixed energy costs and the constellation diagram variables that directly affect the transmit energy consumption.

$$E_{send} = k_{send} \cdot \varepsilon_s \times size + c_{send}, \tag{4.8}$$

$$E_{rec} = k_{rec} \times size + c_{rec}. \tag{4.9}$$

$\varepsilon_s = d_1^2 + d_2^2$ is the energy per symbol as we mentioned in Equation 4.1 in Section 4.2, $k_{send} \cdot \varepsilon_s$ gives us the cost of sending one byte and $k_{send} = \frac{8}{\log_2 M}$ is fixed for a given constellation (*i.e.,* for a given M, which is the number of bits in each symbol). $size$ is the number of bytes per packet. k_{rec} is also fixed and represents the cost of

receiving one byte. c_{send} and c_{rec} are fixed costs of sending and receiving a packet, respectively. Therefore, E_{hop} can be rewritten by using Equations 4.8 and 4.9.

$$E_{hop} = (k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c_{send} + c_{rec}. \tag{4.10}$$

At this point we modify the total energy expression (Equation 4.6) according to our multi-rate multicasting. This is due to the fact that we have two different transmission radii, hence two different hop counts depending on the type of packets transmitted and received by the nodes. Therefore the total energy consumption will be expressed in two parts, namely high-rate and low-rate energy consumption. Note that, since we are using the same constellation for both high and low-rate transmissions, E_{send} and E_{rec} stay the same for both radii, and only the hop count will change.

Before we proceed any further, we need to include the coexistence of two data rates into our analysis. In order to do this, we introduce a new parameter, γ, which expresses the ratio of the high-rate data requesting multicast members to all multicast members. Consequently, $1-\gamma$ represents the ratio of the low-rate requesting members to all multicast members. If γ is close to zero, there is little or no high-rate traffic and a γ close to one means almost all the traffic present in the network is high-rate. Either case is not favorable for multi-rate multicasting and should be handled by single-rate conventional multicasting.

$$\begin{aligned} E_{total} = &\frac{D}{R_1} \cdot (1-\gamma) \cdot [(k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c_{send} + c_{rec}] \\ &+ \frac{D}{R_2} \cdot \gamma \cdot [(k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c_{send} + c_{rec}] \end{aligned} \tag{4.11}$$

Due to the fact that both rates are transmitted in a superposed manner in a single transmission, we use the same fixed per-hop hardware cost for both high-rate and low-rate traffic. Our goal is now to find the optimum values for R_1 and R_2 that minimize Equation 4.11 for a given ε_s and γ. We start by deriving a relationship between the two transmission radii, assuming that the variation in received signal power over distance is due to path loss only.

$$P(R) = P_t \cdot \frac{\alpha}{R^n}. \tag{4.12}$$

In Equation 4.12, α is a constant that stands for wavelength of the signal, antenna gains, speed of light, and possibly for other constants associated with the propagation model. We can write a similar equation in terms of transmitted energy per symbol and received energy per symbol at distance R. The transmission power can be written in terms of ε_s as:

$$P_t = \varepsilon_s \cdot \frac{symbols}{sec}. \tag{4.13}$$

Therefore, the effect of the propagation on transmit energy per symbol ε_s can be expressed by combining Equations 4.12-4.13.

$$\varepsilon(R) = \varepsilon_s \cdot \frac{\alpha}{R^n}. \tag{4.14}$$

Now with the help of Equations 4.1 and 4.14, we rewrite the $\varepsilon(R)$ in terms of the constellation diagram variables d_1 and d_2 (see Figure 4.3).

$$\begin{aligned} \varepsilon(R) &= (\varepsilon_1 + \varepsilon_2) \cdot \frac{\alpha}{R^n} \\ &= (d_1^2 + d_2^2) \cdot \frac{\alpha}{R^n}. \end{aligned} \tag{4.15}$$

This equation assumes that the QAM carriers are orthogonal and only Additive White Gaussian Noise (AWGN) is considered. Moreover, we can also write,

$$\varepsilon_1(R) = d_1^2 \cdot \frac{\alpha}{R^n}, \tag{4.16}$$

$$\varepsilon_2(R) = d_2^2 \cdot \frac{\alpha}{R^n}. \tag{4.17}$$

At distances R_1 and R_2, $\varepsilon_1(R_1)$ and $\varepsilon_2(R_2)$ drop down to a limiting value below which we can not decode the encoded bits correctly. Therefore, we can write $\varepsilon_1(R_1) = \varepsilon_2(R_2) = \varepsilon_{low}$, where ε_{low} is the lowest energy level at which the bits are still decodable. In other words, according to the values of d_1 and d_2 from the constellation diagram, there are two different distances R_1 and R_2 at which ε_1 and ε_2 drop down to the same energy level ε_{low}.

$$\begin{aligned} \varepsilon_s &= d_1^2 + d_2^2 \\ &= \varepsilon_{low} \cdot \frac{R_1^n}{\alpha} + \varepsilon_{low} \cdot \frac{R_2^n}{\alpha} \\ &= \frac{\varepsilon_{low}}{\alpha} \cdot (R_1^n + R_2^n). \end{aligned} \tag{4.18}$$

In our optimization process we assume that ε_s is fixed due to possible power consumption restrictions on the transceiver or battery limitations. For a fixed ε_s, we can obtain a relationship between R_1 and R_2 by using Equation 4.18. This relationship can be utilized to simplify the E_{total} expression (*i.e.,* Equation 4.11).

$$R_2 = \sqrt[n]{\frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}} - R_1^n} \tag{4.19}$$

Therefore, we can write E_{total} as:

$$E_{total} = (\frac{D}{R_1} \cdot (1-\gamma) + \frac{D}{\sqrt[n]{\frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}} - R_1^n}} \cdot \gamma) \cdot [(k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c']. \tag{4.20}$$

The optimization of the transmission radii to minimize the total energy consumption can be carried out by taking the derivative of E_{total} over R_1 and setting $\frac{dE_{total}}{dR_1} = 0$.

$$\begin{aligned} \frac{dE_{total}}{dR_1} &= (-\frac{D}{R_1^2} \cdot (1-\gamma) + \frac{D \cdot R_1^{n-1}}{(\frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}} - R_1^n)^{\frac{n+1}{n}}} \cdot \gamma) \cdot [(k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c'] \\ &= -\frac{D}{R_1^2} \cdot (1-\gamma) + \frac{D \cdot R_1^{n-1}}{(\frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}} - R_1^n)^{\frac{n+1}{n}}} \cdot \gamma = 0 \end{aligned} \tag{4.21}$$

Solving Equation 4.21 provides the optimum R_1 that minimizes E_{total}.

$$R_1 = \sqrt[n]{\frac{A}{1 + (\frac{\gamma}{1-\gamma})^{\frac{n}{n+1}}}} \tag{4.22}$$

where $A = \frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}}$. We can also find the optimum value for R_2 by using Equation 4.19.

$$R_2 = \sqrt[n]{A - \frac{A}{1 + (\frac{\gamma}{1-\gamma})^{\frac{n}{n+1}}}} \tag{4.23}$$

Note that neither of the optimum values for R_1 or R_2 depend on the constant hardware cost. In other words, the values of k_{rec}, $size$, c_{send} and c_{rec} do not play any role

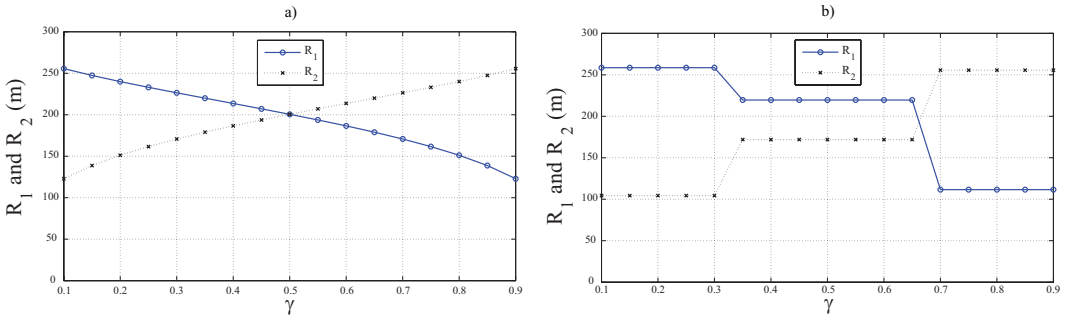


Figure 4.20: **a)** R_1 and R_2 versus γ according to Equations 4.22 and 4.23 for a given A and $n = 2$. **b)** R_1 and R_2 versus γ obtained from in Equation 4.24 for a fixed A and $n = 2$.

in minimizing the total energy consumption. This result is understandable since we are trying to decide how much of the fixed energy should be given to each transmission radius.

4.8.2 Results

In this section, we investigate how R_1 and R_2 change when we vary ε_s and γ. We also compare these results, obtained from Equations 4.22 and 4.23, with the ones that could be obtained by keeping the ceiling function in the E_{total} expression. We discuss the applicability of multi-rate coding according to the results we obtain from both approaches.

In this section, we use the energy model described in [82] and make our calculations for fixed energy according to the detailed measurements for the Lucent WaveLAN IEEE 802.11b wireless card [83]. Measurements indicate that the $Energy = m \times size + b$ model provides an accurate representation for wireless modems, where m and b are constants determined by the specific card and $size$ is the packet size. The free space model is used for the calculations performed in the rest of this section.

Figure 4.20-a illustrates the optimum values of the (R_1, R_2) pair, which minimize the total energy consumption, plotted against γ, the ratio of high-rate requesting multicast members to all multicast members in the network. Here, we keep $A = \frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}}$ constant and vary γ. The symmetrical relationship between the radii around $\gamma = 0.5$ tells

us that if we need to employ multi-rate coding and also want to achieve minimal energy consumption, we can only achieve this by keeping the percentage of high-rate requests below 50%. Above $\gamma = 0.5$ we cannot satisfy the minimum energy consumption and multi-rate broadcasting goals simultaneously since multi-rate broadcasting clearly requires that $R_1 > R_2$.

We mentioned earlier that the ceiling function in Equation 4.6 was dropped to allow an easier optimization process. If we keep the ceiling function and follow the same steps, we end up with a different set of equations.

$$E_{total} = (\lceil \frac{D}{R_1} \rceil \cdot (1-\gamma) + \lceil \frac{D}{\sqrt[n]{\frac{\varepsilon_s \cdot \alpha}{\varepsilon_{low}} - R_1^n}} \rceil \cdot \gamma) \cdot [(k_{send} \cdot \varepsilon_s + k_{rec}) \times size + c']. \quad (4.24)$$

Figure 4.20-b shows the change in the (R_1, R_2) pair, which was calculated according to Equation 4.24, minimizing the total energy consumption. As γ increases from 0.1 to 0.9, the optimum value of R_2 maintains its non-decreasing behavior we observed in Figure 4.20-a. However, the discontinuous nature of the ceiling function dominates the resulting values for R_1 and R_2. Note that in this case we need $\gamma \leq 0.65$ in order to satisfy the requirements of multi-rate coding.

Energy per symbol (ε_s) is another parameter that affects the optimum values of the transmission radii. Although ε_s is primarily determined by the hardware capabilities, we provide results for a range of values, which clearly demonstrates how the optimum transmission radii change with ε_s. Figure 4.21 shows the results obtained from the two different versions of the same equation. Equations 4.20 and 4.24 are almost the same except for the ceiling function in Equation 4.24. The results, however, show perceptible differences at each γ value.

As γ increases (*i.e.,* the number of multicast member nodes requesting high-rate increases), the intervals of the (R_1, R_2) pair satisfying the multi-rate coding approach decreases. The behavior of the plots obtained from Equation 4.24 are mainly dominated by the $\lceil \frac{D}{R_1} \rceil$ term in the equation. For a given γ, at a certain energy per symbol value, there might be more than one (R_1, R_2) pair that minimizes the total energy consumption. In that case, we choose the pair with the largest R_1. This leads to the jumps shown in the figures in the second column of Figure 4.21. When we look at the results obtained from Equation 4.24, we observe that there is more room in choosing the transmission

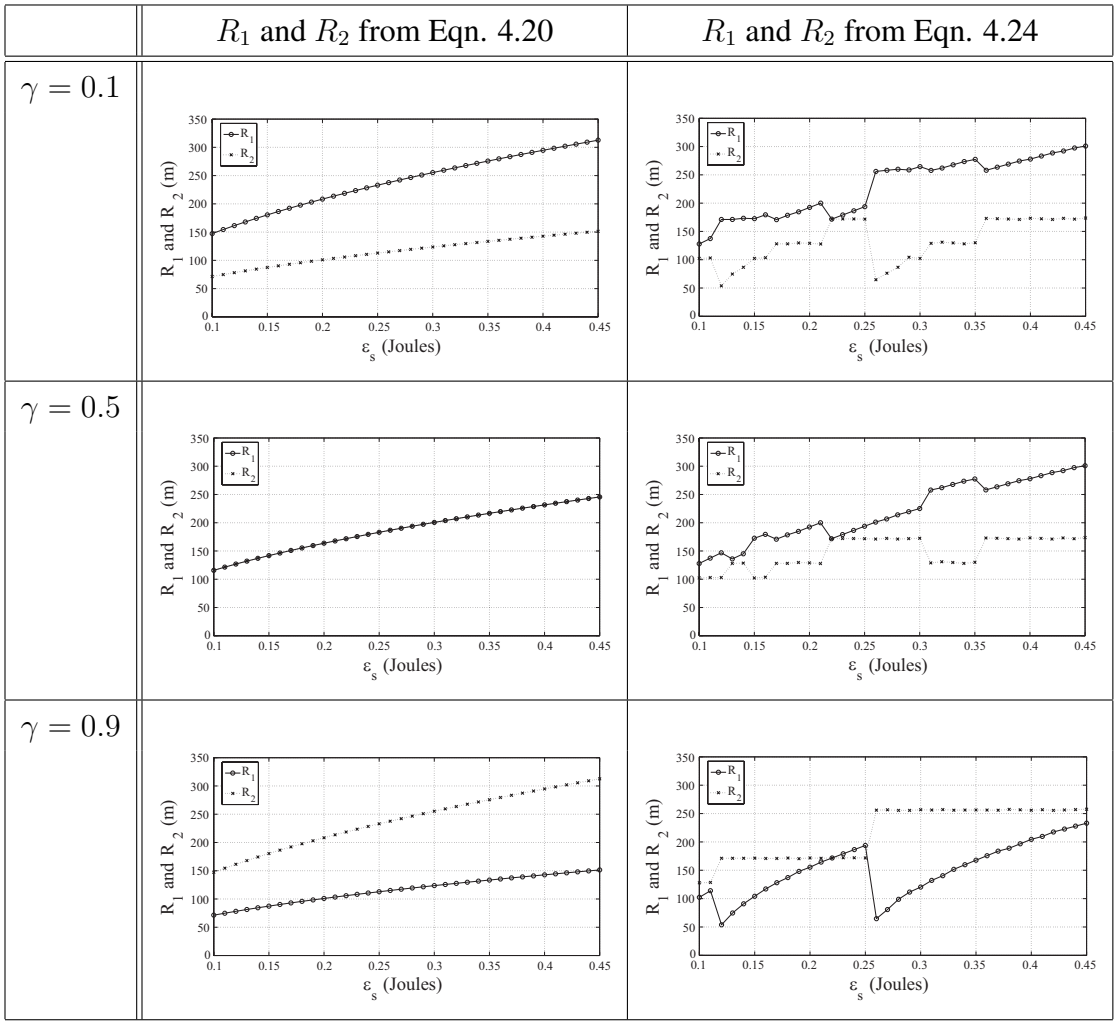


Figure 4.21: R_1 and R_2 vs ε_s for $\gamma = 0.1$, $\gamma = 0.5$, and $\gamma = 0.9$.

radii within the $R_1 > R_2$ limit. On the other hand, the results obtained through Equation 4.20 indicate that for $\gamma \geq 0.5$ there is no way of achieving the minimum energy dissipation goal if we want to employ multi-rate coding.

4.9 Summary

In this chapter we summarized the idea of multi-rate broadcasting and showed the performance gains achievable by using multi-rate broadcasting in network-wide broadcasting [84]. We took the idea of multi-rate broadcasting one step further and utilized it to provide a flexible throughput-delay trade-off in a multicasting scenario [85]. This

trade-off between throughput and delay, when available, can be exploited within the network according to the different requirements of the different members of a multicast group. We investigated the bounds on the selection of transmission radii in multi-rate multicasting, and we also showed through analysis that for a given scenario, one can minimize the total energy consumption of the network by choosing the transmission radii properly.

The need for multi-rate multicasting will continually increase as the number of subscription based wireless services, such as mobile TV broadcasting and news feeds, offered to consumers increase. In addition to the variety of services offered, the different communication capabilities and needs of the continuously improving hand-held mobile devices must be supported by MANETs. We believe that multi-rate multicasting will provide an efficient solution to satisfying the varied needs of all these devices through a single multicast transmission, thereby reducing energy dissipation and bandwidth usage.

In network-wide broadcast or multicast networking, where the routing protocol creates a set of links from cluster heads to gateways to cluster heads until the destination node's cluster head is reached, if a gateway suddenly loses connectivity with one of its cluster heads (due to channel noise or node mobility), the entire route may be broken. We need to develop approaches to minimize the impact of such broken links. Therefore, in the next chapter, we research different adaptive solutions to this problem, including the use of backup routes and/or mesh trees to add reliability to the broadcast and multicast trees created through NB-TRACE and MC-TRACE, respectively.

Chapter 5

Adaptive Mesh Networking

A growing concern with protocols for MANETs is whether or not end-to-end reliability can be provided. The MH-TRACE framework automatically handles reliability at the MAC layer, ensuring that all nodes within the receive range of a node can receive that node's transmission and a node can contend for channel access from any clusterhead in its transmit range. Thus the MH-TRACE protocol has an inherent degree of reliability. However, in Chapter 3 we demonstrated that channel errors can have a severe effect on the performance of MH-TRACE. In this study, we test the behavior of the network-wide broadcast and multicast protocols of TRACE (*i.e,* NB-TRACE and MC-TRACE) under channel errors. We present a mesh networking inspired approach that utilizes variable redundancy in the connectivity of the network to increase the stability of these protocols. We perform simulations to show that adaptive redundancy can indeed overcome some of the performance loss of the TRACE family of routing protocols under lossy channel conditions.

5.1 Motivation

The objective of a multicast routing protocol for MANETs is to support the dissemination of information from a sender to all multicast group members while trying to use the available bandwidth efficiently in the presence of frequent topology and channel condition changes. In MANETs, node mobility and the fact that wireless links are more prone to transmission errors result in higher packet drop probability when compared to wired

networks. Therefore, multicast routing protocols that provide route redundancy (*i.e.,* routing packets along multiple paths from source to receivers), typically outperform multicast routing mechanisms that offer no redundancy. However, increased redundancy can cause significant overhead in a resource-constrained MANET, even though it provides higher packet delivery ratios.

When channel conditions are good (*i.e.,* link reliability is high), having larger redundancy in the network does not significantly increase packet delivery ratio (PDR). However, when channel conditions get worse, having greater redundancy does have a considerable impact on the packet delivery ratio. Our proposed multicasting mechanism with adaptive redundancy varies the redundancy in the network according to the local packet reception history. Redundancy is managed locally and adaptive behavior is controlled by the multicast members (*i.e,* receivers). A multicast member node controls the amount of redundancy depending on the stability of its upstream node and incoming traffic. If the upstream node is volatile and data packets do not arrive in order, the number of upstream nodes is increased. If a consistent upstream node exists and data packets continue to arrive in order, the number of upstream nodes is decreased. This process is repeated by each relay to provide a greater number of non-overlapping redundant branches between the source and the multicast member.

In this study, we first demonstrate the vulnerability of NB-TRACE under channel errors, and employ a mesh networking inspired approach that utilizes redundancy in the connectivity of the network to increase the stability of NB-TRACE. However, redundancy has a cost associated with it, both in terms of energy as well as bandwidth due to the forwarding of a greater number of packets. We implement our approach of varying adaptive redundancy into the NB-TRACE protocol and perform simulations to show that adaptive redundancy can indeed overcome some of the performance loss of NB-TRACE under lossy channel conditions.

We also implement our adaptive redundancy algorithm on MC-TRACE, and compare the performance of adaptive redundancy multicasting against its non-adaptive version and a non-adaptive mesh-based routing protocol, On Demand Multicast Routing Protocol (ODMRP) [46]. Through ns-2 simulations, we show that adaptive redundancy can maintain high packet delivery ratios at higher protocol efficiency when compared to both non-adaptive mesh and tree-based multicast protocols.

5.2 The Need for Redundancy: Effects of Channel Errors

In this section, we utilize flooding with 802.11 and NB-TRACE to show the need for adaptive redundancy. The most important difference between the two approaches is the level of redundancy in the broadcast tree. Flooding with IEEE 802.11 can be considered as the extreme case in terms of redundant number of rebroadcasting nodes in a network. On the other hand, NB-TRACE specifically aims to remove any redundant branches from the broadcast tree in order to optimize energy savings and channel utilization.

To demonstrate the performance characteristics of these two extreme examples of network-wide broadcasting (NWB) protocols, we test NB-TRACE and flooding with IEEE 802.11 against increasing bit error rate (BER) levels. We ran simulations using the ns-2 network simulator [72]. We used a constant bit rate (CBR) traffic generator with UDP transport agent to simulate a constant rate voice codec at 32 Kbps. The source node is positioned in the middle of the network. The channel rate is set to 2 Mbps and the standard IEEE 802.11 physical layer is employed for both protocols. All the simulations are run with 80 nodes, moving within a 1 km by 1 km area for 100 seconds according to the Random Way-Point (RWP) mobility model with node speeds (S) chosen from a uniform distribution where, $0.0m/s < S \leq 5.0m/s$. Pause time is set to zero to avoid any non-moving nodes throughout the simulations. The simulations are repeated with the same parameters ten times, and the data points in the figures are the average of the ensemble. Acronyms, descriptions and values of the parameters used in the simulations are presented in Table 5.1.

Figure 5.1(a) presents the packet delivery ratios (PDRs) of NB-TRACE and flooding with IEEE 802.11 obtained from simulations as functions of BER. Minimum packet delivery ratio (PDR) achieved at any node is plotted against increasing BER, shown in Figure 5.1(b). The PDR of flooding with IEEE 802.11 is around 80%, even with low BERs, because of the fact that during flooding there are a huge number of data packet transmission attempts that result in collisions. However, as we increase the BER, we observe a steeper fall in both average and minimum PDRs of NB-TRACE when compared with that of flooding with IEEE 802.11. There are two mechanisms that decrease the throughput of NB-TRACE with increasing BER.

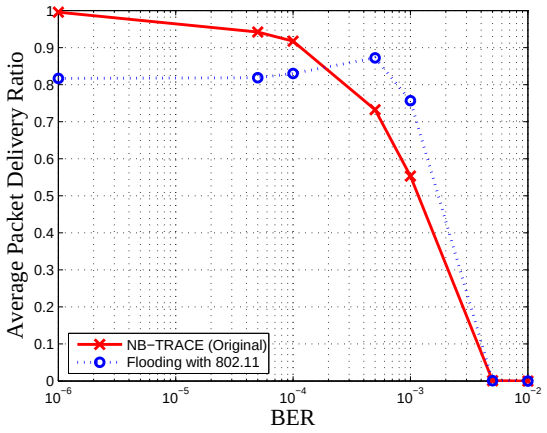
Table 5.1: Simulation Parameters

Acronym	Description	Value
T_{SF}	Superframe duration	25.056 ms
N_F	Number of frames	6
N_{DS}	Number of data slots per frame	6
N_C	Number of cont. slots per frame	10
N/A	Data packet size	110 B
N/A	Header packet size	5-17 B
N/A	All other control packet size	4 B
T_{drop}	Packet drop threshold	150 ms
D_{Tr}	Transmission range	250 m
D_{CS}	Carrier Sense range	500 m
P_T	Transmit power	600 mW
P_R	Receive power	300 mW
P_I	Idle power	100 mW
P_S	Sleep power	10 mW

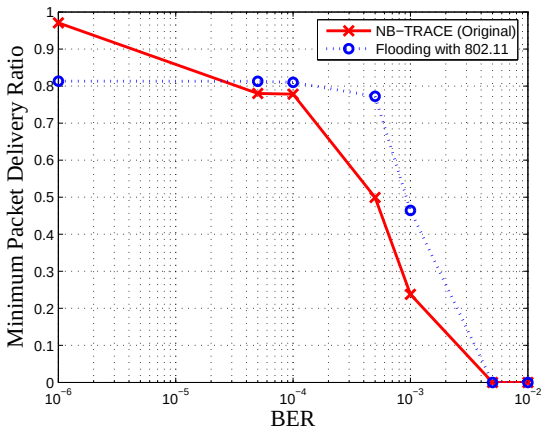
- More data packets are corrupted, which is also true for flooding with IEEE 802.11. However, as the PDR of NB-TRACE continuously decreases with increasing BER, simple flooding benefits from dropped data packets (*i.e.,* reduced data traffic) and manages to deliver better PDR values before the BER reaches 10^{-3}.

- The increase of corrupted control packets, especially header packets, results in unutilized data slots and broken branches in the broadcast tree, hence interrupted data packet flow for NB-TRACE, whereas in flooding with IEEE 802.11 this is not a problem due to the lack of control packets and the added redundancy provided by flooding.

This situation creates an interesting trade-off: while scheduling through header packets results in very high channel utilization and QoS in congested networks, it prevents nodes from channel utilization and maintaining its superior QoS performance in



(a) Average data PDR versus bit error rate (BER).



(b) Minimum data PDR versus bit error rate (BER).

Figure 5.1: Average and minimum PDRs versus BER.

high BER levels. When we examine the figures we see that NB-TRACE's minimum PDR values decrease faster than its average PDR values. This behavior demonstrates that the QoS provided by NB-TRACE is questionable under channel errors.

One of the most important advantages of NB-TRACE over flooding with IEEE 802.11 is its better energy efficiency due to the tree structure rather than the added redundancy of flooding. Average energy dissipation per node per second for NB-TRACE and IEEE 802.11 as a function of BER are presented in Figure 5.2. For BER $\leq 10^{-3}$, NB-TRACE energy dissipation is less than 20 % of the energy dissipation of flooding with IEEE 802.11. Flooding with IEEE 802.11 energy dissipation does not show

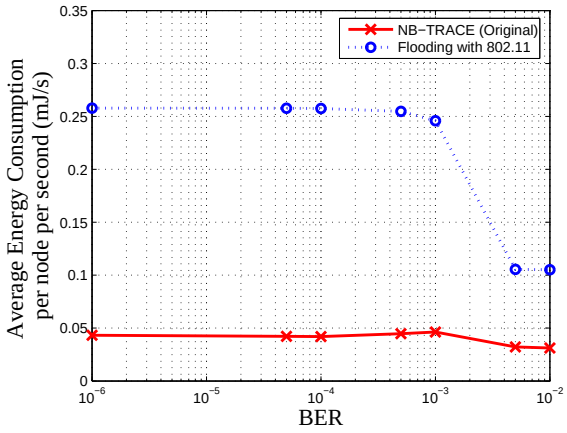
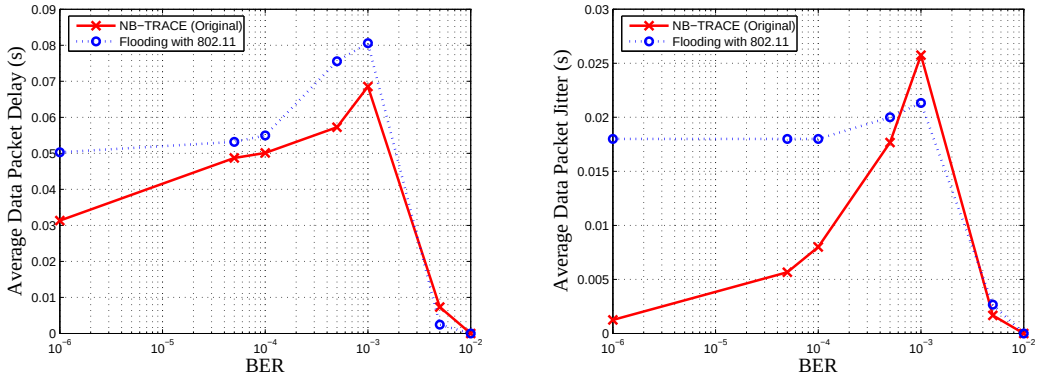


Figure 5.2: Average energy consumption per node per second versus BER.

any significant change with increasing BER due to the fact that the dominant energy dissipation terms in IEEE 802.11 are receive and carrier sensing, and they are not significantly affected by the BER. This is because the energy dissipated for receiving a non-corrupted packet and a corrupted packet is the same. Furthermore, in carrier sensing only the presence of the carrier is important, which is not affected by the BER level of the network. Packet transmissions, however, are related with BER level since in network wide broadcasting most of the data packets are routed from the source or an intermediate node. Therefore, for BER $\geq 10^{-3}$, flooding with IEEE 802.11 energy consumption approaches to NB-TRACE energy consumption, which is optimized by the explicit coordination between the nodes.

Figure 5.3 presents the average data packet delay and delay jitter for NB-TRACE and flooding with IEEE 802.11 as a function of BER. Data packets are dropped at the source node and intermediate nodes if the data packet delay exceeds 50 ms and 150 ms, respectively. NB-TRACE packet delay is lower than flooding with IEEE 802.11 packet delay at all BER levels due to the fact that in NB-TRACE the broadcast tree, formed to ensure QoS, together with the 150 ms packet drop threshold prevents packets from traveling longer routes. On the other hand, flooding with IEEE 802.11 has no control over the route a packet might follow. Therefore, if shortest path delivery fails because of channel errors and collisions, packets will be flooded through other and possibly longer routes.



(a) Average data packet delay versus BER.

(b) Average data packet jitter versus BER.

Figure 5.3: Average delay and jitter of data packets versus BER.

In terms of jitter performance, NB-TRACE suffers from the use of a time division multiple access scheme, which allows nodes to transmit only once in a superframe time (T_{SF}). Channel errors force NB-TRACE to form and reform branches in the broadcast tree, and this process disturbs the periodic behavior of the packet arrival times. As a result of this, jitter values suffer as the BER increases. The delay and jitter values of NB-TRACE and simple flooding tend to converge to similar values when BER $> 10^{-3}$.

These results show that the QoS performance of tree-based routing degrades as channel errors increase. On the other hand, the mechanisms that provide robust QoS performance for flooding under increasing channel errors cause poor performance in terms of bandwidth and energy efficiency when channel conditions are good (*i.e.*, at low BER levels). Consequently, both the tree-based and mesh-based routing protocols suffer from preset performance limitations that cannot adapt to varying channel conditions. Once again, this shows the need for adaptive redundancy, varying the amount of routing redundancy according to the channel conditions.

5.3 Adaptive Redundancy Considerations

In this section, we explore the limits of redundancy in routing both analytically and through simulations. We analyze the limits of adaptive redundancy for a given size of the network under varying bit error rates. In addition to this, we present the trade-

off between redundancy in the network and energy consumption. In the next section we will use this motivation to design an adaptive multicasting technique to obtain the advantages of both tree and mesh-based routing as the link conditions change.

5.3.1 Limits of Redundancy

We aim to achieve adaptive redundancy through varying the number of upstream nodes. When channel conditions are bad, more upstream nodes are needed to create a mesh-like routing that delivers data packets multiple times using different routes. When conditions are good, fewer upstream nodes are needed to create a tree-like routing that reduces the number of unnecessary routes between the source and destination pairs. The minimum amount of redundancy is achieved when there is a single route between each source-destination pair. The maximum level of redundancy can be considered as flooding where all routes between the source-destination pair are utilized. However, as we increase the number of upstream nodes (*i.e.,* transition from tree-based to mesh-based routing), we will reach to a point where increasing the redundancy in the network is either not possible or not worth the price we pay in terms of bandwidth and energy efficiency.

We can demonstrate the need for an upper limit to the number of routes with the help of an example given in Figure 5.4. For a given bit error rate BER, if we assume that a single bit error is enough to corrupt a data packet and cause it to be dropped, the probability of having at least one bit in error in a data packet of size L_D bits becomes,

$$P_{drop} = 1 - (1 - BER)^{L_D}. \tag{5.1}$$

In a real-time traffic scenario, where retransmissions are not utilized, the packet delivery ratio (PDR) of branch I in Figure 5.4 can be written in terms of the number of hops in the branch (N_{hop}).

$$PDR_1 = [(1 - BER)^{L_D}]^{N_{hop}}. \tag{5.2}$$

Following the same idea, we can rewrite the effective PDR of having k non-overlapping branches between the source-multicast member pair.

$$PDR_k = 1 - (1 - PDR_1)^k \tag{5.3}$$

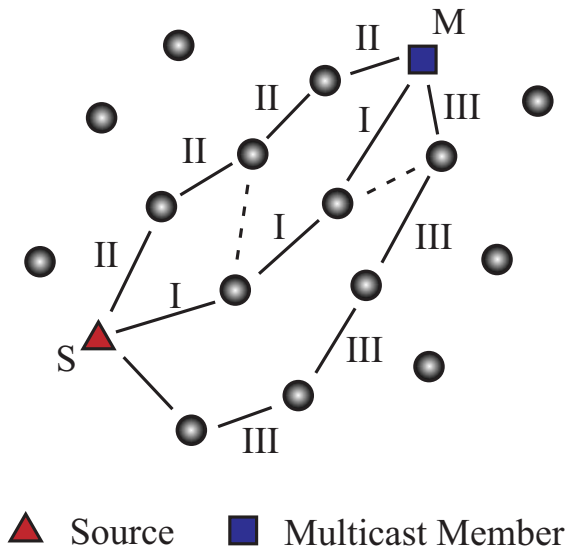


Figure 5.4: Illustration of multiple branches between a source-multicast member pair. Solid lines represent possible non-overlapping routes between the source and member. Dashed lines represent possible interconnections between the branches.

For $k = 3$ in Equation 5.3 and assuming the number of hops in all three branches is the same, we can calculate the effective PDR of the scenario illustrated in Figure 5.4. The existence of the interconnections between the nodes (*i.e.,* the dashed lines in Figure 5.4) do not contribute to the effective PDR since each node rebroadcasts the same data packet only once. Moreover, non-overlapping branches always will outperform partially overlapping routes. As Equation 5.3 suggests, increasing the number of non-overlapping branches will result in an increased effective PDR. However, as the number of branches increases, the number of nodes participating in the relay process also increases ($N_{hop} \times k$). This results in increased traffic, bandwidth usage, and energy consumption.

In Figure 5.5, we plotted the effective PDR equation (Equation 5.3) against increasing BER. As k increases, one can see that the increase in the effective PDR saturates. On the other hand, the amount of traffic generated and the energy consumption are linearly related with k. Therefore, it is necessary and wise to limit the number of non-overlapping branches in order to strike a well balanced trade-off between the amount of redundancy and energy consumption.

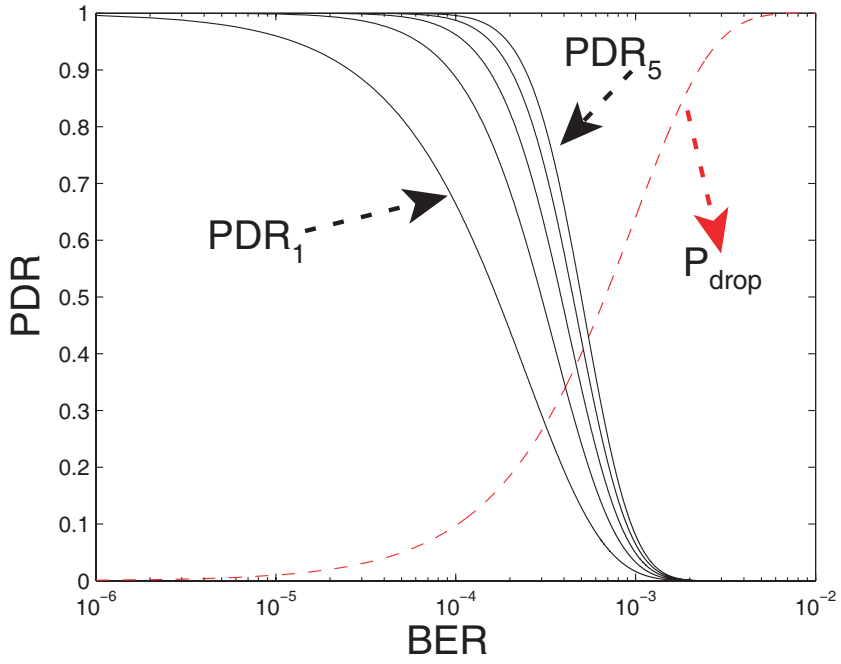


Figure 5.5: Solid lines are plots of Equation 5.3 for $k = [1, 5]$, $L_D = 128 \times 8 = 1024bits$, $N_{hop} = 4$, and varying BER. The dashed line is P_{drop} versus BER.

It is possible to further investigate these results to determine an optimum number of branches for a given set of constraints on bandwidth, energy consumption and PDR. However, our approach, as described in the next subsection, locally increases the number of upstream nodes, and resulting branches may be overlapping depending on the node distribution and density of the network.

5.3.2 The Redundancy vs. Energy Consumption Trade-off

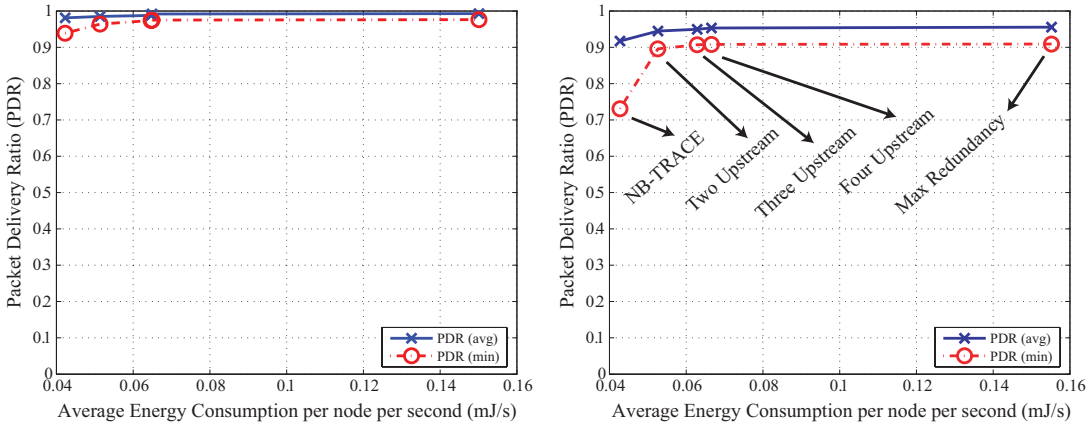
In this section, we introduce redundant branches into the tree formed by NB-TRACE in order to examine the trade-off between redundancy (with improved packet delivery) and energy consumption. In other words, as shown in Section 5.2, pruning all redundant branches in NB-TRACE reduces energy dissipation and improves channel utilization at the cost of making the NB-TRACE protocol more vulnerable to channel errors. Therefore, we would like to see the performance improvement possible by adding some redundancy into the NB-TRACE tree, which can be accomplished by in-

creasing the number of upstream nodes to be acknowledged. This number is set to one in NB-TRACE.

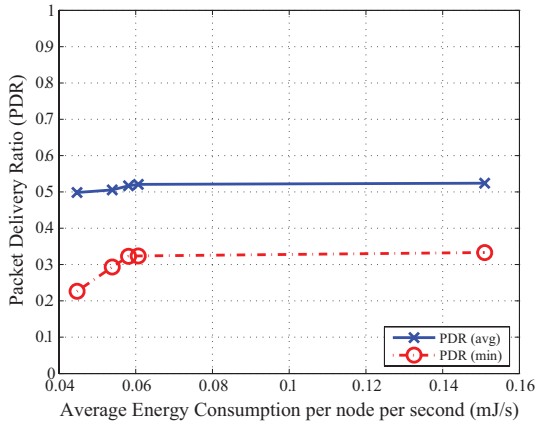
We simulate NB-TRACE with an increasing number of fixed upstream nodes, which are forced to rebroadcast data packets, as they are acknowledged as valid upstream nodes through the IS slots. In Figure 5.6, we present the PDRs as a function of varying channel errors. The first data points in each plot are obtained through the original NB-TRACE protocol. The second, third, and fourth data points correspond to NB-TRACE with two, three, and four upstream nodes, respectively. The last data point represents the maximum redundancy where we employ flooding with MH-TRACE, the MAC protocol used in NB-TRACE. The x-axis of these graphs represents the corresponding energy consumption for each level of redundancy.

For low BER levels (*i.e.,* BER = 10^{-5}) average PDR values do not differ much with increasing redundancy. At BER = 10^{-4} (Figure 5.6(b)), as the redundancy increases, minimum PDR values increase by $\sim 22 - 25\%$. At this level of BER, we start to observe noticeable increases in the average PDR as well ($\sim 4\%$). Figure 5.6(c) shows that when BER = 10^{-3}, nodes in the network fail to receive nearly half of the packets broadcasted by the source. Increasing redundancy, once again, improves the minimum PDR by $\sim 30 - 47\%$. The improvement in the average PDR values stays around $\sim 1 - 5\%$. Starting with BER = 10^{-3}, at higher BERs NB-TRACE control traffic gradually collapses. This results in limited channel access and crippled routing of data packets. We observe a similar behavior in the standard deviation values. Initially, at low BER levels, standard deviation values stay within $\sim 2\%$ of the mean values. However, as the BER increases, the standard deviation values reach up to $\sim 15\%$.

These results indicate that increasing redundancy can improve the QoS of NB-TRACE under channel errors. Moreover, this improvement is made possible by trading-off energy consumption with reliability through redundancy. As we introduced additional upstream nodes to increase the chance of packet delivery to the nodes in the network, in return, we achieved higher PDR and energy consumption values. Higher redundancy results in higher reliability because packets can be delivered even in the presence of links breaking and channel errors. However, in less dynamic environments, in which links break less frequently and channel errors are low, the additional redundancy may not be needed in terms of reliability and may significantly increase overhead



(a) BER = 10^{-5} (b) BER = 10^{-4}



(c) BER = 10^{-3}

Figure 5.6: Average and minimum PDRs versus energy consumption corresponding to the level of redundancy at given BERs.

and energy consumption. For example, in Figure 5.6(a) the rate of performance gain achieved by increasing redundancy (also energy consumption) is much less than that of Figure 5.6(b). Therefore, we propose that there should be an adaptive approach to balance the trade-off between reliability and level of redundancy in the network. In the next section, we present our adaptive redundancy approach.

5.4 The Adaptive Redundancy Approach

Although redundancy in a multicast mesh promises greater stability in the face of changing conditions or failure at single nodes [55, 86, 87], careful steps need to be taken in order to achieve the desired design goals and a balanced trade-off between the increased use of resources and performance improvement. In addition to this, when link breakages are low, even protocols that do not utilize redundancy have excellent packet delivery ratios [88]. Therefore, the additional cost incurred by redundant mesh based protocols is unnecessary and often wasteful. The next section presents our proposed method to adaptively vary the level of redundancy in a multicast mesh according to the channel conditions.

We consider constant bit rate (CBR) real-time data traffic where data packets need to be delivered within a pre-determined delay bound in the order they are transmitted (*e.g.,* real-time voice traffic). Although CBR traffic implies continuous transmission of data packets, our approach can support any type of real-time traffic including sessions with bursty real-time data transmissions and multiple sessions as well. This is due to the fact that with each data transmission source nodes indicate whether or not they are going to transmit again. Therefore, nodes can anticipate the packet arrivals and proactively decide on changing the level of redundancy locally. In other words, if there is a disruption in the incoming data traffic, a node can decide if this was due to a lost packet or because the particular session ended by keeping track of the transmission information obtained from the previously received data packets.

We assume nodes do not utilize any retransmission policy. Moreover, cooperative broadcasting, where receiving the same packet from different sources increases the chance of successful decoding, is not considered. We move from network-wide broadcasting to multicasting as a more general case for routing of real-time data in a mobile ad hoc network. Specifically we look at how the tree-based multicasting protocol MC-TRACE, as described in Section 2.1.3, can utilize adaptive redundancy to improve its performance under harsh channel conditions.

Before going into the details of our approach, we would like to point out that there are mechanisms within the MC-TRACE protocol that significantly improve system performance in the face of node mobility. However, these mechanisms cannot completely solve the problem of broken trees when, in addition to node mobility, channel errors

cause packets to be dropped. For example, if a data packet is randomly dropped due to channel errors, this may not trigger any of the existing repair mechanisms unless this behavior continues for some time. This is because MC-TRACE uses timers that vary between two to five superframe times to react to a disruption in packet flow. Moreover, the repair mechanisms are not designed to deal with random link breaks that last shorter than the pre-defined timers. The problem of random link failures cannot be handled by the original MC-TRACE protocol. Thus, we introduce adaptive redundancy, which, in conjunction with the branch formation and repair mechanisms of MC-TRACE, helps in dealing with link breaks due to channel errors as well as mobility.

Adaptive systems need to react to changing conditions at a rate fast enough to sustain a desired level of performance. In our case we aim to achieve a better QoS under increasing channel errors while keeping the unnecessary redundancy as low as possible. In order to keep our approach as a distributed and lightweight addition to a multicasting protocol, we locally vary the redundancy in the multicast tree according to the data packet reception history. In order to do this, each node monitors data packet receptions and the corresponding upstream nodes to make sure that data packets are arriving regularly from the same upstream node. If there is a disruption of data packet flow, our mechanism will increase the number of upstream nodes until a maximum number of allowed upstream nodes is reached.

We can describe the operation of adaptive redundancy with the help of the following example. Under perfect channel conditions and in the presence of a well maintained broadcast tree, multicast member node A periodically receives data packets broadcasted by the source node through the branch formed between itself and the source node. However, when an expected data packet is not received (*i.e.,* dropped or never routed to node A), node A starts the process to increase its number of upstream nodes. Node A switches to Increase Redundancy status and announces this information via an *INcrease Redundancy* (INR) packet. An INR packet is transmitted by using one of the empty IS slots (see Figure 2.2), which is chosen randomly. Upon receiving an INR packet, all the nodes in the receive range of the transmitting node switch to INR status if their own hop distance to source (HDTS) is less than or equal to the HDTS of the sender (*e.g.,* if node A's HDTS is equal to 4, nodes with an HDTS less than or equal to 4 switch to INR status; however, nodes with an HDTS larger than 4 do not). When a

node switches to INR mode, it starts to relay the data packets if it has data packets for the desired multicast group. Moreover, it propagates the INR request by broadcasting an INR packet to its one-hop neighbors and starts ACKing an additional upstream node (a node with a lower HDTS) in order to sustain the required redundancy level. This procedure is repeated by all the nodes until the source node is reached. After this point, newly established links are maintained by ACK and pruning mechanisms.

We utilize the IS slots because all the nodes listen to the IS slots regardless of their energy saving mode (*e.g.,* clusterheads and ordinary nodes have different energy saving modes and an ordinary node sleeps more than a clusterhead). Upon receiving the first sequential set of data packets from the same upstream node, node A will reduce the number of its upstream nodes by sending out a *DEcrease Redundancy* (DER) packet, again using the IS slot. Therefore, the data packet reception history plays the main role in determining the level of redundancy in the network.

The main functionality of INR and DER packets is to add (remove) another (redundant) gateway. Once the redundant gateway is established, then the flow of data packets, possibly from multiple flows and/or from multiple sources, can reach the nodes with greater reliability. Note that increased redundancy causes more energy consumption, and the improvement in the performance of MC-TRACE slows down after a certain level of redundancy. Therefore, we limit the maximum level of redundancy that can be introduced by INR packets in order to avoid unnecessary energy consumption and traffic.

Figure 5.7 shows a simple flow chart for our adaptive redundancy mechanism. After each superframe nodes check whether or not consecutive data packets have been received from the same upstream node. The level of redundancy is updated periodically after each superframe according to two conditions; (i) a node must receive two packets in a row from the same upstream node to be able to transmit a DER packet, and (ii) failure to receive any consecutive data packets is enough to increase the level of redundancy.

Note that, this is the particular mechanism used for MC-TRACE, but for other multicast tree-based protocols, similar mechanisms could be used, where new branches are formed (while keeping original branches) when packets are not received as expected. We can remove some of the branches when data packets are consistently received.

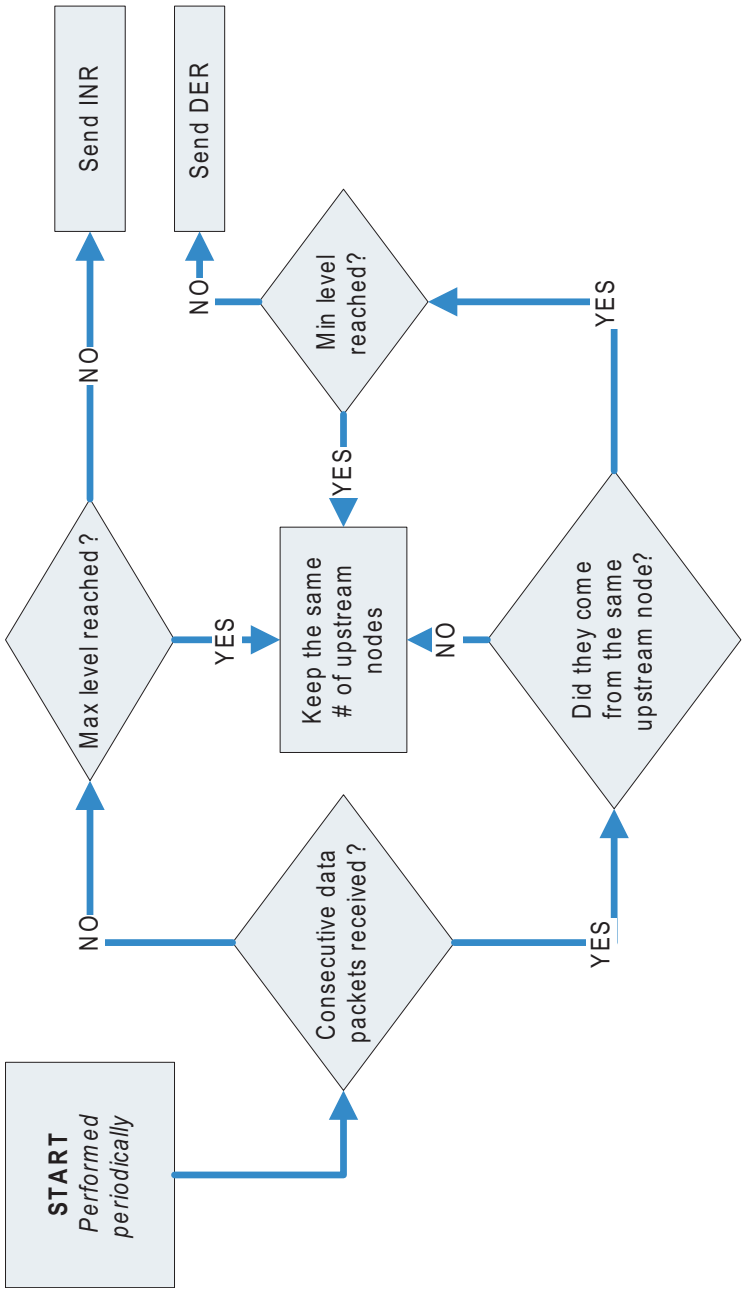
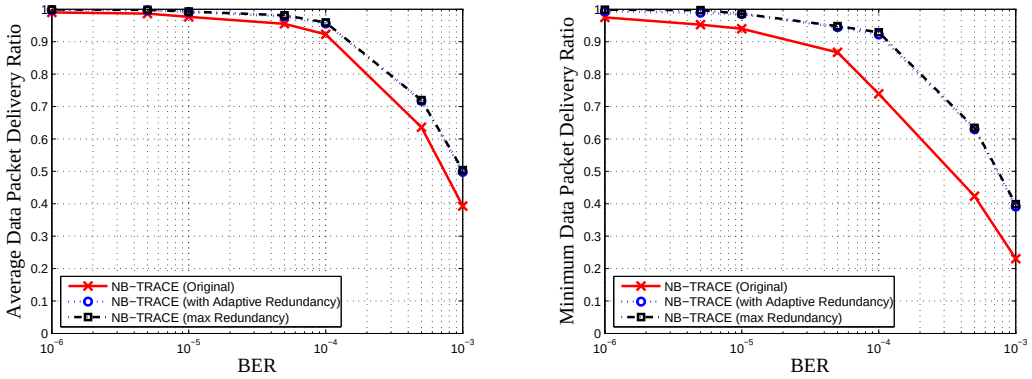


Figure 5.7: Simple flow chart of the adaptive redundancy algorithm.



(a) Average packet delivery ratio versus bit error rate (BER).

(b) Minimum packet delivery ratio versus bit error rate (BER).

Figure 5.8: Average and minimum PDRs versus BER.

5.5 Simulations with Adaptive Redundancy NB-TRACE

Our next step is to perform simulations with the modified version of NB-TRACE and compare it to the original version. We implement the adaptive redundancy approach into the source code of NB-TRACE without modifying the frame structure and initial branch formation mechanisms. Our patch activates itself when a packet drop happens and none of the NB-TRACE repair mechanisms takes action. We investigate the performance increase achieved by the INR mechanism through another set of simulations, much like the ones we performed in Section 5.3.2. We compare NB-TRACE against this new adaptive redundancy version using 120 [1] mobile nodes within a 1 km by 1 km. We include the results obtained through maximum redundancy as a reference, where we use flooding with MH-TRACE. In that case all nodes in the network stay awake at all times and participate in relaying the data packets.

Figure 5.8 shows the average and minimum PDR values obtained with the original NB-TRACE, NB-TRACE with adaptive redundancy, and NB-TRACE with maximum redundancy. The adaptive redundancy enables more packets to be delivered in a more consistent manner to all nodes in the network. This results in an $\sim 1 - 27\%$ increase

[1] The number of nodes is increased to 120 to add more redundant links to the network, but this could not be done in previous simulations because flooding breaks down with this much redundancy in the network.

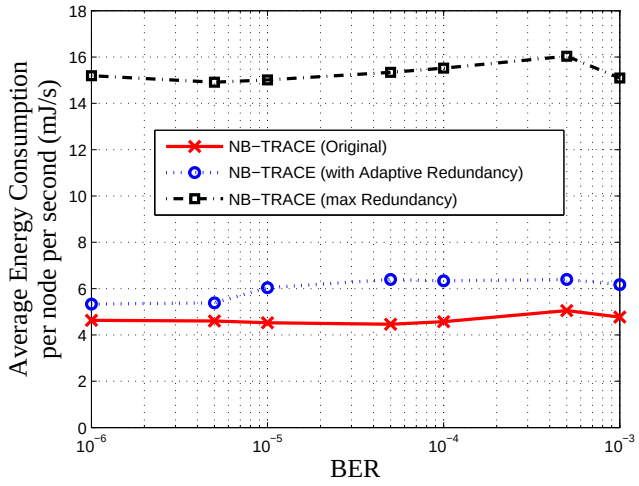
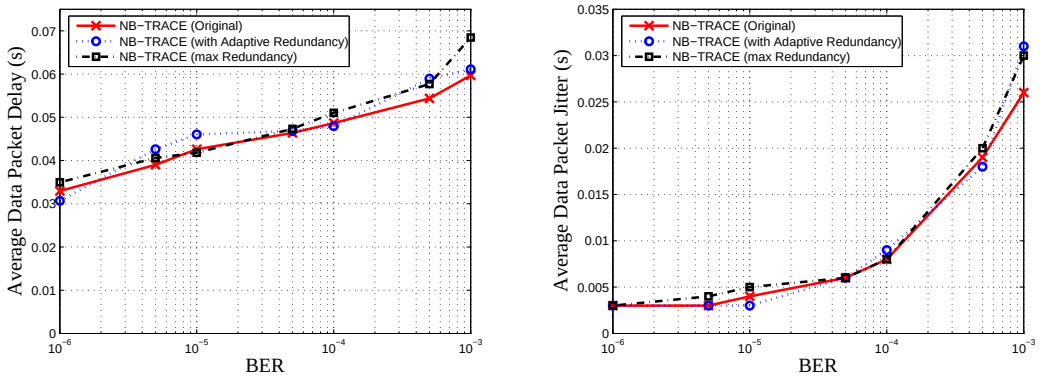


Figure 5.9: Average energy consumption per node per second versus BER.

in average PDR for NB-TRACE. In particular, improvement becomes more visible at higher bit error rates. High minimum PDR values are vital to any protocol that aims to offer QoS. NB-TRACE with adaptive redundancy performs $\sim 2 - 70\%$ better than the original NB-TRACE protocol in terms of minimum PDR achieved at any node in the network. In addition to these results, the adaptive redundancy approach also manages to generate results close to that of maximum redundancy, without the energy expense. This shows the efficiency of our adaptive redundancy approach.

Figure 5.9 shows the energy consumption values for the three versions of NB-TRACE. With these results, our choice of adaptive redundancy becomes clear, mainly because we achieve $\sim 60 - 65\%$ reduction in consumption when compared with the maximum redundancy approach. We would like to point out that the increased energy consumption in NB-TRACE with adaptive redundancy compared with the original NB-TRACE protocol can be seen as a trade-off between PDR and energy consumption.

The data packet delay and jitter values are too close to each other to draw a simple conclusion (see Figure 5.10). However, we would like to point out that relaying data packets under channel errors causes delay and jitter values to increase. This is because data packets are routed through longer routes instead of the shortest path between the source and destination. Once again we observe two different trends in the standard deviation values. At low BER levels, standard deviation values stay within $\sim 5\%$ of the



(a) Average data packet delay versus bit error rate (BER).

(b) Average data packet jitter versus bit error rate (BER).

Figure 5.10: Average delay and jitter of data packets versus BER.

mean values. However, as the BER increases, the standard deviation values reach up to $\sim 18\%$ of the mean values.

We have evaluated the performance gains achieved by incorporating our adaptive redundancy approach into NB-TRACE. Simulation results showed that this new approach introduced a favorable PDR versus energy consumption trade-off. As our next step, we implement the adaptive redundancy approach in a multicasting scenario where a smaller, more focused, and potentially more vulnerable multicast tree is formed between the source and multicast member nodes. Considering the improvement we achieved in network-wide broadcasting, we believe that there is more room for improvement in such a setup.

5.6 Simulations with Adaptive Redundancy MC-TRACE

We perform simulations with the modified version of MC-TRACE and compare it with the original version and ODMRP. We implement the adaptive redundancy approach into the source code of MC-TRACE while keeping the original frame structure intact. However, adaptive redundancy imposes tighter limits on branch formation and repair mechanisms. Moreover, we introduce an advanced acknowledging scheme that enables each node to vary its number of upstream nodes according to the channel conditions and

Table 5.2: Simulation Parameters

Acronym	Description	Value
T_{SF}	Superframe duration	32 ms
N/A	Data packet payload	128 B
T_{drop}	Packet drop threshold	160 ms
D_{Tr}	Transmission range	250 m
D_{CS}	Carrier Sense range	507 m
N/A	Network Area	$1km \times 1km$
P_T	transmit power	600 mW
P_R	receive power	300 mW
P_I	idle power	100 mW
P_S	sleep power	10 mW
C	Channel Rate	2 Mbps
N/A	Max. number of Upstream Nodes	4

data packet reception history. Our patch activates itself when a packet drop happens and before one of the MC-TRACE repair mechanisms can take action. Choosing a shorter reaction time is essential for the performance increase since MC-TRACE does not necessarily guarantee any response for lost data packets due to channel errors.

We investigate the performance increase achieved by the adaptive redundancy approach through a set of ns-2 simulations that compare MC-TRACE and ODMRP against this new adaptive redundancy version of MC-TRACE. We perform various simulations where we employ different channel error characteristics to better demonstrate the performance characteristics of these different approaches. First, we introduce a global BER at the receiving end and track the change in the performance metrics as we gradually increase the BER between the simulations. Next, we employ a different approach and assign a random BER for each node in the network instead of a global BER. The third and the last set of simulations are performed using a well known bursty-noise channel model [26]. In this model, the channel is represented by two states; *good* and *bad*. The transition probabilities between the states determine the burstiness of channel errors. Note that we have simulated our approach with a few other channel error

models where we vary the global BER throughout the simulation and use the bursty-noise model with different setups. However, we have elected to show only the results that revealed further insight into the performance of adaptive redundancy rather than repeating the results of other simulations.

5.6.1 Constant BER

We start with simulations where we vary the global BER between 10^{-5} and 10^{-3}. We choose 8 multicast members among 128 nodes and increase the number of multicast members to 32 while keeping the number of nodes at 128 in order to observe the effect of the traffic load on these three different multicasting approaches. Later, we increase the number of nodes to 256 while keeping the number of multicast members at 8. This reveals the effects of increasing node density on the protocols. In all our simulations we use mobile nodes moving within a 1 km by 1 km simulation area for 100 seconds according to the Random Way-Point (RWP) mobility model with node speeds (S) chosen from a uniform distribution where $0.0m/s < S \leq 5.0m/s$. The average instantaneous node speeds as a function of time vary from $2.3m/s$ to $2.5m/s$ throughout the simulation time. A short summary of the common parameters that we employed in all our simulations can be found in Table 5.2.

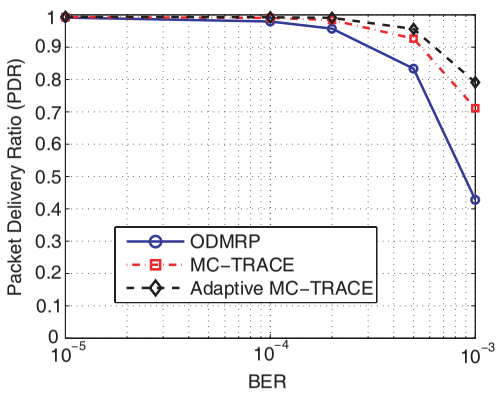
5.6.1.1 Packet Delivery Ratio

Figure 5.11 shows the average and minimum PDR values of simulations with a group of 8 and 32 multicast members out of 128 nodes, and 8 multicast members out of 256 nodes. Results obtained with ODMRP, original MC-TRACE, and MC-TRACE with adaptive redundancy are displayed together to offer a relative comparison between these protocols. For 8 multicast members (with 128 nodes), adaptive redundancy improves both the average PDR and minimum PDR of MC-TRACE at most by $\sim 11\%$ and $\sim 41\%$, respectively.

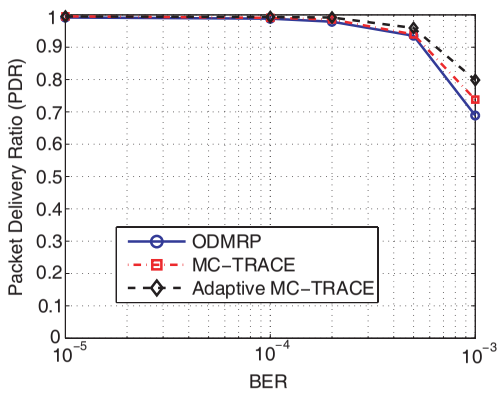
As the number of multicast group members is increased from 8 to 32 while keeping the number of nodes in the network at 128, ODMRP exhibits a closer performance to that of MC-TRACE due to the fact that a bigger more reliable mesh cloud, which covers almost all of the network area, is formed for 32 multicast member nodes. In this

scenario, the improvement in the average PDR and minimum PDR values for adaptive MC-TRACE reaches up to $\sim 8\%$ and $\sim 52\%$, respectively.

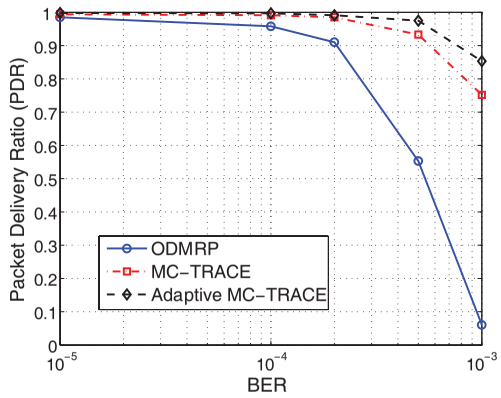
When we double the number of nodes in the network (from 128 to 256), performance characteristics of these three different approaches become more clear. The first thing we notice is the rapid fall in the PDR values of ODMRP. Intuitively, ODMRP should perform better as the node density increases (*i.e.,* as the number of nodes in the routing mesh cloud increases). However, due to increased bandwidth utilization and the resulting congestion in the network ODMRP fails to cope with contention resolution and exhibits an inferior PDR performance to that of MC-TRACE. MC-TRACE PDR performance, which shows a slight decrease at $BER = 10^{-5}$ due to the increasing interference within the network when compared with the 128 nodes scenarios, benefits from increasing node density because of the fact that under increasing channel errors there are more nodes to respond to the branch repair mechanisms of MC-TRACE. The adaptive mesh approach still manages to improve the average PDR and minimum PDR values of MC-TRACE at most by $\sim 13\%$ and $\sim 37\%$. For MC-TRACE and adaptive MC-TRACE, standard deviation values for PDR and minimum PDR stay within $\sim 5\%$ of the mean values. However, for ODMRP, the standard deviation values reach up to $\sim 25\%$ of the mean values.



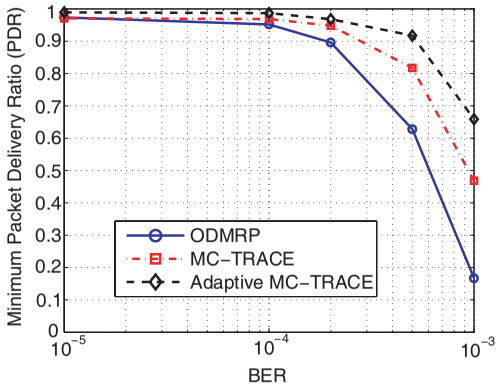
(a) (8 multicast members out of 128 nodes) Average packet delivery ratio versus BER.



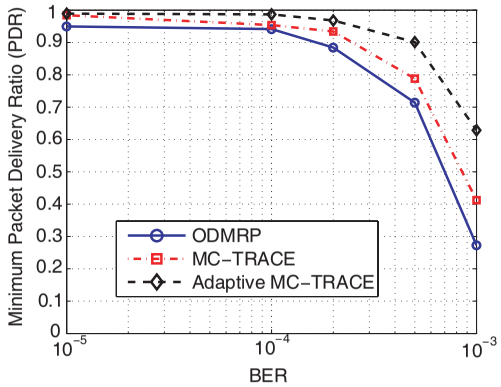
(b) (32 multicast members out of 128 nodes) Average packet delivery ratio versus BER.



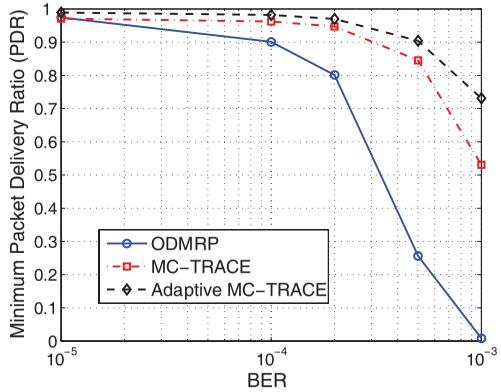
(c) (8 multicast members out of 256 nodes) Average packet delivery ratio versus BER.



(d) (8 multicast members out of 128 nodes) Min packet delivery ratio versus BER.



(e) (32 multicast members out of 128 nodes) Min packet delivery ratio versus BER.



(f) (8 multicast members out of 256 nodes) Min packet delivery ratio versus BER.

Figure 5.11: Average and minimum PDRs versus BER (8 and 32 multicast members with 128 nodes, and 8 multicast members with 256 nodes).

These results can be further explained by referring to the basics of these two protocols. ODMRP is an on-demand routing protocol that creates a mesh cloud between the members and the source node. In this cloud, each node participates in routing in order to increase the redundancy in the packet delivery process. In our simulations, we employed the IEEE 802.11 MAC layer with ODMRP and IEEE 802.11 has an inherent inability to cope with contention resolution under extremely high load [62]. On the other hand, MC-TRACE is a tree based multicasting protocol that incorporates branches between the members and the source of a multicast. Redundancy in the multicasting process is kept at minimum by pruning the unnecessary branches. Under increasing channel errors, these different characteristics of ODMRP and MC-TRACE might have favored the highly redundant mesh cloud of ODMRP to have a better performance when compared with the optimized tree structure of MC-TRACE. However, MC-TRACE's pro-active nature in maintaining its tree structure helps MC-TRACE in outperforming the redundant mesh structure of ODMRP.

We believe that the mesh maintenance parameters of ODMRP can be adjusted to improve the performance under heavy channel errors. In fact a recent improvement on the ODMRP protocol showed that the route refresh interval of ODMRP has critical impact on protocol overhead and thus efficiency. For fast changing channel conditions, when ODMRP cannot keep up with network dynamics, a refresh rate update mechanism that dynamically adapts the refresh interval to the environment is proposed in the enhanced ODMRP (E-ODMRP) protocol [89]. In addition to this, E-ODMRP introduces a local recovery mechanism that detects a broken route and repairs the forwarding mesh, much like the repair branch mechanism in MC-TRACE. However, the energy-efficiency of E-ODMRP is still not comparable to that of MC-TRACE. As we will see in the next subsection, ODMRP suffers from high energy consumption due to its redundant mesh routing approach. Although E-ODMRP dramatically improves the overhead performance for low traffic and less dynamic scenarios, in highly dynamic environments both protocols (ODMRP and E-ODMRP) exhibit identical performances.

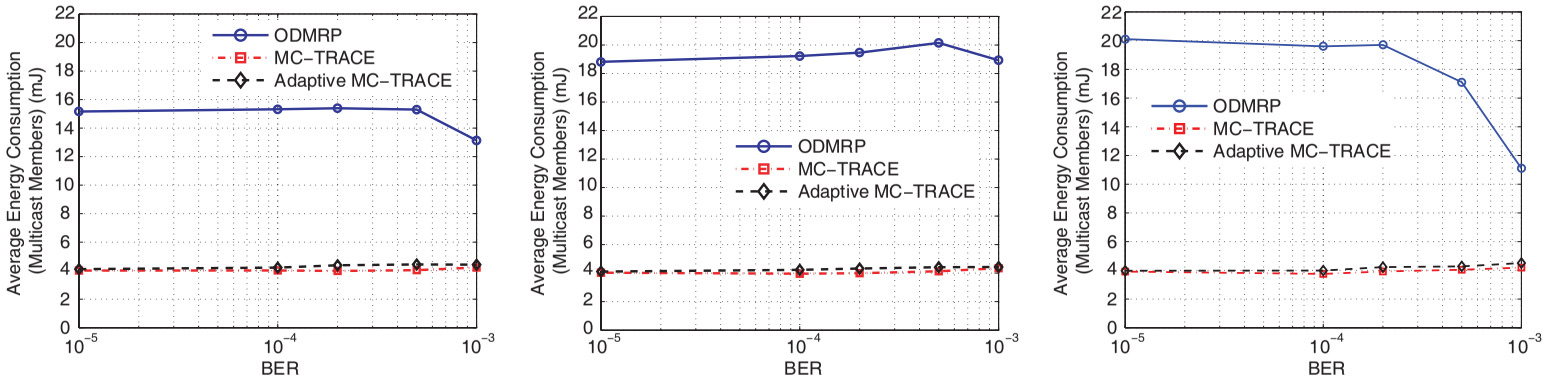
In conclusion, adaptive redundancy enables more packets to be delivered in a more consistent manner to all nodes in the network. In particular, the improvement becomes more visible at higher bit error rates. High minimum PDR values are vital to any protocol that aims to offer QoS.

5.6.1.2 Energy

In our simulations, we model the energy dissipation of a given node as follows: (i) transmit mode, (ii) receive mode, (iii) idle mode, (iv) carrier sense mode, and (v) sleep mode. Transmit and receive energies are dissipated for packet transmissions and receptions, respectively. Carrier sense energy dissipation is another form of receive energy dissipation. The difference is the fact that the source node is located in the carrier sense region rather than the transmit region. In the case where none of the nodes within the communication or carrier sense range are transmitting packets, a node is in the idle mode if not in the sleep mode, which is a low energy state.

Figure 5.12 shows the energy consumption values for the three different protocols. There are three different setups (8 multicast members out of 128 nodes, 32 multicast members out of 128 nodes, and 8 multicast members out of 256 nodes) we employ throughout the constant BER simulations. Average energy consumption values for both the multicast members and also all nodes in the network are provided to show that the overall energy consumption increases as the channel errors force more redundancy to be introduced in the network. However, average energy consumption among the multicast members of Adaptive MC-TRACE stays close to that of original MC-TRACE as BER increases.

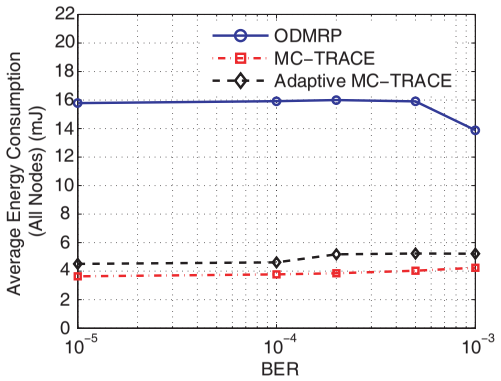
With these results, our choice to improve tree based routing through adaptive redundancy becomes clear, mainly because we consume $\sim 65 - 70\%$ less energy than what ODMRP does. We would like to point out that the increased energy consumption due to the adaptive redundancy approach compared with original MC-TRACE can be seen as a trade-off between PDR and energy consumption. ODMRP's energy consumption reduces for the last data point in both figures since fewer data packets are relayed due to the high BER. For 256 nodes (Figures 5.12(c) and 5.12(f)), ODMRP shows a sharper decline in the energy consumption due to the trend in the PDR results. The standard deviation values for these results stay around $\sim 3\%$ of the mean values.



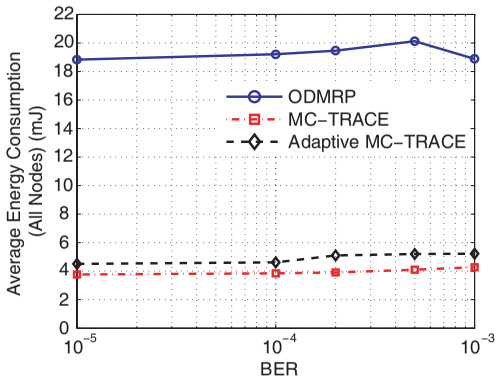
(a) Multicast members only (8 out of 128).

(b) Multicast members only (32 out of 128).

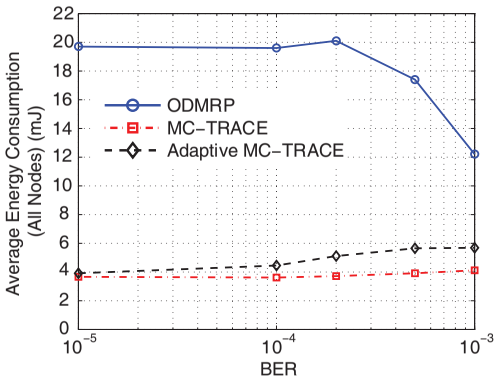
(c) Multicast members only (8 out of 256).



(d) All nodes (128)



(e) All nodes (128)



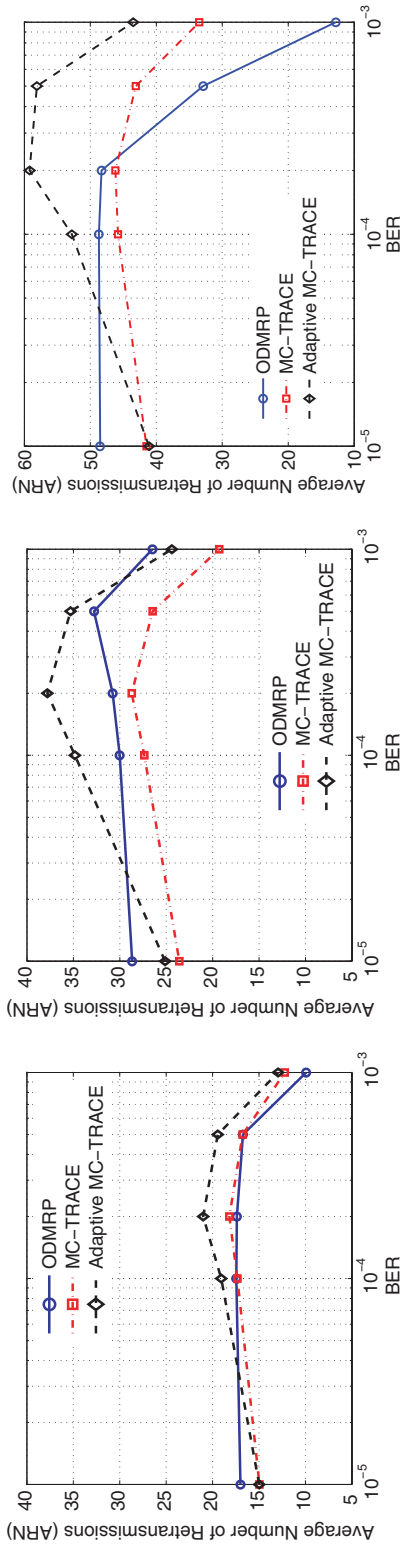
(f) All nodes (256)

Figure 5.12: Average energy consumption per node versus BER (8 and 32 multicast members with 128 nodes, and 8 multicast members with 256 nodes).

5.6.1.3 Average Number of Retransmissions (ARN)

We measure the bandwidth efficiency as the number of required data forwards to cover all the multicast nodes. We call this metric Average Number of Retransmissions (ARN). ARN is the ratio of the total number of data transmissions to the total number of data packets sent down from the application layer. Thus, the higher the ARN is the lower the bandwidth efficiency. Figure 5.13 shows the ARN values for the three different protocols with 8 and 32 multicast group members.

Differences between Figures 5.13(a) and 5.13(b) reveal the behavior of our adaptive redundancy approach. As the number of multicast members is increased from 8 to 32, more nodes can be involved in the redundant mesh formed by both ODMRP and Adaptive MC-TRACE. However, Adaptive MC-TRACE manages to limit its ARN for low BER values where ODMRP results in higher ARN. With increasing BER Adaptive MC-TRACE forces more nodes to participate in the relaying process resulting in increased ARN. As we approach to high BER levels, increased packet drop rate cripples the data traffic causing ARN to drop. When we switch to a higher node density (see Figure 5.13(c)), especially at high BER levels, ODMRP completely collapses and fails to reroute the data packets. Adaptive redundancy, on the other hand, can maintain higher number of retransmissions at around $BER = 5 \times 10^{-4}$, which means better QoS performance than other protocols. The standard deviation values for these results stay around $\sim 5\%$ of the mean values.



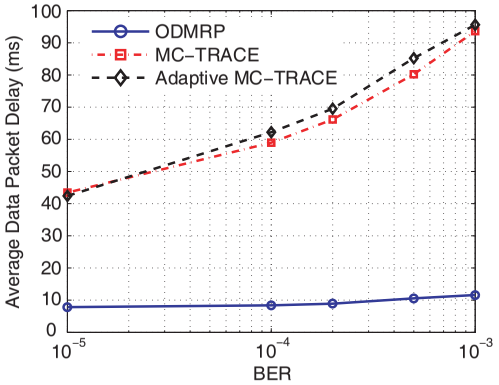
(a) (8 multicast members) ARN versus BER. (b) (32 multicast members) ARN versus BER. (c) (8 multicast members out of 256) ARN versus BER.

Figure 5.13: ARN versus BER (8 and 32 multicast members with 128 nodes, and 8 multicast members with 256 nodes).

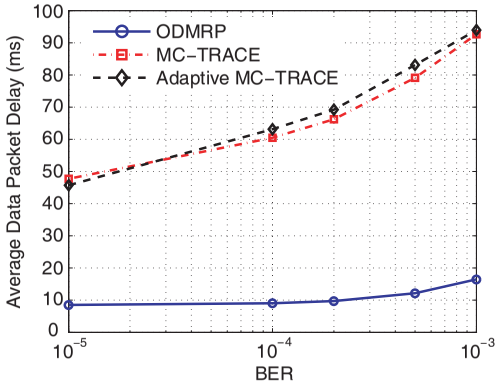
5.6.1.4 Delay and Jitter

Average packet delay is obtained by averaging the delays of all data packets that are received for the first time at multicast nodes. Root Mean Square (RMS) delay jitter, which is a measure of the deviation of the packet inter arrival time from the periodicity of the packet generation period, is the standard deviation of the data packet inter-arrival times. The data packet delay and jitter values presented in Figure 5.14 are too close to each other to draw a simple conclusion. The fact that MC-TRACE's cyclic superframe structure makes it virtually impossible to achieve data packet delay values lower than the superframe time ($32ms$) causes higher delay values for MC-TRACE compared with ODMRP. Jitter values, however, remain comparable for all the protocols for low to mid BER levels. ODMRP performs worse as we reach high BER levels.

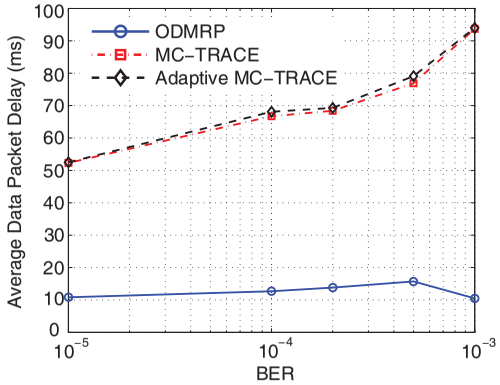
We observe similarity in ODMRP's data packet jitter performance (in Figures 5.14(d) and 5.14(f)). This can be explained by the fact that fewer number of multicast members (8 instead of 32) results in a smaller mesh cloud which is more susceptible to channel errors. Therefore, data packet arrival times fail to maintain their periodical behavior, which leads to a poor QoS. The main reason for MC-TRACE's low jitter is the periodic renewal of channel access (*i.e.,* once a node reserves a data slot it can have periodic channel access without having to contend for the channel). Thus, the nodes forward the sequential packets of the same flow without any deviation from the periodicity. The standard deviation values for most of these results stay around $\sim 6\%$ of the mean values. Only ODMRP with 256 nodes produces high standard deviation values that reach up to $\sim 50\%$ of the mean values.



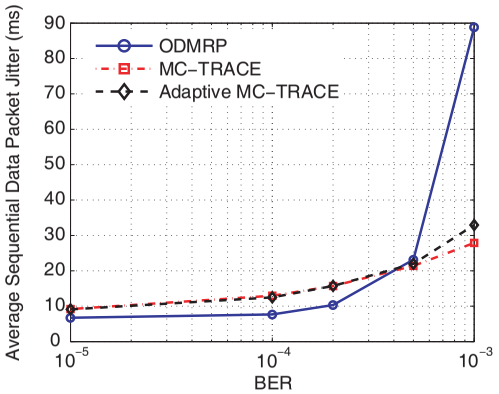
(a) (8 multicast members out of 128 nodes) Average data packet delay versus BER.



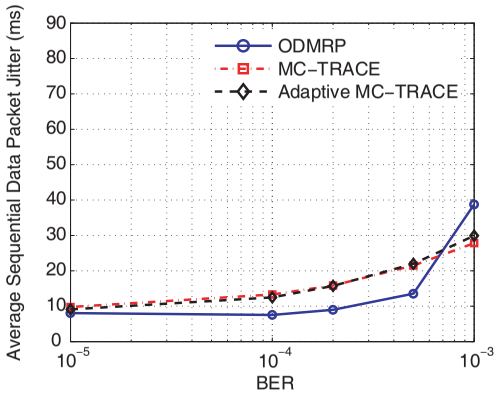
(b) (32 multicast members out of 128 nodes) Average data packet delay versus BER.



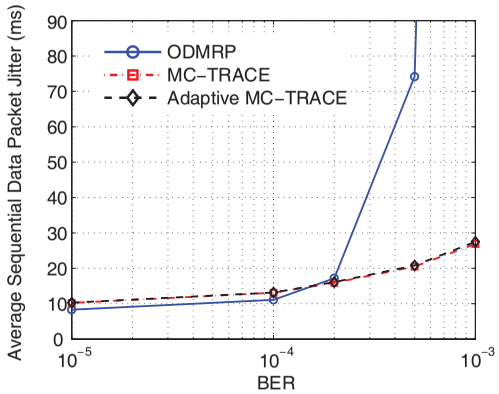
(c) (8 multicast members out of 256 nodes) Average data packet delay versus BER.



(d) (8 multicast members out of 128 nodes) Average data packet jitter versus BER.



(e) (32 multicast members out of 128 nodes) Average data packet jitter versus BER.



(f) (8 multicast members out of 256 nodes) Average data packet jitter versus BER.

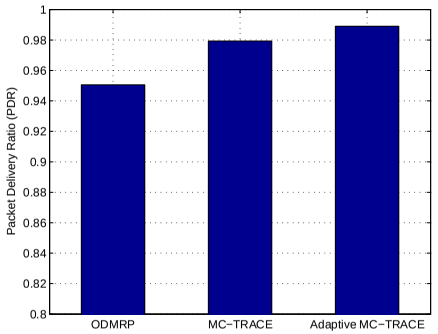
Figure 5.14: Average delay and jitter of data packets versus BER (8 and 32 multicast members with 128 nodes, and 8 multicast members with 256 nodes).

5.6.2 Random BER

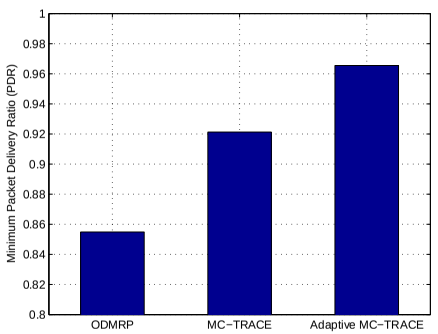
In the previous subsection we focused on the performance of the protocols under a constant global BER. In this subsection, we assign each node a constant BER, which is chosen randomly from the interval of $[5 \times 10^{-6}, 5 \times 10^{-4}]$. This corresponds to a data packet drop rate interval of $[0.5\%, 40\%]$. In other words, random BER values are assigned to the nodes in the network to achieve spatial variation in the BER values throughout the network. We have chosen this setup to reveal the benefits of using local adaptation over mesh-based routing.

In Figure 5.15, results of the simulations are collectively presented. PDR and ARN values are presented in an aggregated fashion to emphasize the performance improvement between the ODMRP, MC-TRACE and Adaptive MC-TRACE protocols. The PDR performance of adaptive MC-TRACE is better than that of ODMRP and MC-TRACE due to the localized adaptive redundancy management approach. The minimum PDR achieved at any multicast member is increased by the adaptive redundancy (from $\sim 92\%$ to $\sim 96\%$ at the expense of ~ 4 extra average number of retransmissions (ARNs)). However, the effect of the increased ARN value does not cause a considerable increase in the energy consumption of Adaptive MC-TRACE when compared with regular MC-TRACE. The adaptive redundancy approach manages to overcome localized bottlenecks (*i.e.,* links with higher BER values) to further improve the QoS provided by MC-TRACE.

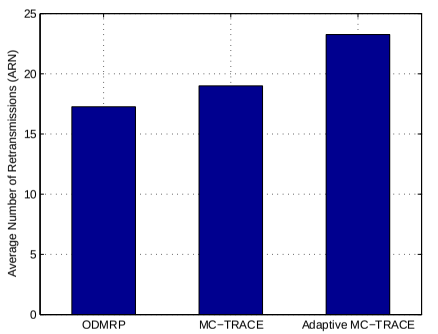
The data packet delay values follow the same trend we observed in Section 5.6.1.4. ODMRP outperforms both MC-TRACE versions due to its IEEE 802.11 based MAC scheme. The jitter values are all very close to each other hence are not displayed here. The standard deviation values stay below $\sim 6\%$ of the mean values.



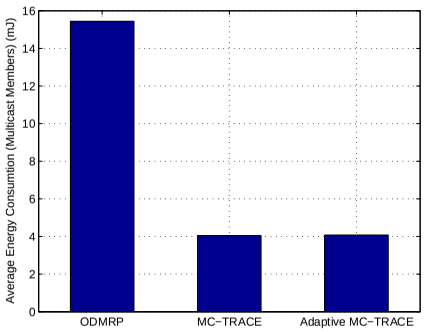
(a) Average PDR.



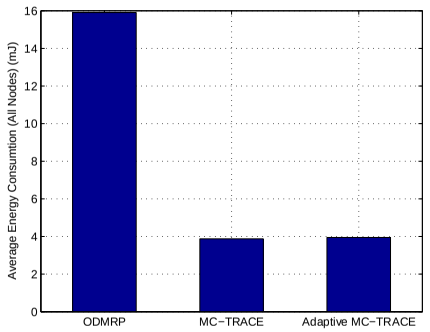
(b) Minimum PDR.



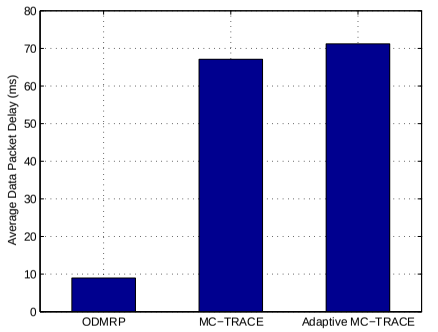
(c) ARN.



(d) Average energy consumption per node (Multicast members only).



(e) Average energy consumption per node (All nodes).



(f) Average data packet delay.

Figure 5.15: Results for ODMRP, MC-TRACE, and Adaptive MC-TRACE (8 multicast members with 128 nodes and BER varies between $[5 \times 10^{-6}, 5 \times 10^{-4}]$)

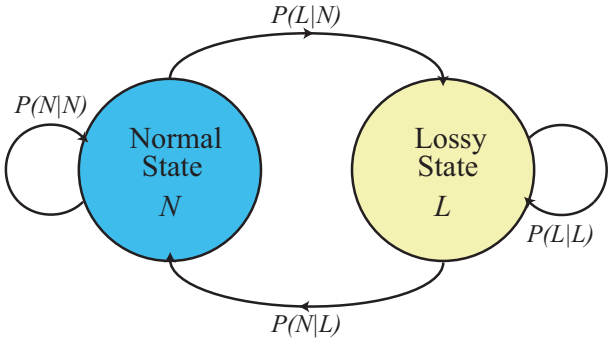


Figure 5.16: Gilbert-Elliott channel model.

5.6.3 Two-state Markov Channel

Using a Markov chain representation assumes that an adequate description of a system is given by a finite number of states. Each state is assigned a probability of the system being in that state. The study of Markov approximation for fading channels dates back to the early work of Gilbert [24] and Elliott [25], who built a two-state Markov channel known as the Gilbert-Elliott channel. In a simplified Gilbert model [26], the error probabilities in "bad" and "good" states are 1.0 and 0.0, respectively. The state transition diagram is shown in Figure 5.16. The Normal state represents a perfect channel in which there is no error present, whereas the Lossy state represents a wireless channel in which no packet can be delivered without error. The channel statistics are controlled by a set of transition probabilities that determine the individual probabilities P(N) and P(L) [71].

For our simulations, the Gilbert model is employed with transition probabilities and state probability distribution shown in Table 5.3. We employ the state transition probabilities proposed according to the empirical data, which is obtained by observing the data communications in a Global System for Mobile communications (GSM) network [30].

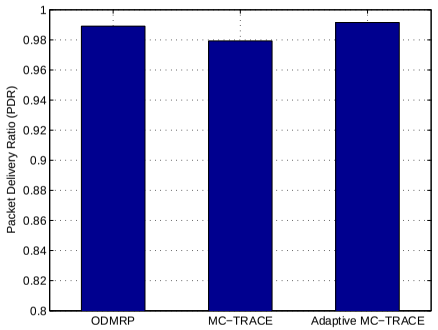
Simulation results (presented in Figure 5.17) indicate a new performance behavior that does not exactly follow the results we presented in the previous sections. The PDR values of ODMRP turn out to be larger than that of MC-TRACE for the first time in our simulation analysis. Considering the ARN results, we conclude that under bursty channel errors, MC-TRACE fails to preserve its pro-active behavior. This can

Table 5.3: Gilbert-Elliott Channel Model Statistics.

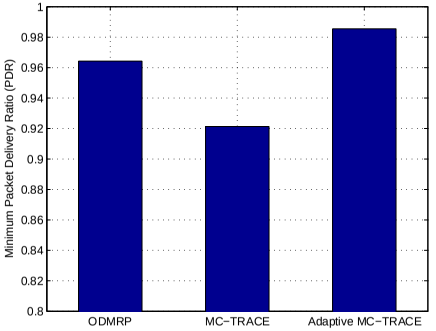
State	**P(State)**	**P(*L*|State)**	**P(*N*|State)**
N (Normal)	0.9375	0.01	0.99
L (Lossy)	0.0625	0.85	0.15

be explained by the fact that for the entire error burst duration (*i.e.,* when the channel is in the lossy state) all the packets are dropped regardless of their size (*i.e.,* BER is equal to one). In the lossy state, the control traffic of MC-TRACE (consisting of small sized packets) is affected more than it was when there was a constant BER and the drop rate of control traffic was close to zero. Consequently, MC-TRACE without the adaptive redundancy (*i.e.,* tree-based routing) fails to maintain its superior performance over ODMRP.

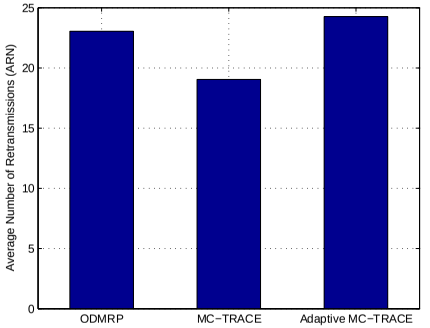
The quick response time of the adaptive redundancy approach helps adaptive MC-TRACE to outperform ODMRP. However, when compared with results obtained with previous channel error models, the performance gains achieved by adaptive redundancy is relatively smaller. The energy consumption and data packet delay values maintain the same behavior that is established throughout this chapter. Once again, the data packet jitter vales are not included here because they are too close to each other ($\sim 15ms$). Note that, the standard deviation values stay below $\sim 8\%$ of the mean values.



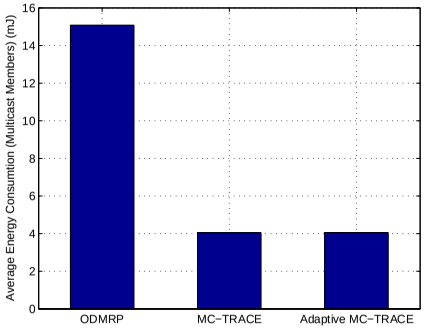
(a) Average PDR.



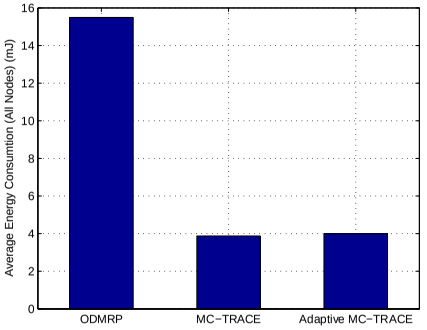
(b) Minimum PDR.



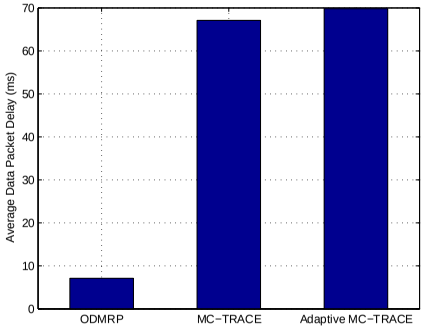
(c) ARN.



(d) Average energy consumption per node (Multicast members only).



(e) Average energy consumption per node (All nodes).



(f) Average data packet delay.

Figure 5.17: Results for ODMRP, MC-TRACE, and Adaptive MC-TRACE (8 multicast members with 128 nodes with bursty channels).

5.7 Summary

In this chapter, we proposed an adaptive redundancy algorithm in order to improve the performance of ad-hoc routing protocols under channel errors. We investigated the impact of channel errors on the performance of a tree-based network-wide broadcast protocol (NB-TRACE) and a mesh-based network-wide broadcast protocol (flooding with IEEE 802.11) [90]. We showed that for mid to high BERs NB-TRACE QoS is effected more than that of simple flooding. On the other hand, simple flooding could not handle the traffic generated at low BER levels as well as NB-TRACE did. Later, we analyzed the limits of redundancy in routing and pointed out the trade-off between the energy consumption and level of redundancy in the network.

Based on the results with network-wide broadcasting, we moved on to exploring adaptive redundancy in multicasting [91], proposing an adaptive redundancy scheme that locally increases/decreases the number of upstream nodes according to the packet reception history and channel errors. We implemented the adaptive redundancy algorithm on MC-TRACE, a tree-based multicasting approach, and compared it to a mesh-based multicast protocol (ODMRP). Through ns-2 simulations, we showed that adaptive redundancy maintains high packet delivery ratios at higher protocol efficiency when compared to both non-adaptive mesh and tree-based multicast protocols under several channel error models [92].

This work showed the advantages provided by adapting the routing redundancy as channel conditions change. While route-layer reliability helps ensure continuous data flow in the presence of node mobility and changing link states, channel noise is often the cause of a "broken" link, rather than simply being caused by node mobility. In such a case, we might do better in terms of overhead by utilizing link-level reliability mechanisms, such as the use forward error correction (FEC) to automatically correct errors in data transmission at the receiver. Thus, in the next chapter, as a continuation of our goal of improving the reliability and the QoS performance of real-time communications in MANETs, we summarize and propose various adaptive ways to overcome the effects of channel errors in mobile ad-hoc networks. We focus on real-time data communications, with an emphasis on multicasting in MANETs.

Chapter 6

Adaptive Techniques for Reliability

In a mobile ad-hoc network (MANET), successful real-time data communications, such as multimedia broadcasting, multicasting and unicasting, require (*i*) minimizing energy dissipation, (*ii*) providing QoS for real-time data (*e.g.,* voice, video) packets, and (*iii*) enabling bandwidth-efficient multi-hop broadcasting and multicasting. Because of the potentially limited shared bandwidth of the network, and the lack of a central controller that can account for and control these limited resources, nodes must negotiate with each other to avoid collisions and manage the resources required for maintaining QoS along the routes. This is further complicated by frequent topology changes and the time/space varying nature of the wireless channel.

Applications that require QoS such as multimedia broadcasting and multicasting are becoming increasingly important applications for MANETs, and protocols are needed to improve the reliability and performance provided by the network. However, as MANETs continue to increase in size and complexity, further QoS functionality and adaptivity towards challenging wireless conditions may be needed, along with ways to provide QoS in an energy-efficient manner.

Many real-time applications require a guaranteed end-to-end data rate and delay, which are two of the most important QoS metrics in ad-hoc networks. However, it is often unrealistic to expect that all of the QoS metrics can be satisfied simultaneously. In fact, most of the time there will be a trade-off between the QoS metrics, and the application must decide at which point on this trade-off curve to operate. This could be a rate-delay trade-off or simply a choice between increased energy dissipation and

increased latency. The lower layer protocols determine how many of these trade-off choices the application will have. Ideally, the trade-off curves in network design should be multidimensional, incorporating as many QoS parameters as possible (*e.g.,* rate, delay, robustness, and energy efficiency trade-offs).

Researchers have developed an extensive number of techniques for combating the effects of poor channel conditions and node mobility in MANETs. These techniques include: (1) physical layer techniques, such as transmission power control, modulation and forward error correction, (2) MAC layer techniques, such as collision avoidance and resource allocation and (3) network layer techniques, such as cost-aware routing and multi-path redundant routing (*e.g.,* mesh-based routing) [93, 94]. Many of these techniques can benefit from cross-layer information sharing, such that information from one layer (*e.g.,* link conditions detected at the Physical layer) is used to inform the decisions of a different layer (*e.g.,* the cost function to use in the routing protocol at the Network layer).

While these techniques have been shown to greatly improve the performance of MANETs, they all add overhead to the system that may not be needed under certain operating conditions. Therefore, it is advantageous to adapt the use of these techniques to the current network conditions rather than design a-priori (oftentimes conservative) estimates of the required protection. This will enable the network to reduce the overhead during periods of good performance yet quickly adjust as the network conditions degrade in order to maintain acceptable performance.

All of the techniques described above have parameters that can be adapted, either temporally as link conditions change or spatially to account for spatial variations in network conditions. For example, transmission power control can dynamically set the transmit power based on estimates of the link quality over time, and links in different parts of the network can utilize different settings of transmit power based on spatial variations in link quality. Similarly, a multi-path routing protocol can adjust the number of redundant routes (or the size of the mesh) based on local and/or temporal network conditions. By determining the potential adaptation options for the various techniques, system designers can optimize the performance vs. overhead (*i.e.,* lifetime) trade-off for supporting real-time data transfer in MANETs.

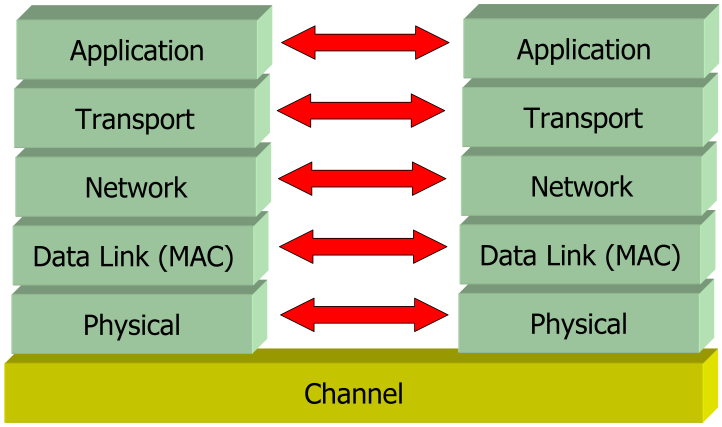


Figure 6.1: TCP/IP reference model.

In this chapter, we survey several existing techniques for combating the effects of lossy channels in MANETs. We begin by discussing physical layer techniques, including transmit power control (TPC), modulation and coding, and we describe the parameters that can be adapted using each of techniques. Next we discuss the MAC layer techniques of collision avoidance and resource allocation and the issues involved in adapting these approaches for mitigating the effects of channel errors. Finally, we detail routing layer techniques such as cost-aware routing and multi-path routing, including mesh-based multicasting, as well as the adaptation parameters for these techniques. Based on this discussion, we provide a guide towards an adaptive optimization manager that can manage these decisions in an integrated and coherent manner.

6.1 Adaptive Techniques

Most of the popular network protocols are created using a modular design methodology, where the modules of a synthesized protocol are arranged in a vertical stack. There are several reference models for describing the layers of a communication network, such as the Open System Interconnection (OSI) reference model [56] and the Transmission Control Protocol / Internet Protocol (TCP/IP) reference model [80, 95]. The objective for organizing the network functions into layers is simple and clear: management of a single complex module is not as easy as a general design rule in the broad field of tech-

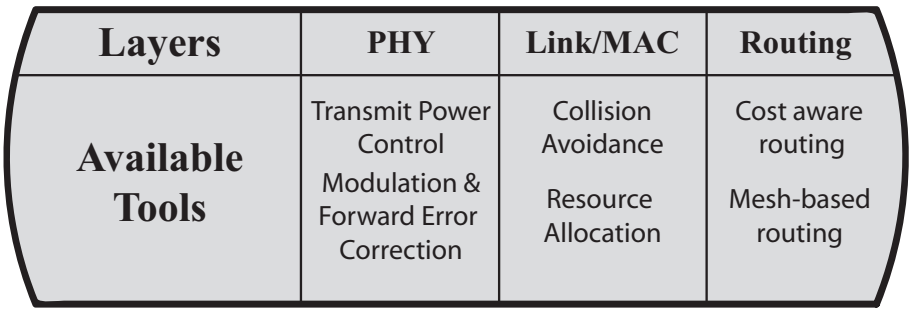


Figure 6.2: Available tools to improve communication reliability at different layers.

nology. Instead, a system created from well-integrated but separable blocks is easier to design, manage and maintain. Just to emphasize the functionality of various layers of a generic communication protocol, we will refer to the layered protocol stack described in [80], which is basically the TCP/IP reference model and is shown in Figure 6.1. Note that we refer to the *Network* layer as the *routing* layer in this thesis.

Figure 6.2 lists some of the available techniques for mitigating the effects of lossy channels. Some of these techniques span multiple layers of the communication stack. For example, transmit power control is optimized at two layers, the physical and data link (*i.e.,* media access) layers. Moreover, resource allocation needs collaboration between the MAC and routing layers. Cross-layer design considers multiple layers of the protocol stack together. This can be either in terms of a joint design or in information exchange between the layers. We believe that cross-layer designs can tailor the protocol stack according to different QoS requirements and achieve the design goals with much higher efficiency when compared to a general architecture [82].

In the following sections, we summarize recent works that utilize one or a combination of these adaptive techniques. We include, where possible, research directions that can further improve these state of the art solutions.

6.2 Physical Layer Adaptation

In this section, we describe the adaptive techniques that originate from the physical layer of the communication stack. Specifically, we look into the recent work on transmit power control and modulation/forward error correction.

6.2.1 Transmit Power Control

Transmission power in wireless ad hoc networks directly determines network connectivity. Moreover, the strategy used for managing the transmission power plays an important role in optimizing the network lifetime. Additionally, transmission power can be used to improve the reliability, latency and efficiency of wireless communications under channel errors [1].

Variations in the transmit power affect the connectivity, capacity, and power conserving properties of wireless ad hoc networks. Using a high transmission power reduces the number of hops needed to reach the intended destination, but this creates excessive interference in a medium that is commonly shared. In contrast, using a lower transmission power reduces the interference seen by other transmitting nodes, but packets require more hops to reach their destination. Connectivity also increases with increasing transmit power since the number of direct links between nodes is increased. However, this results in reduced network capacity [96].

For example, Figure 6.3 shows the resulting branch formations according to two different transmission radii T_{R1} and T_{R2}. Note that higher transmission power ($P_2 > P_1$) results in fewer hops between the source-destination (S-D) pair, however, more nodes are affected. Therefore, the disadvantages of increased transmit power are a larger interference region, reduced shared bandwidth, and increased energy consumption. If these drawbacks are acceptable for a network, one can adapt the transmit power according to the channel noise level to achieve a constant signal to noise ratio (SNR) at each transmission. Consistent SNR leads to simpler modulation and coding.

Additionally, adapting the transmit power according to the distance between the transmitter and receiver can help with CDMA communication since TPC relieves the near-far effect and improves channel capacity [97]. Moreover, under heavy channel loads, individual TPC at each node along with interference feedback from the physical layer can be used to allow capture effects to overcome the increased number of collisions. This way, the packets with larger power in the collision can be decoded due to the capture effect. However, this technique is useful only when nodes are trying to improve their last hop communication (*e.g.,* transmissions to a sink or a base station).

The routing layer can also benefit from varying the transmission power when there is a need for quick route discovery. Especially for real-time data applications, latency is

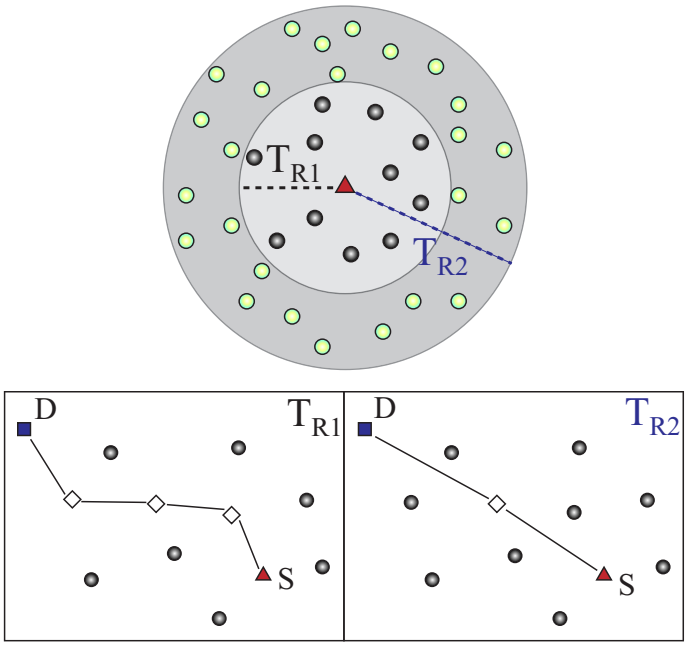


Figure 6.3: Illustration of transmission radii T_{R1} and T_{R2} resulting from transmission power P_1 and P_2, respectively (only propagation loss is considered). Established routes (corresponding to T_{R1} and T_{R2}) between the source (S) and the destination (D) are also shown.

of the utmost importance. In order to reach all nodes in the network and determine the branches or mesh structures to be constructed between the source and the destination(s), larger transmission power reduces the branch/mesh construction time. On the other hand, when latency is not a priority, transmission power can be reduced to satisfy other possible constraints such as energy consumption and network lifetime. On the flip side, this approach requires flooding of the updated transmission power data to all nodes in the network. Therefore, the overhead will increase when the transmission power needs to be adapted according to the ongoing traffic.

Many interesting problems regarding TPC remain to be addressed. As the number of handheld devices with different capabilities continuously increase, interoperability within existing standards and hardware is one of the most important issues. Interoperability requires that newer protocols and schemes have to be backward compatible with the existing transmit power control approaches that are already implemented.

6.2.2 Modulation, Forward Error Correction

For a given channel, achievable rates have been investigated thoroughly and information theoretical limits have been determined [98]. However, it is often impossible to find a practical solution that can get close to the theoretical limit. Modulation and channel coding are two well-known tools for reducing errors in data transmission. In particular, forward error correction (FEC) techniques add controlled redundancy to data in order to allow the receiver to detect and correct errors. This can help avoid retransmissions at the cost of higher bandwidth requirements. Therefore, FEC is applied in situations where retransmissions are relatively costly or impossible. Real-time data traffic is an example of such a situation, where transmitting the next packet in a sequence is always preferred over retransmitting the old one.

Adaptive modulation and coding provides reliable and efficient communications over time-varying channels. Estimation of the channel at the receiver is usually fed back to the transmitter in order to adapt the transmission scheme to the varying channel characteristics. Non-adaptive schemes have to be designed with a considerable margin for error to maintain acceptable performance when the channel quality is poor. To avoid the inefficient utilization of the wireless channel and to increase the average throughput, modulation and coding can be adapted to the current channel conditions. The goal is usually to dynamically adapt system behavior (*e.g.,*coding and modulation parameters) to the channel bit error rate (BER), QoS, available bandwidth, and type of the traffic (*i.e.,* application). Adaptive approaches are also essential for efficient transmission because they promise the best trade-off between transmission overhead and guaranteed QoS [99].

Figure 6.4 shows a transmitter receiver pair that adaptively vary the modulation and coding for the transmission according to the transmitter's estimate of the channel conditions and according to the QoS requirements. Note that the receiver can also help in the modulation and coding scheme (MCS) selection process by sending its estimate to the transmitter as shown in Figure 6.4. Note that in this process, a fixed modulation is used to transmit the packet headers in order to let the receiver(s) know what MCS was used by the transmitter.

In wireless metropolitan area networking (WMAN), a concatenated FEC scheme is used as shown in Figure 6.5. This serial concatenation employs a Reed-Solomon

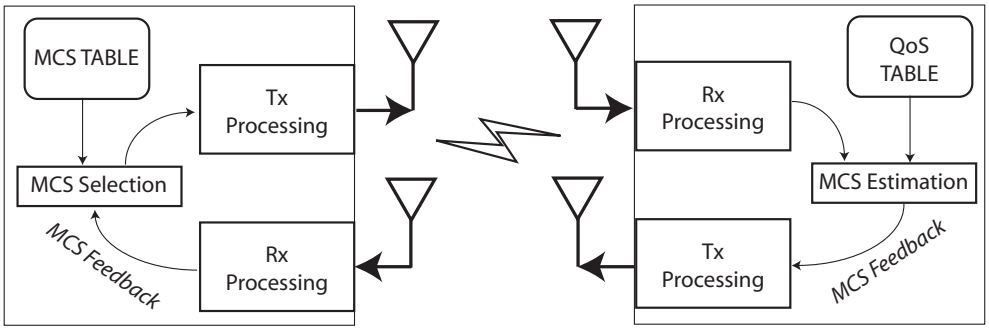


Figure 6.4: Illustration of a practical link adaptation technique utilizing various modulation and coding schemes (MCS) [1].

Figure 6.5: Concatenated coding scheme for wireless metropolitan area networking (WMAN) single carrier modulation format 2 (SC2) [2].

outer code and a rate-compatible convolutional inner code [2]. Between the outer and the inner code, an interleaver can be utilized to spread consecutive bits into separate symbols after modulation. The outer code can be shortened or punctured to enable variable block sizes and variable error-correction capability. Moreover, the inner code is also rate-compatible.

The reason for using two types of codes is to cope with both random errors and burst errors caused by the wireless channel. The inner code protects against the Gaussian channels with random channel bit errors. On the other hand, the outer coding provides extremely reliable error detection, and it can also correct several symbol errors in each block. On the decoding side, the Viterbi decoder does most of the work, and the RS decoder gets rid of most of the remaining errors.

The third generation (3G) cellular communication system utilizes a similar approach in order to provide the best possible service to users. A multilevel puncturing scheme provides symbol-by-symbol adaptive puncturing and interleaving. In the High-Speed Downlink Packet Access (HSDPA) link adaptation scheme, two scenarios

are considered: constant throughput and constant BER. When compared to fixed-rate coding in terms of signal-to-noise ratio (SNR) and throughput, varying code rate and modulation together provides better results [100]. Using HSDPA, modulation type and FEC coding rate are simultaneously adjusted according to the channel conditions as follows:

- **Modulation type:** Employ QPSK for noisy channels and 16QAM for clearer channels. QPSK is more robust and can tolerate higher levels of interference and noise but has lower transmission bit rate. 16QAM has twice the bit rate (*i.e.,* 4 bits/symbol) but is more prone to errors due to interference and noise. Hence it requires stronger FEC coding for high levels of interference.

- **FEC coding rate:** The FEC code-rate is $1/3$, but it can be varied effectively by bit puncturing and Hybrid automatic repeat-request (HARQ) with incremental redundancy. When the channel conditions are good (average throughput of 10 MAC-PDUs per TTI, where PDU and TTI stand for Protocol Data Unit and Transmission Time Interval), more bits are punctured and the information bit rate is increased. On the other hand, in poor link conditions (average throughput of 5 MAC-PDUs per TTI) all redundant bits are transmitted and the information bit rate drops. In very bad link conditions retransmissions occur due to HARQ, and bit rate further drops.

On clear channels using 16-QAM and close to 1/1 coding rate, HSDPA provides very high bit rates, of the order of 14 Mb/s. HSDPA can also provide reliable communications using QPSK and 1/3 coding rate (resulting bit rate is around 2.4 Mb/s) when the channel is noisy. The modulation and code-rate adaptation is performed up to 500 times per second [100].

Both the WMAN's concatenated FEC scheme and HSDPA's link adaptation scheme can be utilized in a MANET with the help of an advanced version of the MCS scheme shown in Figure 6.4. According to the channel error behavior (*i.e.,* burstiness of the channel) and link conditions (*i.e.,* SNR value of the channel) the advanced MCS can select the appropriate technique(s) from a pool of MCS tables. Each MCS table represents the available coding and/or modulation combinations that are offered by a specific

adaptive technique (*e.g.,* concatenated FEC and link adaptation). The chosen MCS configuration can be sent in the packet header or separately prior to the transmission. This way, nodes that receive the data packet will be able to decode it according to the given MCS configuration. We believe that specifics and details of such an advanced MCS need to be explored in order to optimize the MCS selection process and the number of MCS tables to be constructed.

In [101], the authors point out that adaptive FEC schemes used for video streaming applications have to be aware of both the constrained source-rate and the fixed available bandwidth. They present FEC codes that can adapt in real-time to provide higher source-packets recovery without changing the FEC block (N, K) pair constraint. This is accomplished by using partial recovery codes that can be employed to recover partial information in a scenario where the channel conditions prevent the complete recovery. The idea of adapting the protection level without changing the block-length or coding rate proves to be useful in real-time voice and video applications if changing the transmitted packet size is not possible. This will improve the packet recovery rate without changing either the modulation type or packet size.

There are several practical limits to be explored before adaptive modulation can be efficiently used. The need for a feedback channel between the transmitter and receiver is a fundamental drawback of a regular adaptive modulation and coding system. Estimating the channel conditions without the help of feedback is also possible, however, the reliability of the transmitter side estimation cannot be guaranteed because of the space varying nature of the wireless channel (*e.g.,* asymmetric links). Moreover, if the channel conditions are changing faster than can be reliably estimated, adaptive techniques will perform poorly.

6.3 MAC/Link Layer Adaptation

In wireless communications, the channel, which is the common interface that connects the nodes, is a shared resource. Thus, access to this shared resource needs to be coordinated either centrally or in a distributed fashion. The objective of controlled access is to avoid or minimize simultaneous transmission attempts (that will result in collisions) while maintaining a stable and efficient operating region for the whole network

Table 6.1: Media Access Used In Different Wireless Systems.

System	Media access
AMPS	FDMA
PACS	TDMA
PCS-2000	CDMA/TDMA
GSM	TDMA
WiFi	CSMA/CA
Bluetooth	FHSS/TDMA
HomeRF	FHSS/TDMA/CSMA

[77, 95, 102, 103]. MAC protocols are responsible for the access control and the way bandwidth is made available to the nodes of the network. Therefore, a MAC protocol can directly affect the maximum throughput, delay, jitter and energy efficiency of the network.

Table 6.1 shows the media access used in several wireless systems. Systems that must guarantee a certain QoS to the user, such as cellular systems, typically use controlled multiple access techniques. On the other hand, systems that make no guarantees about timely delivery of data, such as wireless data networks, often use random-access techniques. Several systems combine media access technologies, such as using time division multiple access (TDMA) with a code division multiple access (CDMA) protocol (*e.g.,* PCS-2000 [104]), assigning slots to users within a given frequency-hoping spread-spectrum (FHSS) protocol (*e.g.,* Bluetooth [3]), or using either TDMA or CSMA within a FHSS protocol (*e.g.,* HomeRF, where TDMA is used for real-time data delivery and CSMA is used for asynchronous delivery) [105].

The MAC layer can regulate the channel access and improve the QoS performance of the network through *collision avoidance* and *resource allocation*. In MANETs, collisions happen if users transmit their packets over a shared bandwidth and in an overlapping fashion in time, with no additional coding that would allow separation of simultaneously transmitted packets. In this case neither packet may be decoded successfully. Therefore, collision avoidance techniques often try to take away one of the three causes

of collisions (*i.e.,* using the same bandwidth, using the same time slot, and lack of code division). It is also possible to include transmit power control to the list of collision avoidance techniques.

6.3.1 Collision Avoidance

The collisions introduced by hidden terminals are often avoided in wireless networks by a four-way handshake prior to transmission [1]. The RTS/CTS handshake is typically coupled with a random backoff to avoid all nodes transmitting as soon as the channel becomes available. Another technique to avoid the effects of hidden nodes is busy tone transmission through setting a bit in a predetermined field on the control channel. In [106], an adaptive power control scheme, which utilizes handshakes and busy tone transmission, is proposed. The main ides is to adaptively vary the power level of packet transmissions so as to increase the possibility of channel reuse and reduce the amount of handshake overhead needed to avoid collisions. This approach makes use of exchanged RTS and CTS packets between two nodes to determine their relative distance. This information is then utilized to determine the power level of data packet transmission. Using lower power can increase channel reuse, and thus channel utilization. It also saves the precious battery energy of portable devices and reduces co-channel interference with neighboring nodes.

The authors in [107] propose a modified RTS/CTS exchange to help a multicast transmitter in selecting the optimal FEC coding rate that could enable all receivers in the multicast group to successfully receive the data transmission. While an RTS is used to notify the multicast group of a multicast transmission, a CTS time slot is divided into M mini-slots each reflecting a unique coding rate. The receivers select a coding rate based on their channel estimation and send the selected coding rate in the CTS packets sent to the transmitter. When good channel conditions are estimated, no CTS frames are sent. The transmitter encodes the frame with the strongest code rate selected from the busy mini-slots so that all receivers can receive it correctly. However, when no CTS transmission is sensed in any of the mini-slots, the transmitter assumes good channel conditions and sends data frame without FEC encoding. Since FEC is a computationally expensive process, varying the FEC rates based on the channel feedback from neighbors reduces the cost of FEC and helps maintain high throughput.

6.3.2 Resource Allocation

The second MAC layer technique for regulating the channel access and improving the QoS performance is resource allocation, which dictates how network resources such as power and bandwidth are allocated throughout the network. Resource allocation can thus be seen as the ability to accommodate different levels of service at any given time [108]. Different levels of QoS guarantees are important when the network capacity is insufficient, especially for real-time streaming multimedia applications that require fixed bit rate and are delay sensitive. In addition to the following service levels, by default, the control traffic should be given the highest priority.

- **Best-effort service:** (no need for QoS), basic connectivity with no guarantees and no differentiation between flows.
- **Differentiated service:** (soft QoS), some traffic needs to be treated better than the rest. This can be a statistical preference, not a hard and fast guarantee.
- **Guaranteed service:** (hard QoS), this is an absolute reservation of network resources for specific traffic.

We believe that in order to accommodate more types of traffic streams than is possible with simple prioritization, a parameterized specification for each traffic type can be used. Each traffic type is represented with a set of parameters that specifies the characteristics of the type of traffic in detail. Delay, jitter, and packet delivery ratio requirements can be used to specify the traffic type along with the transmission duration and data rate requirements. A MAC layer coordinator (*e.g.,* clusterhead) would decide on the priority level of the traffic by looking at its specified parameters and determine the channel assignment according to the QoS requirements of the traffic stream. Moreover, an adaptive mechanism is also needed to ensure fairness and overall improved QoS under varying traffic loads. This mechanism makes sure that the channel access is shared fairly among differentiated traffic streams by dynamically varying the priority assignment according to available bandwidth.

Adaptive techniques combined with cross-layer design can help achieve global optimization in multiuser multimedia communications in wireless mobile ad-hoc networks.

The MAC layer is where the adaptability is needed in order to accommodate the varying demands of multiuser packet data and multimedia applications. In addition to a flexible MAC that can dynamically allocate the channel resources according to the demand from the nodes, adaptive channel coding can also be utilized to satisfy different data rates and QoS requirements. In [94], there are two different types of allocation periods defined, the static allocation period (SAP) and the dynamic allocation period (DAP). Both allocation periods are contention-free. SAP slots are defined once and remain valid until a new beacon message removes them. On the contrary, DAP slots are assigned on an on-demand basis and announcement is done separately within the beacon associated with the given frame. The joint design of link and physical layers allows for increased data throughput while providing a wide range of QoS requirements due to this centralized, dynamic slot allocation technique.

Future challenges in this research area include determining the most suitable medium access techniques according to system requirements and characteristics along with cost and complexity constraints. In addition to the interoperability issues we mentioned in Section 6.2.1, fusing all the different adaptive techniques together to function as one unified layer is an extremely challenging task. Moreover, we believe that steps towards exploring the interaction between the existing techniques should be taken before trying to introduce new ones.

6.4 Adaptive Routing

The first objective of a routing protocol is to convey packets from a source to a destination (or group of destinations for multicast routing) with an acceptable quality of service (QoS) [109]. Specifically, QoS in multimedia communications requires maintaining a high enough packet delivery ratio (PDR), which is defined as the ratio of the number of data packets received by the destination nodes to the number of data packets generated at the source node. Moreover, keeping the packet delay low enough, and minimizing the jitter in packet arrival times are also crucial.

We can list techniques for improving the QoS of routing in ad-hoc networks according to the following three categories: (i) adapting the channel coding and routing to overcome the varying link conditions [48], (ii) choosing routes according to a cost

function (*i.e.,* cost-aware routing) in order to avoid unnecessary retransmissions and improve the delay performance [49–51], and (iii) varying the number of redundant links between the nodes of the network to increase the efficiency [52–54]. The common goal of these different types of approaches is to increase the reliability of relaying information over multiple hops and to multiple destinations in a wireless medium.

6.4.1 Routing with Adaptive Channel Coding

We have investigated adaptive coding and modulation in Section 6.2.2 in detail. Therefore, we keep our discussion in this section short. Lin et al. [48] point out that the time-varying nature of wireless channels is often ignored in ad-hoc routing scenarios. In order to overcome the performance loss due to the time-varying wireless channel, they propose a cross-layer channel adaptive routing protocol that utilizes an adaptive channel coding and modulation scheme (similar to the one given in Section 6.2.2) to dynamically adjust the amount of error protection. They achieve shorter delays and higher rates at the expense of a higher overhead in route set-up and maintenance.

6.4.2 Cost-aware Routing

The second adaptive routing technique we emphasize in this section is cost-aware routing, which utilizes an optimization criterion to compute the optimal route from source to destination (or tree from source to multicast members). Many researchers proposed that instead of trying to adapt the level of error protection, choosing routes that are more stable or less prone to channel errors would increase the QoS performance of the network [49–51]. The idea is to select better routes according to the predicted state of the links and prevent unnecessary packet losses and retransmissions. As a result, for slowly varying channel conditions one can achieve increased energy efficiency, and a stable and good quality communication route. However, predicting the state of multiple routes between two nodes and choosing a route that is more stable than the others may not be possible for highly dynamic channels.

Cost-aware routing, in general, can be an adaptive approach to routing, where the goal is to change the routing cost function according to channel conditions, QoS requirements, and node properties (*e.g.,* node energy, location, coverage, etc.). This way,

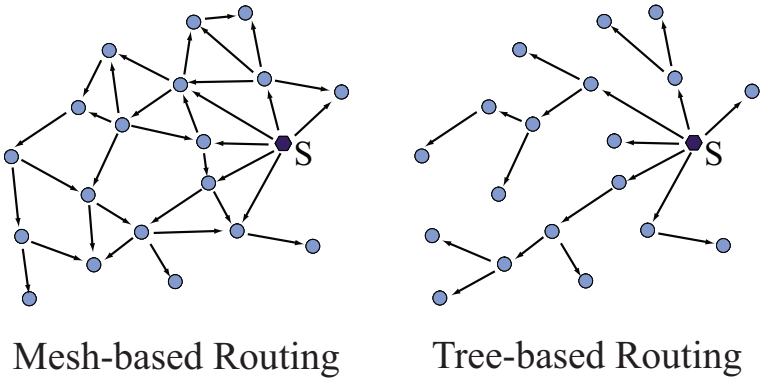


Figure 6.6: Illustration of mesh-based and tree-based multicast routing approaches.

Table 6.2: Routing Cost Functions.

Parameter	**Function**
Remaining Node Energy (E)	choose next hop: $min(\frac{1}{E})$
Link Quality (L_Q)	choose route: $max(\sum L_Q)$
Latency	choose route: $min(hops)$
Traffic Priority Level (TP)	route first: $max(TP)$

established routes tend to be more reliable and satisfy the QoS requirements. Table 6.2 lists some of the key parameters in determining the performance characteristics of the network such as reliability and robustness of the packet delivery, network lifetime, and QoS. For a chosen parameter, the goal of the cost function is to utilize the parameter in order to optimize network performance or satisfy the application requirements. Branches between the source and destination nodes are constructed in a way that minimizes the routing cost while satisfying the QoS requirements.

A multi-objective cost function for multi-hop routing should be able to adapt to the varying nature of the network parameters. For example, if we assume that remaining node energy is the only parameter we would like to take into account, we might end up with highly unreliable network performance. This is due to the fact that avoiding certain routes in order to maximize network lifetime can cause data packets to travel over longer, less reliable routes. Consequently, there will be an increased number of data packet drops and data packet delay will also increase due to longer routes. Although

we have listed ways to overcome the varying link conditions in Section 6.2.2, real-time applications have limited flexibility for data packet delay and jitter. Therefore, a better approach would be to vary the cost function according to the requirements of the data traffic.

- **No QoS (best effort service is enough):** Employ remaining node energy as the cost function in route management to increase the network lifetime.
- **File Transfer (all packets need to reach to the destination):** In addition to the "No QoS" approach, link quality along the route can be included in the route selection process to reduce the number of retransmissions. If there are more than one QoS level, each level needs to be included in the cost function.
- **Real-time Data Traffic (guaranteed service is needed):** Extra measures to ensure the timely arrival of data packets (in the order they have been generated) need to be taken. Node location, link quality, and traffic priority level are included in the cost function in order to reduce the number of hops data packets travel (*i.e.,* packet delay), to increase the reliability of the routes established, and to make sure real-time traffic is is prioritized over other traffic, respectively.

6.4.3 Adaptive Redundancy Routing

The third adaptive technique in ad-hoc routing is to vary the number of redundant links between the nodes of the network. Due to the success of mesh based routing protocols in maintaining stable communication in the face of changing link conditions, researchers investigated ways to reduce or adapt the number of redundant links in a route between source and destination according to the dynamic behavior of the network [52–54]. The common motivation is the fact that although utilizing multiple paths from senders to receivers results in higher reliability, in less dynamic environments the additional redundancy may not be needed in terms of reliability, and increased redundancy causes increased overhead. For example, Figure 6.6 shows the differences between mesh-based and tree-based multicast routing. In mesh-based routing, the same packet is received multiple times leading to a higher packet delivery ratio than that of tree-based routing at the expense of higher overhead.

We believe that in order to satisfy the multiple QoS requirements under varying channel conditions, a detailed and comprehensive ad-hoc routing protocol, which can adaptively vary its routing technique to meet these QoS requirements, is needed. As a simple example, a drastic change in the channel conditions will force the network to change its routing cost function parameter (*e.g.,* from latency to link quality). Moreover, there can be a gradual switch from multicast tree routing to mesh routing as the channel conditions continuously degrade. Ignoring or focusing on a specific version of the time and space varying nature of the wireless channels is a severe design drawback because the varying channel quality can lead to very poor overall route quality in turn.

Challenges will arise when it is not possible to work with a cross-layered stack that is needed for most of the adaptive approaches we surveyed so far. Moreover, sometimes relaying the link state information through the layers may not be feasible because of the time sensitive nature of the information. Therefore, it makes sense for each protocol layer to adapt to variations that are local to that layer instead of trying to achieve a global adaptation. However, if local adaptation is not sufficient multi-layer adaptation can be triggered for better performance. For example, poor link conditions can be detected easily at the physical layer and locally handled by increasing the transmit power or the level of error correction coding. When the physical layer remedies are not sufficient, the routing layer can activate the adaptive redundancy approach via multi-path routing to overcome the problem.

In conclusion, we would like to point out the fact that there are countless possible combinations of all the adaptive approaches we surveyed so far. This motivates us to define a new entity that would coordinate how and when these adaptive techniques should be used for a given set of channel conditions and QoS requirements.

6.5 A Qualitative Model for Adaptive Techniques

In this section, we provide a guideline towards modifying or combining the adaptive techniques we presented in the previous sections. It is possible to create new adaptive techniques that can adapt multiple parameters of the transmission scheme (including rate, power, and coding), MAC scheme, and routing protocol. However, in order to do this, joint optimization of the different techniques has to be used to meet a given per-

formance requirement. Adaptive modulation and coding schemes need to be combined with power adaptation to maximize spectral efficiency. In addition to this, information sharing between the layers or a cross-layer architecture can broaden the scope of the optimization process into the MAC and routing layers.

In the previous sections, we have summarized several adaptive approaches to increase the reliability and performance of multi-hop mobile ad-hoc networks. Each adaptive technique has its benefits, drawbacks, and a trade-off between two or more performance parameters. Utilizing one or more of these adaptive techniques requires a new way of processing the information shared between the communication stack layers and deciding according to the information whether or not a certain adaptive technique can be or needs to be used. Figure 6.7 shows the diagram for our adaptive optimization management approach, which handles the information and requirements provided by the communication layers and optimizes the adaptive technique selection process accordingly. The list of adaptive techniques may be modified depending on the hardware specifications, type of application, and all other factors that impose limits on or directly limit the overall system performance.

From a high level perspective, the adaptive optimization manager (AOM) can be seen as a decision making entity that enables the use of different adaptive techniques according to the outcome of a pre-determined logical decision making process. There can be two different ways to establish a multi-objective adaptive optimization and decision making process. (i) Implement all possible adaptive techniques along with a complex decision making manager that determines which techniques should be utilized for the given communication scenario and the channel condition. (ii) Limit the number of adaptive techniques to be utilized in order to enable a simpler decision making process (*e.g.,* just a look-up table with conditions and corresponding sets of techniques to be used) and avoid the need for a separate management entity. Either way, adaptation and cross-layer design (at least a certain level of information sharing between layers) are indispensable in achieving optimal performance in the face of varying conditions of the wireless channel.

There is an inherent redundancy between the adaptive approaches we summarized in Section 6.1 due to the fact that adaptation requires feedback mechanisms between the layers of the communication stack. Therefore, when implementing the adaptive opti-

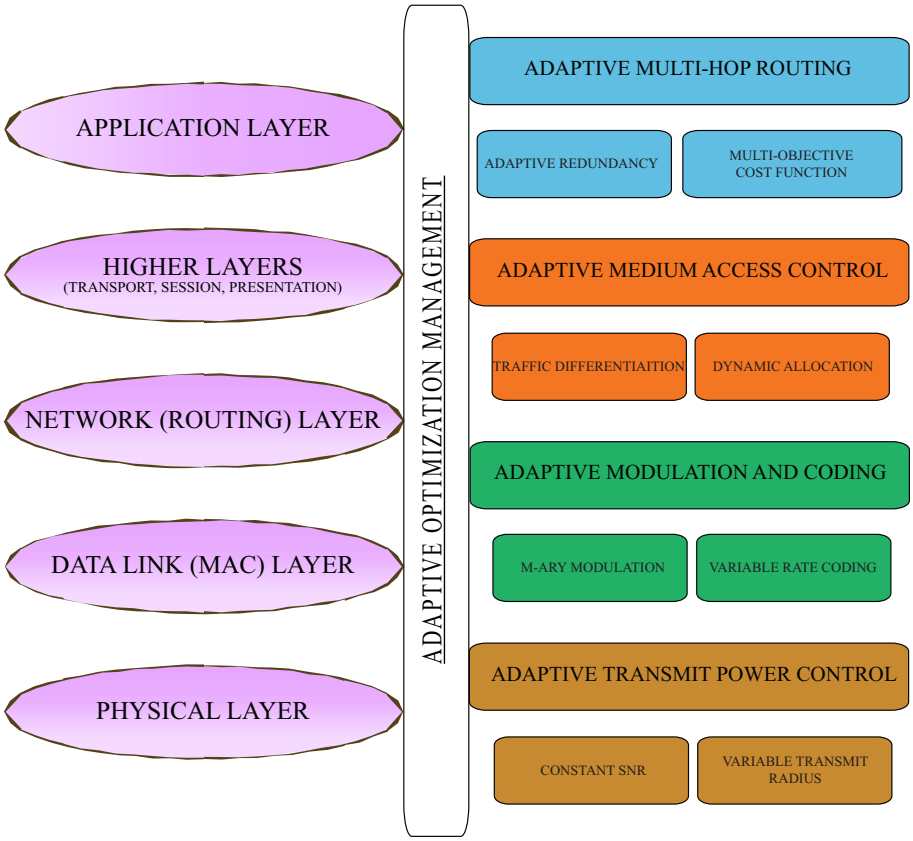


Figure 6.7: Diagram of the adaptive optimization management.

mization manager there will be shared entities within layers, which serve and distribute the information from the corresponding layers to multiple adaptive protocols. Determining the functions and responsibilities of these entities should be a part of the overall system performance optimization. For example, there may be a need for a physical layer entity that monitors the channel conditions, channel utilization, and packet reception history. If we design a separate entity for each one of the three parameters and store the corresponding information separately, we may be wasting our resources because of the possible redundancy between the separately collected information. On the other hand, if we design a single entity to keep track of the received packet information only, it will be hard if not impossible to estimate the channel conditions and utilization information. Therefore, the design of the information entities has to be carried out along with the adaptive optimization manager design.

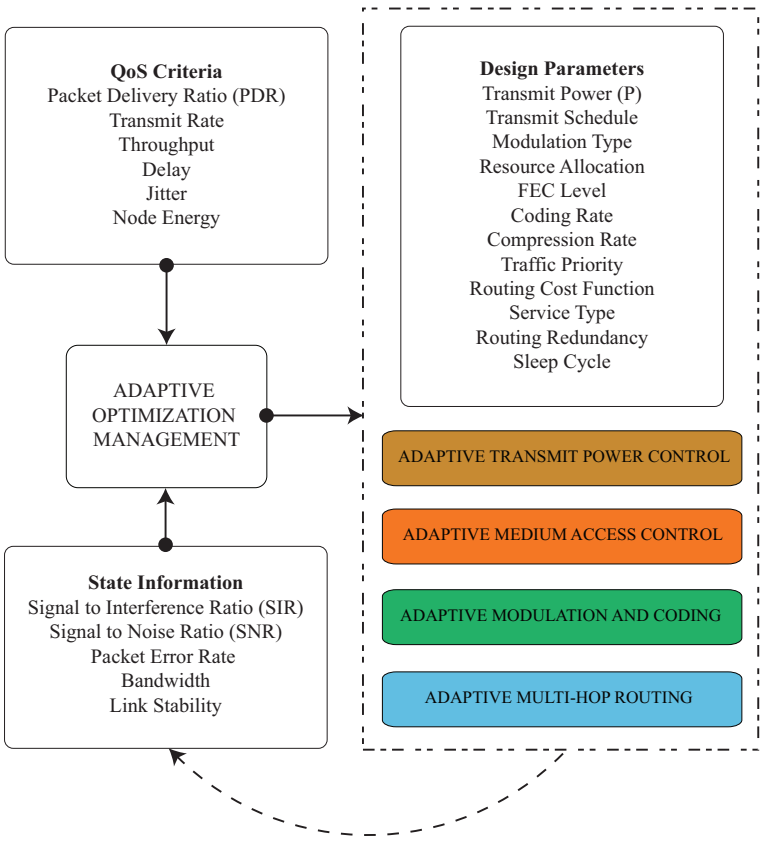


Figure 6.8: Information flow diagram for the AOM.

In Figure 6.8, we illustrate the information exchange and feedback mechanisms for the AOM. The AOM initially gathers the QoS criteria and then assesses the channel conditions in order to decide on which of the available techniques need to be employed. Moreover, according to the severity of the conditions or the difference between the current and desired levels of QoS, the AOM determines the required setup for a given technique (*e.g.,* coding level, modulation type, mesh redundancy level, etc.). After each adjustment, the QoS criteria are updated and a new optimization process begins. This process is recursive, and the QoS criteria history plays an important role in both predicting the next best configuration and determining the rate of optimization. Varying the rate of optimization can be considered as damping the system behavior in order to reach the steady state (*i.e.,* meet the QoS criteria) as soon as possible.

Optimization will generally focus on improving just one or two aspects of performance and will usually require a trade-off, where one factor is optimized at the expense of others. If the AOM fails to provide the required QoS levels due to extreme conditions, it needs to strike a balance between the QoS criteria that have to be sacrificed. Therefore, if a criterion has a higher priority than the others, the AOM has to be notified. In order to avoid conflicts between the criteria, pair-wise rules have to be set as to which criterion will have priority over the other.

As an example, consider the case where we want to achieve a certain level of PDR in a real-time routing (*e.g.,* unicast, multicast, and broadcast) scenario. There are several adaptive tools and schemes that we can employ to achieve this PDR goal. First, we have to determine what is crucial and what can be sacrificed in achieving consistent PDR performance in multicasting. Then, we choose the set of QoS criteria that we want to keep track of in order to provide the desired level of PDR. In particular, for the real-time routing scenario, we can aim to minimize the dropped packets (*e.g.,* vary the forward error correction on the packets or use adaptive TPC to achieve a desired level of SNR at each hop in the route), create delay optimized, redundant routes between each source-destination pair (*e.g.,* utilize cost function routing as well as adaptive routing redundancy), and prioritize the channel access for the nodes participating in the routing process and require a guaranteed service for the real-time routing stream.

Figure 6.9 shows the process diagram for the real-time routing example. The *status update*, generated by the adaptive optimization management, carries the update information as to which design parameters need to be changed (*e.g.,* increased or decreased). As we mentioned before, in some cases, local adaptation may not be feasible and/or enough to satisfy the required design parameters. For adaptive modulation and coding, and adaptive redundancy routing, a node can get away with informing only its neighboring nodes (*local adaptation*). However, in order to update a routing cost function or for routing service level updates, all the nodes participating in the routing process must be notified (*route adaptation*). Moreover, for adaptive medium access control updates, the medium access controller (*e.g.,* a clusterhead) may need to be notified before a node can get prioritized channel access. As for TPC, in some cases, all nodes (*global adaptation*) in the network have to update their transmission power in order to achieve lower packet delay values in ad-hoc routing. Therefore, the AOM must take the type of

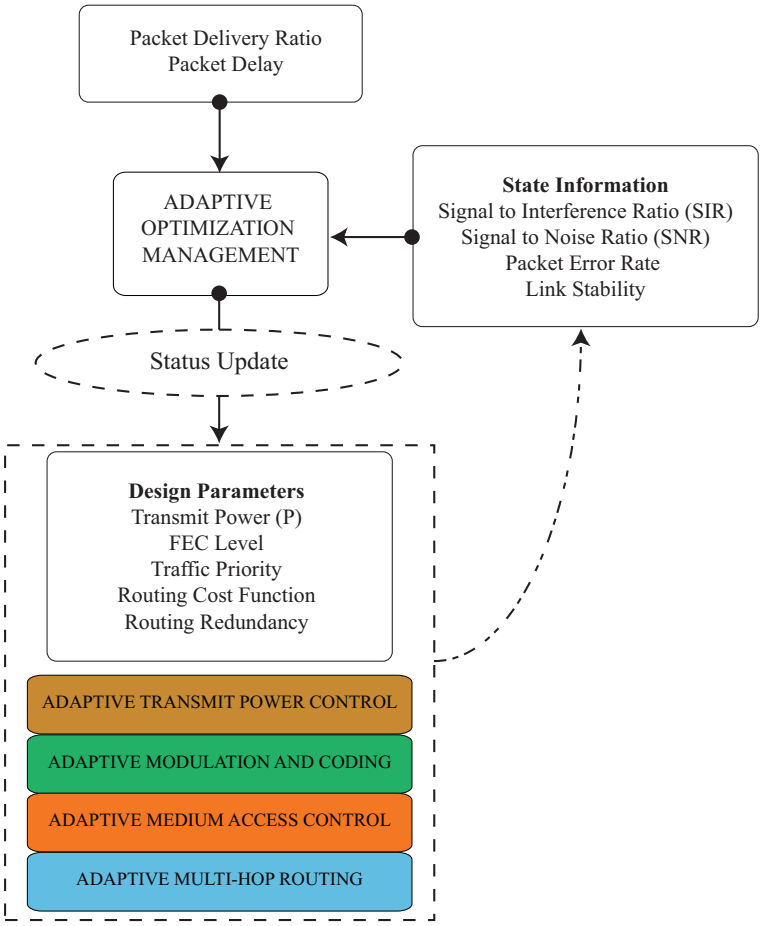


Figure 6.9: Example for real-time routing with PDR and packet delay requirements.

the adaptation (local, route, global) into consideration before selecting a specific design parameter to update.

In general, the AOM performs optimization locally at each node, however, if adapting locally is unsuccessful the AOM should increase the information exchange with higher layers for a broader response to the problem. We believe that for a given scenario with QoS requirements, system restrictions, and available adaptive techniques, it is possible to design an AOM to sustain the desired level of QoS.

6.6 Summary

The mobility of the nodes and the dynamic nature and poor performance of the underlying wireless communication channel require that mobile ad-hoc networks be optimized in order to be robust, and that they be adaptive to the link variations. In this chapter, we investigated various adaptive ways to overcome the performance loss due to the dynamic behaviors of the wireless channel and mobile ad-hoc networking. First, we looked at transmit power adaption, which can be used to keep the interference level under control in a wireless broadcasting scenario. Modulation and coding schemes together promise a greater flexibility and reliability while keeping the trade-off between the overhead and guaranteed QoS under control. Later, we emphasized the MAC layer's important role under the presence of differentiated streams. Then, we summarized adaptive routing techniques that help increase the flexibility and reliability in mobile ad-hoc networks by varying the amount of redundant transmissions or even the cost function according to the channel conditions. We pointed out many research directions that can be pursued to achieve an optimal system performance for a given set of conditions and QoS requirements [110].

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This research thesis jointly addressed several open research issues in improving the reliability and performance of mobile ad-hoc networks. We presented a multi-objective performance analysis of different classes of MAC protocols in the presence of channel noise. Moreover, we offered new protocols that provide better QoS support, including multi-level QoS and adaptive performance to meet the required QoS.

First, we investigated the impact of channel errors on the performance of MH-TRACE and IEEE 802.11, which are examples of coordinated and non-coordinated MAC protocols, respectively, through *ns*-2 simulations using the Gilbert-Elliot channel model. As expected, the impact of channel errors is more severe on MH-TRACE than IEEE 802.11 due to the dependence of MH-TRACE on robust control packet traffic. Nevertheless, the performance of MH-TRACE remains superior to that of IEEE 802.11, even in the presence of large channel errors. Hence, the major conclusion of this study is that coordinated MAC protocols are preferable over non-coordinated MAC protocols even under noisy channel conditions. If the channel conditions improve due to either a better channel or forward error correction utilized in the control traffic, then MH-TRACE performance loss due to the non-perfect channel will be similar to that of IEEE 802.11.

Next, we took a deeper look into the impact of channel errors on the energy efficiency and QoS performance of coordinated and non-coordinated MAC protocols. We

developed an analytical model for the performance of MH-TRACE as a function of network area, number of nodes and bit error rate (BER) of the channel. We presented ns-2 simulations both to demonstrate the validity of the analytical model and to show the degradation in the MAC protocols' (*i.e.,* IEEE 802.11 and MH-TRACE) performance with increasing BER. As expected, the impact of channel errors is more severe on MH-TRACE than IEEE 802.11 at extremely high BER levels due to the dependence of MH-TRACE on robust control packet traffic. Nevertheless, as the node density increases, MH-TRACE performs better than IEEE 802.11 (in terms of throughput and energy efficiency) even under very high BER levels due to its coordinated channel access mechanism.

In this study, we explored the performance of coordinated and non-coordinated MAC protocols as stand-alone entities under noisy channel conditions. However, by building upon our current results it is possible to extend our analysis to other layers, such as the network layer. Furthermore, we considered only real time voice communications in ad-hoc networks and this work can be extended into sensor networks with different flow models and expiration deadlines.

Lessons learned from the results of this study are not specific to MH-TRACE or IEEE 802.11. In fact, we developed our model to account for a generic schedule based coordinated MAC protocol, and the analytical model is shown to be in agreement with the simulations, which are specific to MH-TRACE and IEEE 802.11. Thus, the major conclusion of this study is that the energy efficiency and QoS performance of coordinated MAC protocols are superior to those of non-coordinated MAC protocols. The relatively better QoS performance of non-coordinated MAC protocols at extremely high BER levels is actually deceiving due to the fact that such a low level of QoS is not beneficial to the application layer. Finally, we point out that for higher data rates or node densities coordinated protocols are expected to perform better in terms of initial throughput due to their controlled access mechanisms.

In the second part of the thesis we focused on room for improvements within the traditional way of broadcasting. In order to bring us closer to the ultimate goal of optimal use of the broadcast channel, in this work we proposed and investigated the performance of a superposed multi-rate coding scheme in network-wide broadcasting and multicasting scenarios in MANETs. Superposed coding makes different informa-

tion rates simultaneously available, and this lets multicast group members decide which set of upstream nodes they need to listen to in order to maximize the overall ratio of data rate/delay and minimize the energy dissipation. We investigated both ends of the delay vs. quality trade-off by using MC-TRACE and multi-rate MC-TRACE (MMC-TRACE). We also explored the characteristics of multi-rate multicasting through detailed simulations with MMC-TRACE. MMC-TRACE offers low and high rates simultaneously (*i.e.,* through a single transmission), and, as a result, each multicast group member can individually request either of these available rates.

The need for multi-rate multicasting will continually increase as the number of subscription based wireless services, such as mobile TV broadcasting and news feeds, offered to consumers increase. In addition to the variety of services offered, the different communication capabilities and needs of the continuously improving hand-held mobile devices must be supported by MANETs. We believe that multi-rate multicasting will provide an efficient solution to satisfying the varied needs of all these devices through a single multicast transmission, thereby reducing energy dissipation and bandwidth usage.

In the third part of the thesis, we proposed an adaptive redundancy algorithm, implemented on MC-TRACE (a tree-based multicasting approach), and compared it to a mesh-based multicast protocol (ODMRP). We explored the limits of redundancy while trying to strike a balanced trade-off between the amount of redundancy and energy consumption. Through ns-2 simulations, we showed that adaptive redundancy maintains high packet delivery ratios at higher protocol efficiency when compared to both non-adaptive mesh and tree-based multicast protocols.

In this part of the thesis, we summarize the work that has been done to overcome the performance loss introduced by lossy channels. We surveyed several existing techniques for combating the effects of lossy channels in MANETs. We began by discussing physical layer techniques, and continue with the MAC layer techniques of collision avoidance and resource allocation and the issues involved in adapting these approaches for mitigating the effects of channel errors. We detailed routing layer techniques such as cost-aware routing and multi-path routing, including mesh-based multicasting, as well as the adaptation parameters for these techniques. Based on this discussion, we provided a guide towards an adaptive optimization manager that can

manage these decisions in an integrated and coherent manner. We pointed out the fact that solutions targeting individual layers of the communication stack often provide relief for specific situations, where it is possible to modify only a single layer. However, significant performance gains are possible if one combines the effort at different layers and adaptively optimizes this collective effort according to varying channel conditions and QoS constraints.

7.2 Future Work

Following the research described in this thesis, a number of projects, which would further improve the reliability and performance of real-time communications, could be taken up in the future. Below there is a list of future work directions:

- The TRACE family of protocols rely heavily on the performance of the TDMA based channel access. There are several parameters (*e.g.,* number of frames in a superframe, data slot length, clusterhead formation rules, etc.) in this TDMA based access scheme that can be adapted to the varying network topology, channel characteristics and traffic load. Specifically, the optimum value of the number of frames is the one that minimizes the combined effect of both collisions and dropped packets. By developing an analytical model that determines the optimal TDMA structure under various settings, one can determine the performance gains that could be achieved by adapting protocol parameters as network conditions change.

- The multi-rate multicasting approach we presented in this thesis can be used with other multicasting protocols and therefore, future work concentrating on implementing this approach into other multicasting protocols can focus on the delay-throughput trade-off characteristics achieved by using additional layers in the superposed coding process. Moreover, when using multicasting for applications such as file sharing, different forwarding schemes can be combined with multi-rate coding to achieve better energy efficiency.

- The adaptive redundancy approach presented in Chapter 5 showed the advantages provided by adapting the routing redundancy as channel conditions change. As a

future work, one can look at combining the approaches we presented in Section 2.5.2 (*i.e.,* adapting the channel coding to overcome the varying link conditions, choosing routes according to link conditions in order to avoid unnecessary retransmissions and improve the delay performance, and varying the number of redundant links between the nodes of the network to increase the efficiency) in order to provide the best trade-off between reliability and energy consumption, allowing multicast protocols to support QoS in mobile ad hoc networks for minimum cost.

- In Chapter 6, we presented a summary of adaptive techniques for combating the problem of lossy links in mobile ad hoc networks. We pointed out several future research directions including the design of an adaptation optimization manager (AOM) and its rules. The adaptation may include all layers, from the physical layer to the application layer. This type of cross-layer protocol design is a challenging task, requiring a broad spectrum of techniques focusing on many different areas such as communications, signal processing, and network/information theory. While challenging, future work regarding this approach promises great improvements in reliability and efficiency for mobile ad hoc networks.

Bibliography

[1] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[2] T. Cooklev, *Wireless Communication Standards*. IEEE Press, 2004.

[3] "Bluetooth Protocol." http://www.bluetooth.com.

[4] W. R. Crowther, "A system for broadcast communications: reservation aloha." Presented at the Hawaii International Conference on System Sciences (HICCS), 1973.

[5] J.-F. Frigon, V. C. M. Leung, and H. C. B. Chan, "Dynamic reservation TDMA protocol for wireless ATM networks," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 370–383, 2001.

[6] D. J. Goodman, R. Valenzuela, K. Gayliard, and B. Ramamurthi, "Packet reservation multiple access for local wireless communications," *IEEE Transactions on Communications*, vol. 37, pp. 885–890, 1989.

[7] D. J. Goodman and S. W. Wei, "Efficiency of packet reservation multiple access," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 170–176, 1991.

[8] B. O'Hara and A. Petrick, *The IEEE 802.11 Handbook: A Designer's Companion*. IEEE Press, 1999.

[9] T. M. Siep, I. C. Gifford, R. C. Braley, and R. F. Heile, "Paving the way for personal area network standards: an overview of the IEEE 802.15 working group for wireless personal area networks," *IEEE Personal Communications Magazine*, vol. 7, pp. 37–43, 2000.

[10] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 1265–1275, 1997.

[11] C. Zhu and M. S. Corson, "A five phase reservation protocol (FPRP) for mobile ad hoc networks," *ACM/Kluwer Wireless Networks Journal*, vol. 7, pp. 371–384, 2001.

[12] Y. Yi, M. Gerla, and T. J. Kwon, "Efficient flooding in ad hoc networks: a comparative performance study," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1059–163, 2003.

[13] A. Ephremides and T. Truong, "Scheduling broadcasts in multihop radio networks," *IEEE Transactions on Communications*, vol. 38, pp. 456–460, 1990.

[14] H. Lim and C. Kim, "Multicast tree construction and flooding in wireless ad hoc networks," in *Proceedings of the ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWIM)*, pp. 61–68, 2000.

[15] S. Ni, Y. Tseng, and J. Shen, "The broadcast storm problem in a mobile ad hoc network," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MOBICOM)*, pp. 151–162, 1999.

[16] W. Peng and X. Lu, "On the reduction of broadcast redundancy in mobile ad hoc networks," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC)*, pp. 129–130, 2000.

[17] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint relaying: an efficient technique for flooding in mobile wireless networks," Tech. Rep. 3898, The French National Institute for Research in Computer Science and Control (INRIA), 2000.

[18] J. Sucee and I. Marsic, "An efficient distributed network-wide broadcast algorithm for mobile ad hoc networks," Tech. Rep. 248, Rutgers University Center for Advanced Information Processing (CAIP), NJ, 2000.

[19] K. Tang and M. Gerla, "MAC layer broadcast support in 802.11 wireless networks," in *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pp. 544–548, 2000.

[20] B. Williams, "Network wide broadcasting protocols for mobile ad hoc networks," m.sc. dissertation, Colorado School of Mines, Golden, CO, 2002.

[21] B. Tavli and W. B. Heinzelman, "MH-TRACE: Multi hop time reservation using adaptive control for energy efficiency," *IEEE Journal on Selected Areas of Communications*, vol. 22, pp. 942–953, June 2004.

[22] B. Tavli and W. B. Heinzelman, "NB-TRACE: Network-wide broadcasting through time reservation using adaptive control for energy efficiency," in *Proceedings of the IEEE WCNC*, pp. 2076–2081, 2005.

[23] B. Tavli and W. B. Heinzelman, "MC-TRACE: Multicasting through time reservation using adaptive control for energy efficiency," in *Proceedings of the IEEE Military Communications Conference*, 2005.

[24] E. N. Gillbert, "Capacity of a burst-noise channel," *Bell System Tech. J.*, vol. 39, pp. 1253–65, Sept 1960.

[25] E. O. Elliott, "Estimates of error rates for codes on bursty-noise channels," *Bell System Tech. J.*, vol. 42, pp. 1977–97, Sept 1963.

[26] J. R. Yee and E. J. Weldon, "Evaluation of the performance of error correcting codes on a Gilbert channel," *IEEE Trans. Commun.*, vol. 43, pp. 2316–23, Aug 1995.

[27] J. L. Lemmon, "Wireless link statistical bit error model," Tech. Rep. NTIA Report 02-394, U.S. Department of Commerce, Jun 2002.

[28] M. Zorzi and R. R. Rao, "Lateness of probability of a retransmission scheme for error control on a two-state Markov channel," *IEEE Trans. Commun.*, vol. 47, pp. 1537–48, Oct 1999.

[29] J. McDougall and S. Miller, "Sensitivity of wireless network simulations to a two-state Markov model channel approximation," in *Proceedings of the Global Telecommunications Conference (GLOBECOM)*, 2003.

[30] A. Konrad, B. Zhao, A. Joseph, and R. Ludwig, "A Markov-based channel model algorithm for wireless networks," *Wireless Networks*, vol. 9, pp. 189–199, May 2003.

[31] A. Gurtov and S. Floyd, "Modeling wireless links for transport protocols," *ACM Computer Communication Review*, vol. 34, no. 2, pp. 85–96, 2004.

[32] E. Modiano, "An adaptive algorithm for optimizing the packet size used in wireless arq protocols," *Wireless Networks*, vol. 5, no. 4, pp. 279–286, 1999.

[33] T. Issariyakul, E. Hossain, and D. I. Kim, "Medium access control protocols for wireless mobile ad hoc networks: Issues and approaches," *Wiley Interscience Wireless Communications and Mobile Computing*, vol. 3, pp. 935–958, Dec 2003.

[34] J.-C. Chen, K. M. Sivalingam, P. Agrawal, and S. Kishore, "A comparison of MAC protocols for wireless local networks based on battery power consumption," in *Proceedings of the Seventeenth International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 1998.

[35] Y. Lu, Y. Zhong, and B. Bhargava, "Packet loss in mobile ad hoc networks," tech. rep., Purdue University, Report CSD-TR 03-009, 2003.

[36] M. Takai, J. Martin, and R. Bagrodia, "Effects of wireless physical layer modeling in mobile ad hoc networks," in *Proceedings of the 2001 ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2001.

[37] P. Lettieri and M. B. Srivastava, "Adaptive frame length control for improving wireless link throughput, range, and energy efficiency," in *Proceedings of the Seventeenth International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 1998.

[38] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, Jan 1972.

[39] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Transactions on Information Theory*, vol. 19, pp. 197–207, Mar 1973.

[40] K. Jung and J. M. Shea, "Simulcast packet transmission in ad hoc networks," *IEEE Journal on Selected Areas of Communications*, vol. 23, pp. 486–495, March 2005.

[41] T. Sun, R. Wesel, M. Shane, and K. Jarett, "Superposition turbo TCM for multirate broadcast," *IEEE Transactions on Communications*, vol. 52, pp. 368–371, Mar 2004.

[42] V. Gaddam and D. Birru, "A newly proposed ATSC DTV system for transmitting a robust bit-stream along with the standard bit-stream," *IEEE Trans. on Cons. Electronics*, vol. 49, no. 4, pp. 933–938, 2003.

[43] H. Moustafa and H. Labiod, "A performance comparison of multicast routing protocols in ad hoc networks," in *IEEE PIMRC*, pp. 497–501, 2003.

[44] D. Bertsekas and R. Gallager, *Data Networks*. Prentince Hall, 1992.

[45] C. W. Wu and Y. C. Tay, "AMRIS: a multicast protocol for ad hoc wireless networks," in *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pp. 25–29, 1999.

[46] S. J. Lee, W. Su, and M. Gerla, "On-demand multicast routing protocol in multihop wireless mobile networks," *ACM/Kluwer MONET*, vol. 7, pp. 441–453, 2002.

[47] J. G. Jetcheva and D. B. Johnson, "Adaptive demand-driven multicast routing in multi-hop wireless ad hoc networks," in *ACM MOBIHOC*, pp. 33–34, 2001.

[48] X.-H. Lin, Y.-K. Kwok, and V. Lau, "A quantitative comparison of ad hoc routing protocols with and without channel adaptation," *IEEE Transactions on Mobile Computing*, vol. 4, pp. 111–128, Apr 2005.

[49] R. Agrawal, "Performance of routing strategy (bit error based) in fading environments for mobile adhoc networks," in *IEEE Personal Wireless Communications Conference*, pp. 554–556, Jan 2005.

[50] N. Wisitpongphan, G. Ferrari, S. Panichpapiboon, J. Parikh, and O. Tonguz, "QoS provisioning using BER-based routing in ad hoc wireless networks," in *IEEE Vehicular Technology Conference*, pp. 2483–2487, May 2005.

[51] L. Tan, P. Yang, and S. Chan, "An error-aware and energy efficient routing protocol in MANETs," in *Computer Communications and Networks ICCCN*, pp. 1095–2055, Aug 2007.

[52] R. Vaishampayan, J. Garcia-Luna-Aceves, and K. Obraczka, "An adaptive redundancy protocol for mesh based multicasting," *Elsevier Computer Communications*, vol. 30, pp. 1015–1028, Mar 2007.

[53] M. Burmester, T. V. Le, and A. Yasinsac, "Adaptive gossip protocols: Managing security and redundancy in dense ad hoc networks," *Elsevier Ad-hoc Networks*, vol. 5, pp. 313–323, Apr 2007.

[54] M. Yoshida, M. Terada, and T. Miki, "Adaptive sector-based flooding for mobile ad hoc networks," *IEICE Transactions on Communications*, vol. 90, no. 4, pp. 788–798, 2007.

[55] S. Lee and C. Kim, "Neighbor supporting ad hoc multicast routing protocol," in *MobiHoc*, Aug 2000.

[56] H. Zimmermann, "Osi reference model - the iso model of architecture for open systems interconnection," *IEEE Transactions on Communications*, vol. 28, pp. 425–432, Apr 1980.

[57] A. Chandra, V. Gummalla, and J. O. Limb, "Wireless medium access control protocols," *IEEE Communications Surveys and Tutorials*, vol. 3, pp. 2–15, 2000.

[58] W. Stallings, *Wireless Communications and Networks*. Prentice-Hall, 2002.

[59] D. J. Goodman and S. X. Wei, "Efficiency of packet reservation multiple access," *IEEE Trans. Vehic. Tech.*, vol. 40, pp. 170–176, Feb 1991.

[60] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal of Selected Areas in Communications*, vol. 15, no. 7, pp. 1265–1275, 1997.

[61] B. Tavli and W. B. Heinzelman, "TRACE: Time reservation using adaptive control for energy efficiency," *IEEE Journal on Selected Areas of Communications*, vol. 21, pp. 1506–1515, Dec 2003.

[62] V. Kanodia, C. Li, A. Sabharwal, B. Sadaghi, and E. Knightly, "Distributed multi-hop scheduling and medium access with delay and throughput constraints," in *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2001.

[63] D. A. Maltz, J. Broch, and D. B. Johnson, "Lessons from a full-scale multihop wireless ad hoc network testbed," *IEEE Personal Commun. Mag.*, vol. 8, pp. 8–15, Feb 2001.

[64] D. McKenna, "Mobile platform benchmarks, a methodology for evaluating mobile computing devices," tech. rep., Transmeta Corporation, 2000.

[65] L. van Hoesel, T. Nieberg, H. Kip, and P. Havinga, "Advantages of a TDMA based, energy-efficient, self-organizing MAC protocol for WSNs," in *Proceedings of the IEEE Semiannual Vehicular Technology Conference, VTC Spring*, 2002.

[66] H. P.J.M. and S. G.J.M., "Energy-efficient TDMA medium access control protocol scheduling," in *Proceedings of Asian International Mobile Computing Conference*, 2000.

[67] D. Kotz, C. Newport, and C. Elliot, "The mistaken axioms of wireless-network research," Tech. Rep. TR2003-467, Dartmouth College Computer Science Department, Jul 2003.

[68] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[69] J. G. Proakis, *Digital Communications*. New York: Mc Graw-Hill, 1989.

[70] B. Tavli and W. B. Heinzelman, "A survey of energy efficient network protocols for wireless networks," *Kluwer Wireless Networks Journal*, vol. 7, pp. 443–458, 2001.

[71] T. Numanoglu, B. Tavli, and W. B. Heinzelman, "The effects of channel errors on coordinated and non-coordinated medium access control protocols," in *Proceedings of IEEE International Conference on Wireless and Mobile Computing*, vol. 1, pp. 58–65, Aug 2005.

[72] "Network Simulator - ns." http://www.isi.edu/nsnam/ns, 2003.

[73] T. Numanoglu, B. Tavli, and W. B. Heinzelman, "An analysis of coordinated and non-coordinated medium access control protocols under channel noise," in *Proceedings of Military Communication Conference (MILCOM)*, 2005.

[74] T. Numanoglu and W. B. Heinzelman, "Energy efficiency and error resilience in coordinated and non-coordinated medium access control protocols," *Elsevier Computer Communications Journal*, vol. 29, pp. 3493–3506, Nov 2006.

[75] H. Dong and J. Gibson, "Structures for SNR scalable speech coding," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, pp. 545–557, March 2006.

[76] H. T. Friis, "A note on a simple transmission formula," in *Proceedings of Institute of Radio Engineers*, vol. 24, pp. 254–256, 1946.

[77] T. S. Rappaport, *Wireless Communications*. Prentice-Hall, 1996.

[78] B. Rimoldi, "Successive refinement of information: characterization of the achievable rates," *IEEE Transactions on Information Theory*, vol. 40, pp. 253–259, Jan 1994.

[79] P. Gupta and P. R. Kumar, *Critical power for asymptotic connectivity in wireless networks*. Birkhuser, Boston, 1998.

[80] W. B. Heinzelman, *Application-specific protocol architectures for wireless networks*. PhD thesis, MIT, Cambridge, 2000.

[81] Y. Chen, E. G. Sirer, and S. B. Wicker, "On selection of optimal transmission power for ad hoc networks," in *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*, vol. 2, p. 300, 2003.

[82] W. B. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Negotiation-based protocols for disseminating information in wireless sensor networks," *Wireless Networks*, vol. 8, pp. 169–185, 2002.

[83] L. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad-hoc networking environment," in *IEEE INFOCOM, Alaska*, pp. 1548–1557, Apr 2001.

[84] T. Numanoglu, W. B. Heinzelman, and B. Tavli, "Multi-rate support for network-wide broadcasting in MANETs," in *Lecture Notes in Computer Science: Networking*, May 2007.

[85] T. Numanoglu and W. B. Heinzelman, "Multi-rate Multicasting in MANETs," *Under revision, previously submitted to IEEE Transactions on Mobile Computing Journal*.

[86] S. Lee and M. Gerla, "On-demand multicast routing protocol," in *IEEE WCNC*, Sep 1999.

[87] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Communications Magazine*, vol. 43, pp. 23–30, 2005.

[88] B. Tavli and W. B. Heinzelman, "Energy and spatial reuse efficient network wide real-time data broadcasting in mobile ad hoc networking," *IEEE Trans on Mobile Computing*, vol. 10, pp. 1297–1312, 2006.

[89] S. Oh, J.-S. Park, and M. Gerla, "E-ODMRP: enhanced ODMRP with motion adaptive refresh," *Wireless Communication Systems, 2005. 2nd International Symposium on*, pp. 130–134, Sept. 2005.

[90] T. Numanoglu and W. B. Heinzelman, "Improving QoS under lossy channels through adaptive redundancy," in *Mobile Ad Hoc and Sensor Systems, 2008.*, pp. 509–510, 29 2008-Oct. 2 2008.

[91] T. Numanoglu and W. B. Heinzelman, "Improving QoS in multicasting through adaptive redundancy," in *(to appear) IEEE Wireless Communications and Networking Conference, 2009.*, Apr 2009.

[92] T. Numanoglu and W. B. Heinzelman, "Improving QoS in multicasting through adaptive redundancy," *Under review, submitted to IEEE Transactions on Mobile Computing Journal*.

[93] S. Nanda, K. Balachandran, and S. Kumar, "Adaptation techniques in wireless packet data services," *IEEE Communication*, vol. 38, pp. 54–56, 2000.

[94] R. Badra and B. Daneshrad, "Adaptive link layer strategies for asymmetric high-speed wireless communications," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 429–438, 2002.

[95] A. S. Tanenbaum, *Computer Networks*. Prentice-Hall, 1996.

[96] J. Gomez and A. T. Campbell, "Variable-range transmission power control in wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 87–99, 2007.

[97] K. Mori, "Adaptive transmission power control in CDMA slotted-ALOHA radio communications," in *International Conference on Universal Personal Communications*, vol. 2, pp. 1137–1141, Oct 1998.

[98] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4–5, pp. 265–444, 2008.

[99] N. Nikaein, H. Labiod, and C. Bonnet, "MA-FEC: a QoS-based adaptive FEC for multicast communication in wireless networks," in *IEEE International Conference on Communications*, vol. 2, pp. 954–958, Jun 2000.

[100] "3rd Generation Partnership Project - 3GPP." http://www.3gpp.org/specs/specs.htm, 2008.

[101] S. Karande and H. Radha, "Rate-constrained adaptive fec for video over erasure channels with memory," in *International Conference on Image Processing*, vol. 4, pp. 2539–2542, Oct 2004.

[102] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*. Wiley, 1995.

[103] L. L. Peterson and B. S. Davie, *Computer Networks*. Academic Press, 2000.

[104] V. Garg and J. Wilkes, *Wireless and personal communication systems*. Prentice-Hall, 1996.

[105] "HomeRF Protocol." http://www.homerf.com.

[106] S. Wu, Y. Tseng, and J. Sheu, "Intelligent medium access for mobile ad hoc networks with busy tones and power control," *IEEE JSAC*, vol. 18, pp. 1647–1657, 2000.

[107] A. Basalamah and T. Sato, "Adaptive FEC Reliable Multicast MAC Protocol for WLAN," in *IEEE Vehicular Technology Conference*, pp. 244–248, Oct 2007.

[108] S. Vegesna, *IP Quality of Service*. Indianapolis: Cisco Press, 2000.

[109] J. Janssen, D. D. Vleeschauwer, G. H. Petit, R. Windey, and J. M. Leroy, "Delay bounds for voice over IP calls transported over satellite access links," *ACM/Baltzer Mobile Networks and Applications Journal*, vol. 7, pp. 79–89, 2002.

[110] T. Numanoglu and W. B. Heinzelman, "Adaptive techniques for mitigating the effects of lossy channels in mobile ad hoc networks," *Under review, submitted to ACM Mobile Computing and Communications Review*.