Cross-Layer Design with Multi-Packet Reception, MAC, and Network Coding in Multi-Hop Networks

by

Jason M. Cloud

B.S., Colorado School of Mines (2002)

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Signature of Author ....
Department of Electrical Engineering and Computer Science
May 20, 2011

Certified by ....
Muriel Médard
Professor of Electrical Engineering
Thesis Supervisor

Certified by ....
Linda Zeger
Technical Staff, MIT Lincoln Laboratory
Thesis Supervisor

Accepted by ....
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Cross-Layer Design with Multi-Packet Reception, MAC, and Network Coding in Multi-Hop Networks

by

Jason Cloud

Submitted to the Department of Electrical Engineering and Computer Science
on May 20, 2011 in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

A cross-layer design approach is proposed that can be used to optimize the cooperative use of multi-packet reception (MPR) and network coding. A simple and intuitive model is constructed for the behavior of an opportunistic network coding scheme called COPE proposed by Katti *et. al.*, MPR, the 802.11 MAC, and their combination. The model is then applied to key small canonical topology components and their larger counterparts. The results obtained from this model match the available experimental results with fidelity. Using this model, fairness allocation by the 802.11 MAC is shown to significantly impede performance and cause non-monotonic saturation behaviors; hence, a new MAC approach is devised that not only substantially improves throughput by providing *monotonic* saturation but provides fairness to *flows* of information rather than to *nodes*. Using this improved MAC, it is shown that cooperation between network coding and MPR achieves super-additive gains of up to 6.3 times that of routing alone with the standard 802.11 MAC. Furthermore, the model is extended to analyze the improved MAC's asymptotic, delay, and throughput behaviors. Finally, it is shown that although network performance is reduced under substantial asymmetry or limited implementation of MPR to a central/bottleneck node, there are some important practical cases, even under these conditions, where MPR, network coding, and their combination provide significant gains.

Thesis Supervisor: Muriel Médard
Title: Professor of Electrical Engineering

Thesis Supervisor: Linda Zeger
Title: Technical Staff, MIT Lincoln Laboratory

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Acknowledgments

First and foremost, I would like to thank my wife, Jenny, for all of the love and support she has given me in my pursuit of furthering my education. She has made a lot of sacrifices throughout my studies and I am eternally thankful for everything she has done. I would not have been able to do it without her patience, understanding, love, and “appreciation”.

I am also thankful for the support and advice from my research advisors Professor Muriel Médard and Dr. Linda Zeger. Their advice, encouragement, patience, and understanding were essential to my academic development and have helped me through the difficult times at MIT. Thank you for providing me with the inspiration and motivation to complete my master studies.

Finally, I would like to thank my parents and brother for their support and unwavering love. They are a constant source of encouragement and I thank them for giving me the strength to tackle any challenge I come across.

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Contents

List of Figures

List of Tables

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Chapter 1

Introduction

The utility and use of wireless networks for a variety of applications has revolutionized how people live and conduct business. Whether it is the use of wireless home networks or cellular phones, wireless technologies inundate almost every aspect of everyday life. Unfortunately, network design and characterization of performance is still a difficult task especially with the multitude of network technologies available. Efficiently integrating these technologies into a coherent network is largely done through ad-hoc implementations without consideration of the impacts incurred on other aspects of the network design. This thesis provides a simple and intuitive approach to cross-layer wireless network design and shows the benefits of efficiently combining multiple performance enhancing technologies.

A simple model is the workhorse that is used to first understand the implications of various design strategies and then to provide a rough order of magnitude for the achievable gains from implementing different network technologies in multi-hop wireless networks. Primarily, the model will be used to determine the design implications of combining various medium access control (MAC) approaches, opportunistic network coding (NC), and multi-packet reception (MPR) into a single network. This thesis will consider the question of how these technologies interact.

1.1 Background and Motivation

Instead of treating information in a network as a "fluid" where it is simply routed and replicated, Ahlswede *et. al.* [3] proposed a method which employs coding at each node in the network that "mixes" this information together. In traditional computer networks, each node behaves like a switch or router. For example, each node receives information on an input link, or set of input links, and the node either forwards this information to an output link or replicates this information and sends it to a set of output links. From an information theoretic view point, there is no reason that each node should be restricted to simply behave as a router; rather, each node could receive information from all of the input links, perform an encoding operation, and then send the encoded information to all of the output links. The contribution from Ahlswede *et. al.* showed that the traditional methods of satisfying multicast requirements (one source to multiple sinks) in computer networks is, in general, not optimal and that network coding can optimally achieve these multicast requirements.

The initial work by Ahlswede *et. al.* did not suggest methods in which the network capacity could be achieved, but only showed that from an information theoretic perspective, that it was possible. This led to a wide range of research to find a coding technique that would provide the proposed gains in an arbitrary network. Using the results from [3], Li *et. al.* [4] provided a linear code for multicast in a network that achieves capacity; but noted that, in general, the linear codes that are produced using their algorithm are not necessarily the simplest possible. For example, they specifically noted that the linear code produced for a cyclic network is time varying which makes it difficult to implement in practice. Developing an algebraic framework limited to linear network codes for arbitrary networks, Koetter and Médard [5] proved that simple, linear time-invariant codes are sufficient to achieve the multicast capacity. Ho *et. al.* [6] and Lun *et. al.* [7] extended these results and provided decentralized approaches for implementing network coding and achieving the multicast capacity.

The majority of the above research focused on intra-session network coding. For example, only information from a single source is encoded together. Dougherty *et. al.* [8] asked the question of whether or not linear codes are sufficient for inter-session network (i.e., information from multiple sources are encoded together) and provided a counter-example showing that networks exist where there is no linear solution that achieves the multicast capacity. In addition, it is known that inter-session network coding is difficult and even if linear network codes are used, determining how to perform coding is a NP-hard problem [5]. Regardless, numerous heuristic inter-session network coding schemes have shown significant throughput gains [9, 10, 11]. One such heuristic coding scheme that will be the primary motivation for this thesis is COPE, proposed by Katti *et. al.* [1]. COPE uses the piggybacking technique developed by [9] and generalizes it to arbitrary networks. Chapter 2 will provide a more detailed explanation of COPE and develop a simple, intuitive method to model its performance.

The development of COPE led to various methods to model and analyze the experimental COPE results shown in Figure 1.1. Sengupta *et. al.* [12] modeled the unicast throughput as a linear programming problem considering the coding of only two packets at a time in a random mesh network. However, the experimental COPE results on the right side of Figure 1.1 show that on average nearly three packets are coded together at a time. Thus, the majority of the achievable gain originates from coded packets containing three or more native, or non-coded, packets and was therefore not considered in [12]. Le *et. al.* [13] provided an upper bound to the coding gain by using the number of codable packets as a key performance measure and focusing on a subset of the possible sets of all coding structures. Unfortunately, they did not address the interaction between network coding and the fairness provided by the MAC. As a result, their analysis provides throughput gains that are considerably smaller than those shown in the left side of Figure 1.1.

Zhao and Médard [2] accurately modeled the experimental results produced by [1] and showed that the fairness imposed by the 802.11 MAC explains the non-monotonic throughput behavior in the empirically obtained data. Fig. 1.2 shows a sample of the results

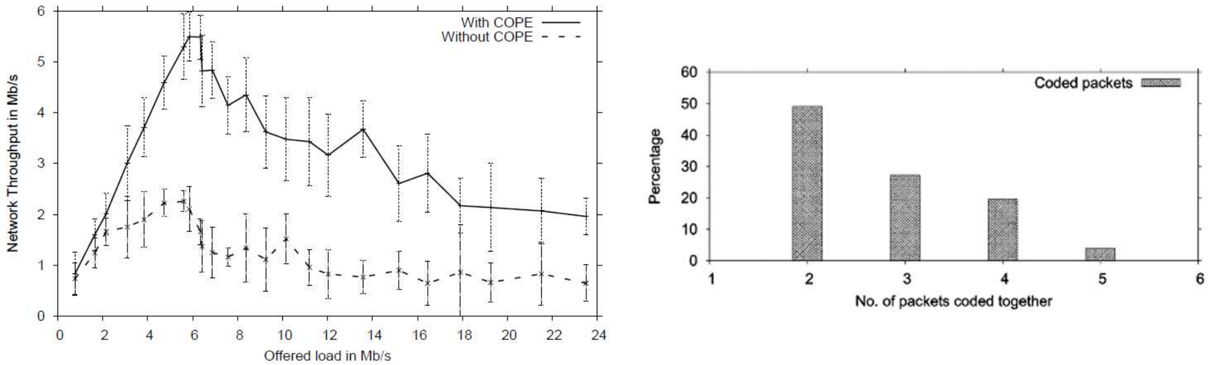

Figure 1.1: The empirical COPE performance data collected from a 20-node 802.11 wireless ad-hoc network test bed (left), and the distribution of the number of native packets encoded together at the peak throughput in the performance data. (right) [1]

obtained using their simplified model. Comparing the left side of Figure 1.1 with Figure 1.2 demonstrates that their model is consistent with the empirical data collected by [1], and illustrates the importance of considering the MAC implementation by showing the negative impacts of using an incorrect MAC in a multi-hop network. While Zhao and Médard showed that the MAC impacts the overall network performance, they only considered a single topology component, or coding structure. It was shown by Le *et. al.* [13], and by Seferoglu and Markopoulou [14] using TCP traffic with a coding scheme similar to COPE, that the throughput gains obtained from opportunistic network coding are dependent on the network topology. Chapters 3 and 4 will explore the expected coding gains when considering the interaction of network coding with the MAC as well as the effects of different topologies on network performance.

While the performance of COPE significantly increases network throughput, it does not completely alleviate the limitation of multi-user interference. With the development of new radio technologies such as orthogonal frequency division multiple access (OFDMA) and multiple-input multiple-output (MIMO), the ability to receive multiple packets simultaneously makes it possible to increase throughput and significantly increase network performance by reducing contention issues among users [15]. Extensive research has been conducted on multi-packet reception with uncoded traffic [16, 17, 18, 19]. For instance, the stability of

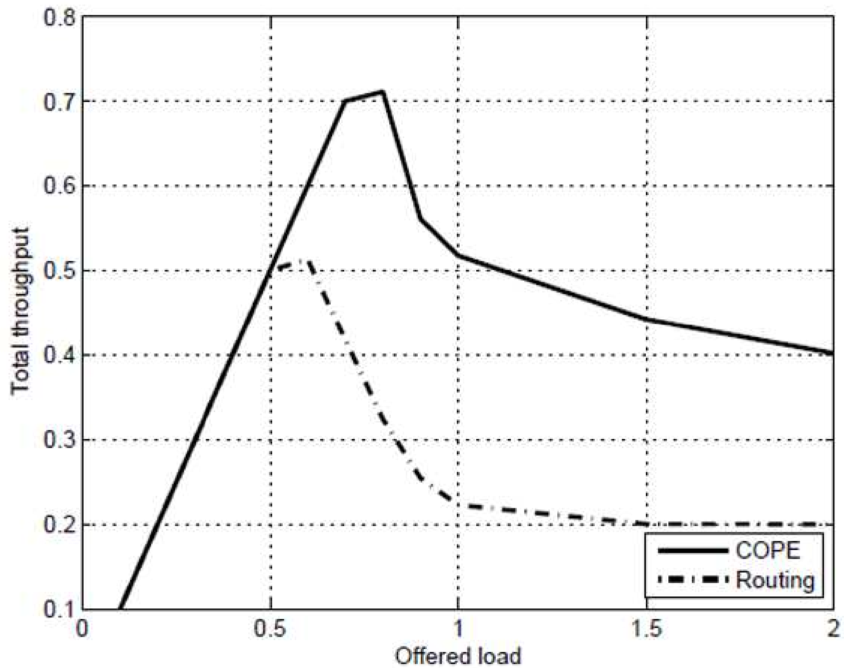

Figure 1.2: Throughput for COPE and non-COPE systems in a cross network with cross traffic and traffic generated at the center node from [2].

slotted ALOHA with multi-packet reception was studied by [18] and several protocols implementing multi-packet reception have been proposed by [17] and [19]; but little on the joint use of multi-packet reception and network coding exists. Garcia-Luna-Aceves *et. al.* [20] *compared* the use of network coding to multi-packet reception, but did not consider their *combined* use. Furthermore, Rezaee *et. al.* [21] provided an analysis of the combined use of network coding and multi-packet reception in a fully connected network, but did not consider the effects of bottlenecks or multi-hop traffic. Chapters 3 and 4 will characterize the performance resulting from combining the two network technologies in a multi-hop network and demonstrate some of the implementation challenges that must be overcome when joining the two.

1.2 Main Contribution and Thesis Outline

This thesis will first develop a model that is both simple and intuitive which will aid in the cross-layer design of multi-hop networks. Using the initial research performed by [2], the model developed will identify the fundamental behavior of the basic network elements (i.e., medium access control, network coding, multi-packet reception, etc.), and use these behaviors to characterize the gains from implementing each element both independently and jointly. Using the model, a new MAC approach will be designed that eliminates the non-monotonic saturation behavior shown in both the experimental COPE data and the initial model proposed by Zhao and Médard [2] in addition to providing fairness to *flows* rather than to *nodes*. The gains resulting from changes in topology, asymptotic traffic across the bottleneck in the network, and the asymptotic behavior in terms of per-node throughput and delay will also be determined with and without network coding and multi-packet reception. Primarily, the main contributions of this thesis are:

1. A network model will be developed that predicts the expected behavior from implementing various cross-layer technologies and also highlights specific design challenges. The primary examples used will be network coding, multi-packet reception, and MAC. This model will then be used to emphasize the impact of implementing the 802.11 MAC in multi-hop networks and the gains achieved from using network coding, multi-packet reception, and their combination will be explored.

2. A new MAC approach will be developed that eliminates the non-monotonic saturation behavior resulting from the use of the 802.11 MAC. Furthermore, it will be shown that super-additive gains of approximately 6.3 times that of routing alone are achievable when using network coding, multi-packet reception, and the new MAC in combination.

3. The flexibility of the model will be demonstrated by evaluating how the effects of topology changes and asymmetric traffic across bottlenecks in the network effect overall throughput performance. In addition, extensions to the model will characterize

additional gains from combining network coding and multi-packet reception in terms of per-node throughput and delay. Primarily, the extended model will show that gains of up to $8/3$ can be achieved in terms of the time it takes to complete all unicast and broadcast sessions.

The remainder of the thesis is organized as follows. Chapter 2 will provide a detailed explanation and derivation of the model used throughout the thesis. Network element models for opportunistic network coding, multi-packet reception, and the 802.11 MAC will be developed. Furthermore, the set of canonical topology components that are fundamental in determining the benefits from implementing each of these elements will be described. Finally, an explanation of the various model parameters needed will be presented. These parameters will be used extensively in the rest of the thesis.

Chapter 3 will use the model to characterize the performance of implementing the 802.11 MAC, network coding, multi-packet reception, and their combination in multi-hop networks. A full explanation of the analysis using the model is provided and simulation results are presented. Furthermore, verification of the model will be included by comparing the results obtained from the model with those obtained by the experiments performed by Katti *et. al.*

In Chapter 4, a new MAC approach is developed that eliminates the non-monotonic saturation caused by the 802.11 MAC, and the effects on the overall network throughput will be re-evaluated. Various design trade-offs, such as the use of a pseudo carrier sense multiple access (CSMA) and differing types of traffic (i.e., unicast or multicast/broadcast) will be explored. The model used in previous chapters will then be extended in order to determine the per-node throughput, as well as the delay performance of the new MAC approach. These extensions will help in determining the performance of network coding and multi-packet reception under asymmetric traffic across the bottleneck node in the network and limited implementation of multi-packet reception at only the bottleneck node. Finally, Chapter 5 will summarize the contributions of this thesis and provide additional avenues of research that are saved for future endeavors.

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Chapter 2

Network Models and Parameters

The development of a simple, intuitive model in order to understand fully the design trade-offs of a cross-layer solution requires that various parts of the network be represented by an abstract model of their primary behaviors. Section 2.1 provides a description of the network element models and assumptions used to develop a tractable design problem. Section 2.2 gives a detailed description of the parameters used in the discussions and analysis found in Chapters 3 and 4.

2.1 Network Model

The main goal in developing a tractable model is to identify the fundamental behavior of each network element. Once the fundamental behaviors are identified, each element must be modeled using simple, intuitive methods so that various performance measures can be evaluated and design trade-offs can be weighed. The model developed within this section will focus on the inclusion of three network technologies: network coding, medium access control (MAC), and multi-packet reception (MPR). Subsequent sub-sections will identify specific behaviors of these technologies and will describe the abstractions and simplifications needed to make the model tractable.

Before describing each of the element models, a brief overview of the general scenario which is of interest is required. A wireless packet network is considered that is operated in fixed-length time-slots. Each node in the network is half-duplex (i.e., cannot receive and transmit in the same time-slot), and only one packet can be either sent or received by any given node per time-slot. If multiple packets are sent to a node in the same time-slot, it is assumed that a collision occurs and all packets are lost. Furthermore, all nodes within the network act as both sources and sinks of information. A more detailed explanation of the general scenario and model follow in the subsequent sub-sections.

2.1.1 Network Coding

Network coding, first introduced by Ahlswede *et. al.* [3], is a method in which mixing of information is performed by intermediate nodes within a network. This mixing of information enables simultaneous use of a communication link by multiple flows of information, or packets, as opposed to conventional networks where intermediate nodes simply forward a single packet one at a time to the next-hop. The basic idea in a packet network is that each packet received by a node is combined with other received packets before it is forwarded. A simple example using a piggybacking scheme proposed by Wu *et. al.* [9] is illustrated in Figure 2.1. Node 1 sends packet A to node 2 and node 2 sends packet B to node 1, but both A and B must first be sent to node 3. The relay, node 3, then forwards each packet to the appropriate destination. Without network coding, the relay must send each packet individually using two time-slots. With network coding, the relay will generate one encoded message, $A \oplus B$ (where $\oplus$ indicates mod 2 addition), and then broadcasts the encoded packet to both node 1 and 2 in a single time-slot. Since both nodes have their original packets, each can decode the message and extract packets B and A respectively.

A significant amount of research has been performed related to network coding, but some of the more significant results are related to linear network codes [4, 5, 6]. While Dougherty *et. al.* [8] provided an example showing that not all networks have a linear solution, the networks found in which capacity cannot be achieved with a linear code are very complex

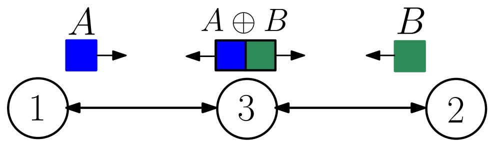

Figure 2.1: A simple example of network coding. By linearly combining packets A and B, the relay can send two messages per unit time.

and somewhat contrived. With this in mind, Katti *et. al.* [1] developed a simple, heuristic coding scheme, COPE, that operates over a finite field of size 2 and decoding of encoded packets occurs at each hop in order to avoid situations where a linear solution does not exist. COPE is a generalization of the piggybacking scheme used in the previous example that exploits the broadcast nature of the wireless medium and is the basis of the model developed in this section. The coding scheme operates using two main concepts:

1. Opportunistic Listening: The wireless broadcast medium allows nodes to overhear packets transmitted from their neighbors. COPE uses this to snoop on all communications over the wireless channel and stores each overheard packet for a period of time. Each overheard packet is then used at a later time to help in decoding any received encoded packets.

2. Opportunistic Coding: The goal is to maximize throughput by coding packets together that maximizes the number of original, or native, packets delivered in a single transmission. In order to do this, every node must ensure that each intended next-hop has enough information to decode any coded message it broadcasts.

As an example of these two concepts, consider the network shown in Figure 2.2. Nodes 1 and 2 overhear the others broadcasted transmission to node 5. Similarly, nodes 3 and 4 also overhear the others broadcasted transmission to node 5. Node 5 then codes the maximum number of packets together with the constraint that nodes 1 through 4 must be able to decode any encoded packets generated and broadcasts these packets back to nodes 1 through 4. As shown in Figure 2.2, node 5 must generate a minimum of two encoded packets, each containing a single native packet from either node 1 or 2 and a single native

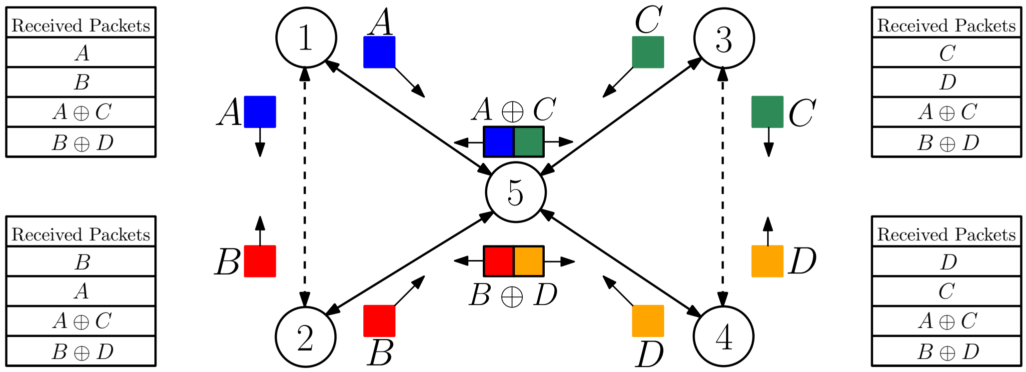

Figure 2.2: An example of opportunistic listening and the coding decisions made by COPE.

packet from either node 3 or 4, where a native packet is defined to be a non-encoded packet generated by nodes 1 through 4. This enables each node to use the stored packets from its neighbors to decode the broadcast transmissions from node 5 and retrieve the necessary native packet(s). If instead node 5 transmits the coded messages $A \oplus B$ and $C \oplus D$, nodes 1 and 2, and similarly nodes 3 and 4, will be unable to decode one of the encoded messages. As a result, COPE never codes packets headed towards the same next-hop. In addition, any packet generated by node 5 must be sent unencoded. This prevents a decoding error which would occur at each of node 5's neighbors.

Implementing COPE, Katti *et. al.* found it necessary for each node to learn each of its neighbor's states. This is done through the use of both reception reports sent in the header of each packet and the use of a probabilistic mechanism that generates an estimate of the native packets in which each node's neighbor has received. Furthermore, nodes do not wait for an optimal coding opportunity. If a node does not have more than one codable packet, it does not wait for another codable packet to arrive. Rather, it sends the packet unencoded at the first opportunity.

The model used to reflect the performance of network coding is a simplification of the above concepts. Each packet is linearly coded with another if and only if it can be decoded by the intended recipients. While COPE uses a field size of 2, Chapters 3 and 4 show that it is sometimes necessary to use a higher field size in order to achieve the maximum throughput gains. The model assumes that feedback is perfect and that each node knows the native packets overheard by its neighbors. Each packet is sent as a broadcast transmission on the channel at the first opportunity without delay and each information flow does not exercise congestion control (i.e., each packet generated is part of a UDP session). Finally, neither the complexity of the coding or decoding operations nor any other aspects of the network coding implementation found in [1] are considered since their contributions to the overall network performance is small in relation to the specific implementation aspects mentioned above.

2.1.2 Medium Access Control (MAC)

The MAC is modeled by identifying its primary behavior in the network with fidelity (i.e., fairness) and is then simplified by assuming optimal performance from the other aspects of its implementation. This gives an intuitive approach in determining the potential throughput while ensuring that the fundamental characteristics of the network are understood. Chapter 3 uses the model of the 802.11 MAC developed in this section while Chapter 4 develops the underlying behavior needed in a MAC so that throughput is optimized in the presence of multi-packet reception and network coding. As an example, this section will give a brief overview of the 802.11 MAC and then show how a simple model of its behavior can be developed.

The 802.11 MAC [22] uses two methods of medium access: the distributed coordination function (DCF), and the point coordination function (PCF). The DCF uses random access as the primary mode of operation while the PCF, which uses polling and scheduling, is optional. As such, the model of the 802.11 MAC developed here will focus solely on the DCF and assume that the PCF is not in use. The DCF uses carrier sense multiple access with collision avoidance (CSMA/CA) as the method in which a node accesses the channel. If a node has a packet to send, it will first select a random back-off interval (i.e., CA) which decrements only while the channel is idle and then monitor the channel to ensure that no other node is transmitting (i.e., CSMA). The back-off interval is uniformly distributed between the interval

0 and the value of the contention window (CW) parameter where CW is an integer starting with aCWmin (e.g., equal to 7 back-off time-slots) and exponentially increments to aCWmax (e.g., equal to 255 back-off time-slots) after each unsuccessful transmission/collision. Each node transmits its packet after the channel has been idle for the period of the random back-off interval. If the transmission is unsuccessful, the CW parameter is exponentially incremented and another back-off interval is randomly chosen between 0 and CW. This process continues until the node receives an acknowledgment that its transmission was successful and then resets CW to aCWmin. In addition to CA, an optional virtual carrier sense (CS) mechanism may be used. This mechanism uses short request-to-send (RTS) and clear-to-send (CTS) messages to reserve the channel for transmission. This mechanism is intended to prevent against hidden terminals creating collisions at the receiving node. While it is important to note that this option exists, its use is normally disabled while the 802.11 MAC is in the ad-hoc mode and was similarly disabled throughout the COPE experiments [1]. As a result, the 802.11 MAC model will assume that the RTS/CTS feature is not used.

Development of the 802.11 MAC model requires that a general idea of the behavior of the MAC over a sufficiently long period of time is identified. Zhao and Médard [2] identified that the non-monotonic behavior of the experimental COPE results shown in Figure 1.1 is a result of the 802.11 MAC, which essentially distributes channel resources equally among competing nodes. For example, if there are three nodes with information to send, the 802.11 MAC allows each node to use the channel $1/3$ of the time. This realization is also consistent with the analysis and simulation results presented by Duda [23], who showed that for a small number of stations, the probability of a node successfully accessing the channel converges to $1/N$ for N competing nodes. While this does not necessarily hold in terms of short-term fairness as the number of competing nodes increases, the network topologies used in the model, see Section 2.1.4, are sufficiently small such that this approximation is accurate.

The model developed captures the fairness aspect of the 802.11 MAC with fidelity, while the random access protocols are simplified to match the experimental throughput behaviors found in [1]. Chapter 3 will show that the 802.11 fairness mechanism is the major contributor to overall network performance. As a result, the model will assume optimal performance from all other aspects of the MAC. For example, the non-monotonic behavior in Figure 1.1 is a result of both collisions and fairness; but the total effects of collisions from either hidden nodes or identical back-off times on throughput are small in relation to the effects of the 802.11 MAC fairness mechanisms. Furthermore, the model does not consider the additional effects on overall throughput associated with various aspects of the DCF such as the potential of lost channel resources due to the random back-off. Since the DCF introduces a constant overhead that lowers the throughput to about 20% to 30% of the bit rate depending on the variant of 802.11 used [23], these assumptions provide upper bounds to the achievable throughput in the various networks that employ the MAC while still providing a rough order of magnitude for the throughput gains that can be achieved through the use of network coding and multi-packet reception.

Two additional assumptions must be made in order to simplify the model further. Since each back-off time-slot is very small in relation to the 802.11 data frame size, the model assumes that each packet is sent on or very close to the beginning of a time-slot. In other words, the contention window is included in the duration of each of the model's time-slots. This assumption allows for each transmitted packet in the following analysis and simulations to be sent on an integer time-slot. Furthermore, acknowledgments are treated in the same manner. Any required acknowledgments verifying the successful transmission of a packet are included as part of each time-slot. Based on these assumptions, the model allows for a new packet to be transmitted in each integer time-slot.

2.1.3 Multi-Packet Reception

In general, multi-packet reception allows for the correct reception of one or more packets involved in a collision. Several techniques can be used to implement multi-packet reception in a wireless network, for example: Code Division Multiple Access (CDMA), Space Division

Multiple Access (SDMA), Orthogonal Frequency Division Multiple Access (OFMDA), etc. The fundamental concept between each of these technologies is that a receiver is able to separate signals transmitted simultaneously from different nodes and then extract the required data from each transmission.

The analysis and simulations in Chapters 3 and 4 will evaluate the potential throughput gains using two models of multi-packet reception. In both cases, the number of simultaneous transmissions that a node can successfully receive without a collision is m. If a node receives strictly more than m simultaneous transmissions, only m transmissions are received while the rest are lost. In the first model, CSMA/CA is strictly enforced. If a node senses another node transmitting, it will follow the 802.11 DCF algorithm and halt its back-off timer until the channel is idle again. This, in essence, uses multi-packet reception to minimize the hidden terminal problem. In the second model, a generalization of the traditional CSMA/CA is required. A node will not halt its back-off timer unless the number of simultaneous transmissions sensed is equal to or greater than m. If a node senses fewer than m simultaneous transmissions, it will continue decrementing its back-off timer and transmit when the timer has expired. In either case, up to m packets can be sent simultaneously in the same time-slot. This will be referred to as MPR-adapted CSMA in the remainder of the thesis.

2.1.4 Network Topologies

The model will use four basic canonical topology components containing only five nodes, where each node is both a source and a sink. The development of these components is a result of two contributing factors. First, a set of components are needed which form the primary structures in larger networks that create bottlenecks and congestion. Forming each component such that every flow crosses at a single point helps model congestion and the performance gains of multi-hop traffic under both low and high loads. Second, the experiments conducted by Katti *et. al.* [1] showed that the majority of gain, as a result of network coding, originates from coding three or more native packets together. This

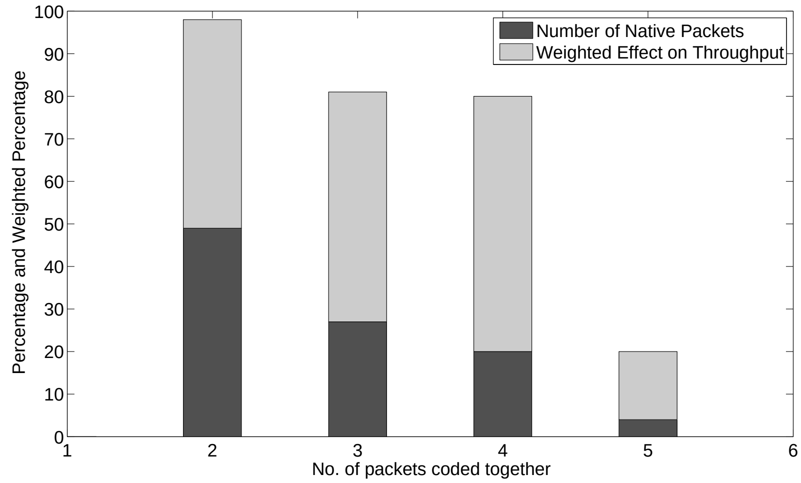

Figure 2.3: Distribution of the number of native packets encoded together at the peak point of Figure 1.1 (left), [1], with a weighted percentage reflecting the effect of coding 2 or more packets together on COPE's throughput gain.

is reflected in Figure 2.3 which shows the distribution of native packets at the maximum observed throughput in the left side of Figure 1.1 along with a weighted percentage reflecting the effect of each encoded packet on the overall throughput gain. Each component used in the model must then be of sufficient size to capture the majority of the gains seen in the experimental COPE data. As a result, the canonical topology components shown in Figure 2.4 reflect all of the possible combinations of five node multi-hop networks that allow for the potential coding of up to four native packets. Each of these components will be used in the analysis of combining network coding, multi-packet reception, and MAC in wireless networks.

Each component found in Figure 2.4 has specific constraints due to its structure and each will affect the performance of the MAC, network coding, and multi-packet reception in different ways. These constraints are defined through the use of a solid edge that depicts active, or primary communication, and a dotted edge that depicts passive, or overhear/listening communication. The absence of an edge between any two nodes indicates that all communication between the two nodes must be routed through a relay/center node.

The center node n_5 in each component is fully connected regardless of the topology, and

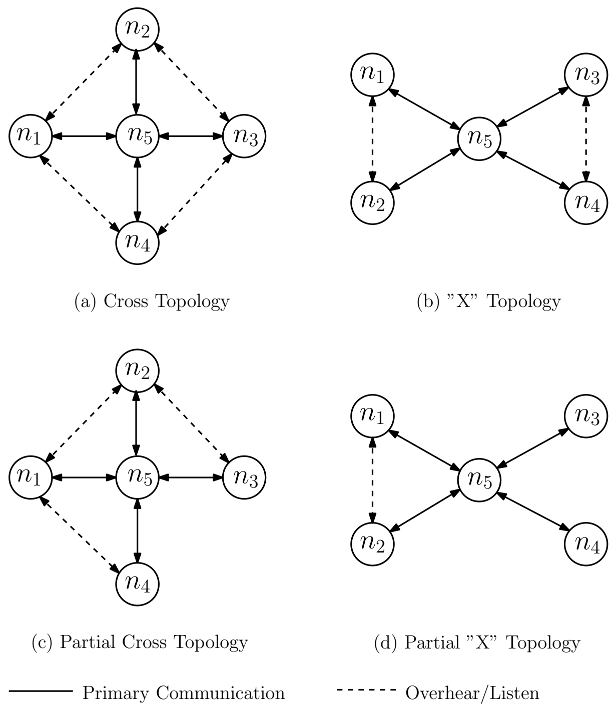

Figure 2.4: Basic network structures responsible for traffic bottlenecks and congestion in larger networks. All nodes are sources and all flows originating from n_j, $j \in [1, 4]$ cross at n_5.

traffic flows originating from the center require only a single-hop to reach their destination. The cross topology component assumes the majority of nodes are interconnected where each node can directly transmit and receive information from every other node. The only exception is that each node, n_1 through n_4, is not connected with the node on the opposite side of the center. For example, n_1 in Fig. 2.4(a) can directly send packets to n_2, n_4, and n_5 while it must send packets intended for n_3 through n_5. The "X" topology assumes that nodes n_1, n_2, and n_5 in Fig. 2.4(b) are connected and nodes n_3, n_4, and n_5 are also connected, but n_1 and n_2 are not connected to n_3 and n_4. All traffic between any node in the set $\{n_1, n_2\} \in X_1$ and a node in the set $\{n_3, n_4\} \in X_2$ must travel through the center. Finally, Figures 2.4(c) and 2.4(d) have similar constraints as the cross and "X" topology components except that a single overhear/listen edge has been removed. Specifically, each topology component imposes a constraint on the achievable throughput when network coding is used where these constraints will be discussed in greater detail in subsequent chapters.

The study of topology components extended to an arbitrary number of transmitting nodes, N, is of intrinsic interest for two reasons. These extended topologies allow for the the determination and evaluation of performance in larger networks. In addition, they provide insight into the design trade-offs of combining network coding, multi-packet reception, and MAC. For example, the analysis of both the effects of asymmetric traffic across the relay node and the determination of the delay is simplified when increasing the number of nodes in the topology component. The analysis in Chapter 4 will use the variants of the cross and "X" topology components shown in Figure 2.5 to provide insight into the achievable gains and cross-layer design of networks employing the various technologies described here. For the cross topology component, there are $N-1$ transmitting edge nodes and a single center, or relay, node. All edge nodes are connected with the center node and connected with all other edge nodes except the one directly opposite the center. Each node generates traffic destined only for the node directly opposite the center. For the "X" topology component, there are also $N-1$ transmitting edge nodes and a single center node. The edge nodes are split into two sets, X_1 and X_2. All edge nodes within a given set are fully connected and are also connected to the center. Each node generates traffic destined for a node within a different set. Furthermore, it is assumed that the set of transmitting nodes N is stable and does not frequently change. This eliminates the need to consider a decision mechanism for determining which nodes are transmitting.

2.2 Additional Parameters

Sections 2.1.1 through 2.1.4 developed the element models required to provide an intuitive, simple approach to determine the gains from network coding, multi-packet reception, MAC, and their combination. Now what is required is to take each of these element models and formulate the full model which will be used in subsequent chapters. Within the full model, network coding will be modeled by taking the linear combination of native packets over $\mathbb{F}_2$,

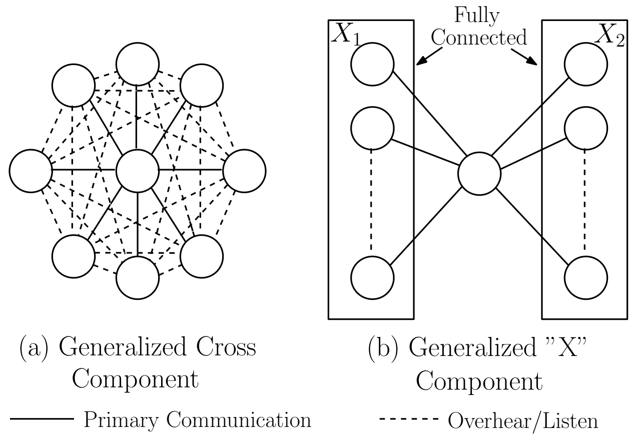

Figure 2.5: Generalized topology components for N nodes.

although extensions to larger field sizes will be required for broadcast sessions. Nodes coding packets together will only do so if the next-hop is able to decode any transmitted encoded packet and packets will not be delayed for the purposes of increasing the number of native packets in each encoded packet. If the coding node has an opportunity to transmit a packet, it will do so with whatever information it has at the time of transmission.

The MAC is modeled using the primary behavior of the protocol chosen. In the case of the 802.11 MAC studied in Chapter 3, channel resources will be distributed to each competing node equally when the channel is congested. If the channel is not congested, channel resources will be distributed to those nodes needing additional time to transmit their data. In Chapter 4.1, an alternate MAC approach will be proposed that divides channel resources differently. The element model for multi-packet reception allows for m packets to be sent from m different sources in a single time-slot. If m different sources do not have packets to send, each source will transmit in the same time-slot conditioned on the specific MAC model. For example, two nodes within range of each other will not transmit at the same time if the model chosen for the MAC uses CSMA. On the other hand, the same two nodes will be allowed to transmit at the same time if the MAC model is not using CSMA. Subsequent chapters will evaluate the benefits of using a CSMA like scheme with multi-packet reception.

The channel is divided into 100 time-slots where each time-slot uses $1/100$ of the total amount of channel resources available to the N transmitting nodes. Successful transmission of each packet requires a full time-slot therefore requiring $1/100$ of the total amount of channel resources. The network component performance is evaluated at various values of k_T ranging from 1 to 200 where k_T is deterministic and is defined as the sum of the number of packets that each node is required to send. These k_T packets are stochastically distributed to each node according to the joint PMF:

$$\begin{aligned} & P_{K_1,\ldots,K_N}(k_1,\ldots,k_N) \\ & \quad = P_{K_1}(k_1)\, P_{K_2|K_1}(k_2|k_1)\cdots P_{K_N|\underline{K}_1^{N-1}}\left(k_N|\underline{k}_1^{N-1}\right) \\ & \quad = \prod_{i=1}^{N} \binom{n_i}{k_i} p_i^{k_i}(1-p_i)^{n_i-k_i} \end{aligned} \tag{2.1}$$

where the random variable K_i is the number of packets distributed to node $i \in N$, $n_i = k_T - \sum_{j=1}^{i-1} k_i$, $p_i = (N-i+1)^{-1}$, and $\underline{k}_1^{N-1} = k_1, \ldots, k_{N-1}$.

The number of packets each node has to send will be referenced in both the analysis and simulations as the fraction of the total channel resources, or load L_i, required to send all K_i packets one hop if each packet is sent, or routed, individually. Formally, the random variable L_i is defined as:

$$L_i = \frac{K_i}{100}. \tag{2.2}$$

In addition, the total offered load to the network P is deterministic given k_T and is defined as:

$$P = \sum_{i \in N} l_i = \frac{k_T}{100} \tag{2.3}$$

where l_i is the sample value of the number of packets originating from node n_i.

In the following analysis and simulations, there are three regimes which are of particular interest. These include the unsaturated throughput regime, the maximum throughput, and the saturated throughput regime. In order to differentiate among these three regimes, it is also necessary to define the total network component load, P_T. This random variable is the actual load induced in the network component as a result of network coding and multi-packet reception. The network component is considered to be in the unsaturated throughput regime for sample values $p_T < 1$, the throughput is maximized for sample values $p_T = 1$, and is in the saturated regime for sample values $p_T > 1$. It consists of the load L_R induced by relaying packets through the center node n_{center}, and the load L_M required to send each native packet one-hop, i.e.,

$$P_T = L_R + L_M \tag{2.4}$$

where

$$\frac{1}{c} \sum_{i \in N \setminus n_{center}} L_i \leq L_R \leq \sum_{i \in N \setminus n_{center}} L_i, \tag{2.5}$$

$$\frac{1}{m} \sum_{j \in N \setminus n_{center}} L_j + L_{center} \leq L_M \leq \sum_{j \in N} L_j, \tag{2.6}$$

the coefficient c is the number of packets that can be encoded together by n_{center}, and L_{center} is fraction of time, or load, needed to send all of the packets originating at the center node one-hop. The relay load L_R is a function of the number of packets that can be encoded together by n_{center} and only counts the load required to send relayed packets a second hop. The one-hop load L_M consists of the load needed to send all of the edge node's packets to n_{center}, which is a function of m, and the load L_{center} required by n_{center} to send its own packets to the edge nodes. Both the relay load, L_R, and the one-hop load, L_M, are lower bounded by the equations (2.5) and (2.6) respectively. The lower bounds are functions of the topology component's configuration as well as the difference in each node's initial load. Each lower bound is met with equality if each $l_i = l_j$, $i, j \in N \setminus n_{center}$, and $i \neq j$. The upper bound is met with equality for L_R if no coding opportunities occur at n_{center} and for L_M if no simultaneous transmissions occur. Given sample values of each node's load l_i, $i \in N$, and the topology component, both L_R and L_M are deterministic. Chapter 3 will provide additional clarification on how to find each as well as examples using the various topology components.

Furthermore, the allocated load, S_i, is defined as the amount of channel resources given to each node in the network as a result of the MAC. When $P_T \leq 1$, each node is allocated enough time-slots to send all of its packets. The allocated load in this case is $S_i = L_i$ for $i \in N \backslash n_{center}$ and the load allocated to the center node is the sum of the load originating from the center L_{center} and the load resulting from the relaying of packets (i.e., $S_{center} = L_{center} + L_R$). As the MAC saturates (i.e., $P_T > 1$) the allocated load for each node is $S_i \leq L_i$, $i \in N \setminus n_{center}$, and $L_{center} \leq L_{center} + L_R$.

Finally, the throughput S is defined for $P_T \leq 1$ as the total number of packets in the topology component, k_T, divided by the total number of time-slots T_{slot} needed to complete either all unicast or broadcast sessions:

$$S = \frac{k_T}{T_{slot}} \tag{2.7}$$

For $P_T > 1$, the MAC limits the allocated load for each node and the throughput saturates to the amount of information that the center node can transmit per time-slot. For example, the 802.11 MAC distributes channel resources equally among each transmitting node. As a result, the center node will only receive $1/N$ of the available resources. The amount of information that the center node transmits with this allocated load is then the throughput in the saturated regime. Sections 3.1 and 3.2 will provide greater detail into calculating the throughput with and without network coding and multi-packet reception in both the cross and "X" topology components respectively.

2.3 Summary

This chapter focused on the development of a simple, intuitive model that can be used to predict the gains and to help understand the design trade-offs as well as the implementation

Term/Parameter	Definition
Native Packet	A non-encoded packet.
Encoded Packet	A packet that is a linear combination of multiple native packets.
m	Number of simultaneous packets that a node can receive in a single time-slot.
c	Number of native packets that can be linearly combined and decoded by each of the receiving nodes .
N	Total number of competing nodes in the network.
K_i	Number of packets in node n_i's queue where each K_i is distributed according to equation (2.1).
k_T	Total number of packets in every queue. This parameter is deterministically chosen in subsequent chapters.
L_i	Node n_i's individual load to the network component. $L_i = \frac{K_i}{100}$.
P	Total offered load to the network component. $P = \frac{k_T}{100}$.
L_R	Component load needed to relay packets through the center node.
L_M	Component load needed to send all packets one hop.
P_T	Total network component load.
S_i	Fraction of channel resource allocated to node n_i by the MAC.
S	Total throughput.

Table 2.1: Summary of the model parameters.

details for combining multiple communication technologies in a network. The model incorporates the primary behaviors of the various elements that contribute to network performance. Network coding was modeled using COPE as a baseline, the 802.11 MAC is modeled by focusing on the fairness aspects of the protocol, and multi-packet reception is modeled by allowing for simultaneous transmissions to occur in any given time-slot. A set of key canonical topology components that simulate multi-hop traffic and bottlenecks in larger networks were identified. Each topology component provides insight into the effects of implementing network coding, multi-packet reception, MAC, and their combination. Finally, the general framework for the evaluation of the performance of network coding, multi-packet reception, and the MAC was setup. A summary of the terms used in Chapters 3 and 4 is provided in Table 2.1.

Chapter 3

Combining Multi-Packet Reception, Network Coding, and 802.11 MAC

This chapter will provide a case study for combining network coding, multi-packet reception, and 802.11 MAC in addition to providing details concerning the achievable gains obtained from their combination. The model developed in Chapter 2 will be the primary work-horse for this evaluation and will provide valuable insight into the various cross-layer design aspects that need to be addressed in real networks. Furthermore, the overall network throughput will be used throughout this chapter to validate the model and provide a comparison of the benefits from combining the various network technologies. An analysis and simulation results for the throughput obtained with the model will presented and a brief discussion of the cross-layer design implications will be included for each 5-node topology component shown in Figure 2.4.

3.1 Cross Topology Component Analysis

The cross topology component, shown in Figure 2.4(a), is useful in the analysis of the performance gains resulting from the combination of network coding, multi-packet reception, and 802.11 MAC for several reasons. First, this component will be used to validate the model through a comparison of the results obtained from the model and the experimental results from [1]. Second, the component simulates a dense network. Since each node can overhear every other node except the node directly opposite the center, the gains from the use of opportunistic network coding will be maximized. Finally, combining network coding and multi-packet reception in this component is relatively straight-forward and will provide insight into combining these technologies in other components.

3.1.1 Routing (No Network Coding and $m = 1$)

The routing of packets without network coding or multi-packet reception provides a baseline for the achievable gains in each topology component. Since the routing throughput is the same regardless of the topology component configuration (i.e., the performance of routing is not dependent on the component configuration), each case in both this section and subsequent sections will be compared with its performance. As a brief introduction, each of the edge nodes, n_j where $j \in [1, 4]$, in Figure 2.4(a) sends all of their packets to the center node, n_5. The center node then forwards each edge node packet in addition to sending its own packets to the required destination. For example, node n_1 sends all of its packets to n_5. The center, n_5, then forwards n_1's packets to node n_3. This also occurs for each of the flows originating from n_2, n_3, and n_4; and the flow originating from n_5 is sent one hop to one, or all, of the edge nodes. Since the center node forwards each packet allowing each node to receive everything, the time and resources needed to complete each unicast session is the same as the time and resources needed to complete each broadcast session (e.g., the sinks for n_1's packets are n_2, n_3, n_4, and n_5).

When the total network component load $p_T < 1$, enough channel resources are available to route every packet from every source to their intended sinks. The fraction of channel resources allocated to each node by the MAC is $s_j = l_j$ for $j \in [1, 4]$ and $s_5 = l_5 + l_R$ where n_5's allocated load is the load originating from n_5 plus the load required to forward each relayed packet to the next hop, $l_R = \sum_{j=1}^{4} l_j$. When $p_T = 1$, the channel is fully utilized

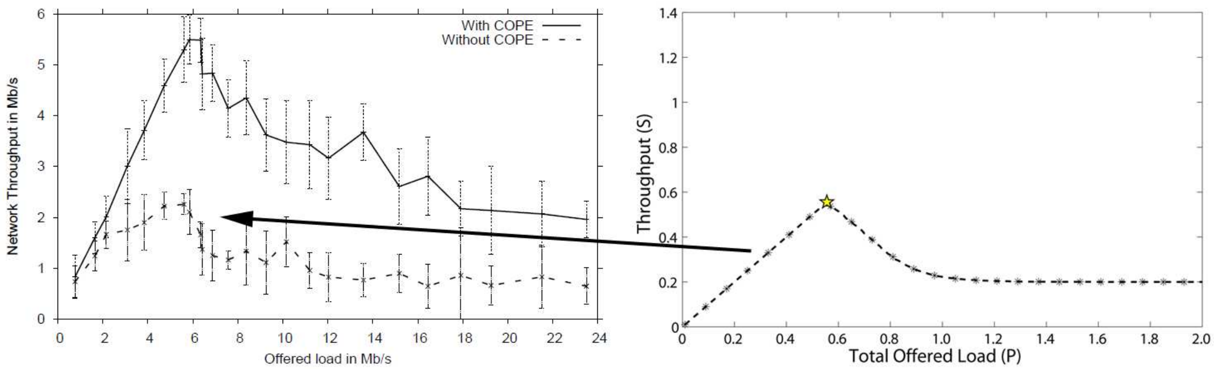

Figure 3.1: Comparison of the COPE experimental data [1] (left) with the 5-node cross topology component's routing unicast and broadcast throughput generated using the model defined in Chapter 2 (right). While the trend in the throughput behavior matches the experimental data with fidelity, the throughput found using the model, which is in packets per time-slot, is not a one-to-one comparison of the experimental throughput, which is in Mb/s. Furthermore, the experimental results contain a mixture of topology components in contrast to the simulated results which contains only one.

and the maximum throughput is reached. For symmetric loads (i.e., $l_i = l_j$ for $i, j \in [1, 5]$, $i \neq j$), this occurs when:

$$p_T = l_R + l_M = \left(\sum_{i=1}^{4} l_i\right) + \left(\sum_{i=1}^{4} l_i + l_5\right) = 1, \quad l_i = 1/9 \tag{3.1}$$

for $i \in [1, 5]$. Consistent with the analysis performed in [2], the maximum throughput when $p_T = 1$ is therefore the amount of information that the center node sends, or $S = s_5 = 5/9$. As the total offered load increases, the 802.11 MAC limits channel resources to each node and the throughput saturates. This occurs for $p_T > 1$. The 802.11 MAC initially limits the channel resources allocated to n_5 and the resulting throughput in this regime is $S = s_5 = 1 - \sum_{i=1}^{4} s_i$ which decreases as P increases. The network component completely saturates when each node requires a large fraction of the available time-slots but the MAC restricts each node's access to the channel by ensuring fairness among each competing node (i.e., $s_i = 1/5$ for $i \in [1, 5]$). For large enough P, the throughput saturates to the total amount of information that n_5 can transmit, or $S = s_5 = 1/5$.

The right side of Figure 3.1 provides a summary of the above analysis using a simulation based on the model. The maximum throughput in the figure is indicated by a star whereas the throughput in each of the regimes discussed above is the dotted curve. In addition, the experimental throughput from the COPE experiments [1] is plotted on the left side of the figure for comparison. The experimental, without COPE, data initially peaks at an offered load of approximately 5 Mb/s and saturates for higher offered loads. While the throughput obtained from the model is not a direct reflection of the experimental data due to the assumptions concerning the network topology and overhead induced by the 802.11 DCF protocol, the throughput obtained using the model does accurately reflect the trend in throughput as the offered load to the network is increased.

3.1.2 Network Coding ($m = 1$)

Opportunistic network coding is now implemented at the center node. Each node transmits each of their non-encoded, or native, packets to the center. Since each edge node is adjacent to two other edge nodes, every edge node is able to receive three degrees of freedom (two degrees of freedom through the use of opportunistic listening plus one degree of freedom from the packet originating at the given node). After each edge node has completed transmission, node n_5 transmits a single encoded packet which is sufficient for each edge node to obtain the single degree of freedom it still requires to complete both the unicast and broadcast sessions.

Similarly to the routing case, the 802.11 MAC does not limit channel resources when $p_T < 1$; and the fraction of channel resources allocated to each node by the MAC is $s_j = l_j$ for $j \in [1, 4]$ and $s_5 = l_5 + l_R$ where $l_R = {}^1\!/\!_c \sum_{j=1}^{4} l_j$. In the case of the cross topology component, $c = 4$ because each edge node only requires a single degree of freedom and the center node can provide this by encoding four native packets together. When $p_T = 1$, the channel is fully utilized and the maximum throughput is reached. For symmetric loads (i.e., $l_i = l_j$ for $i, j \in [1, 5]$, $i \neq j$), this occurs when:

$$p_T = l_R + l_M = \left({}^1\!/\!_4 \sum_{i=1}^{4} l_i \right) + \left(\sum_{i=1}^{4} l_i + l_5 \right) = 1, \quad l_i = {}^1\!/\!_6 \tag{3.2}$$

for $i \in [1, 5]$. The maximum throughput when $p_T = 1$ is again the amount of information that the center node sends, or $S = 5/6$. As the total offered load increases, the 802.11 MAC limits channel resources to each node and the throughput saturates. The throughput does not initially saturate to the same throughput as the routing case. As the MAC saturates, the center node is able to transmit a larger fraction of information than the edge nodes due to the coding operations. As P increases, the gain provided by network coding diminishes. The number of packets reaching n_5 from each edge node is limited by the MAC while packets directly introduced into the network component by n_5 are not. Essentially, the rate at which packets arriving at n_5 from the set of edge nodes is much smaller than the rate at which packets arrive from the source directly connected to n_5. As a result, the center node will send more of its own packets rather than packets obtained from the set of edge nodes. The coding gain, therefore, approaches zero as $P \to \infty$ and the throughput asymptotically approaches the saturated routing throughput.

The right side of Figure 3.2 provides a summary of the above analysis using a simulation and matches both the analysis and simulation results found in [2]. When $P \in [0, 5/9)$, network coding is seen to provide no additional gains over the use of routing alone since n_5 can forward each packet received without the MAC limiting its channel use. For $P \in [5/9, 5/6)$, network coding is instrumental in achieving the throughput shown. For $P > 5/6$, the throughput initially declines rapidly which is a result of the MAC redistributing channel resource equally among the set of nodes and then gradually decreases as the total offered load increases. It is important to note that the network coding (NC) curve shown in the right half of Figure 3.2 does not reach the maximum indicated by a star. This is due to the stochastic distribution of packets each node receives. This distribution is then averaged over in order to obtain the curve shown in the figure. This is an indication that while the maximum number of packets the center node can code together is four, there are times in which it does not have four native packets to code. A detailed discussion of this asymmetry is postponed until Chapter 4.

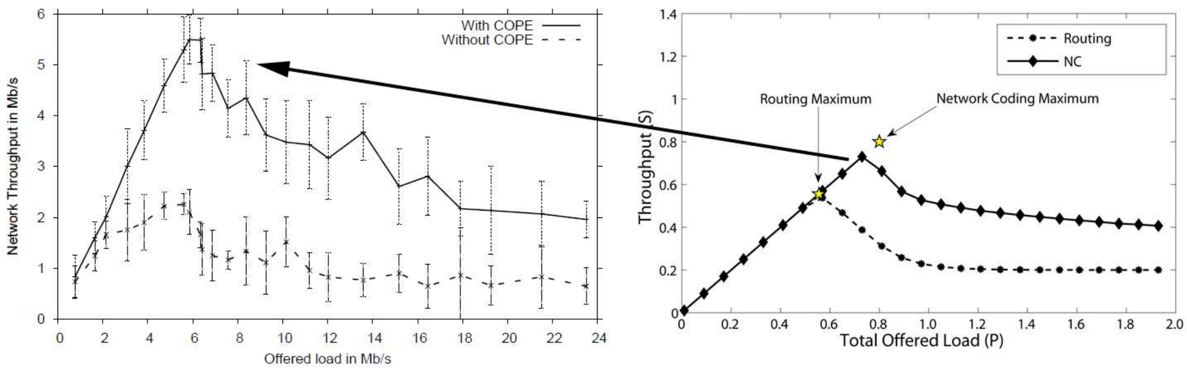

Figure 3.2: Comparison of the COPE experimental data [1] (left) with the 5-node cross topology component's network coding unicast and broadcast throughput generated using the model defined in Chapter 2 (right). While the trend in the throughput behavior matches the experimental data with fidelity, the throughput found using the model, which is in packets per time-slot, is not a one-to-one comparison of the experimental throughput, which is in Mb/s. Furthermore, the experimental results contain a mixture of topology components in contrast to the simulated results which contains only one.

The analysis and simulation results for the network coding case can also be compared to the results obtained in the COPE experiments. Similar to the routing curve, the network coding curve obtained using the developed model accurately reflects the non-monotonic saturation behavior of the experimental throughput. The experimental results show a peak in throughput at an offered load of approximately 6 Mb/s and then a saturation behavior that is similar to the saturation behavior of the model. This comparison, along with the comparison of the non-network coding case, is used as a validation of the model developed in Chapter 2 and expanded upon in later sections.

3.1.3 Multi-Packet Reception (No Network Coding and $m = \{2, 4\}$)

Multi-packet reception is similar to the routing case described earlier except a maximum of m edge nodes are allowed to transmit within a given time-slot. For $m = 2$, the total time used by all of the edge nodes to transmit their packets to n_5 is $1/2$ that needed by routing while the center node cannot transmit multiple packets simultaneously and must transmit each received packet individually. In the traditional use of the 802.11 MAC (i.e., CSMA), nodes opposite each other in this topology component are restricted to transmit at the same time although relaxing this constraint to allow the transmission of any two edge nodes at the same time does not effect either the unicast or broadcast throughput. This is because each edge node receives the necessary degrees of freedom directly from n_5 when it forwards each native packet unencoded. Multi-packet reception, in this case, can be viewed as eliminating the hidden terminal problem that is an issue for traditional 802.11 networks.

When the total network component load $p_T < 1$, enough channel resources are available to route every packet from every source to their intended sinks. The fraction of channel resources allocated to each node by the MAC is $s_j = l_j$ for $j \in [1,4]$ and $s_5 = l_5 + l_R$ where $l_R = \sum_{j=1}^{4} l_j$. When $p_T = 1$, the channel is fully utilized and the maximum throughput is reached. For symmetric loads (i.e., $l_i = l_j$ for $i, j \in [1,5]$, $i \neq j$), this occurs when:

$$p_T = l_R + l_M = \left(\sum_{i=1}^{4} l_i\right) + \left({}^{1}\!/_{m} \sum_{i=1}^{4} l_i + l_5\right) = 1, \quad l_i = {}^{1}\!/_{7} \tag{3.3}$$

for $i \in [1,5]$ and $m = 2$. This yields the maximum throughput of $S = s_5 = {}^{5}\!/_{7}$. The throughput saturates to the same throughput as routing for values of $p_T > 1$ and the gain for $m = 2$ in the saturated regime is 1 due to the suboptimal saturation behavior of the protocol.

Implementing multi-packet reception when $m = 4$ requires the relaxation of the CSMA constraint. All edge nodes, regardless of whether or not they are within transmission range of one another, are allowed to transmit at the same time. The center node, on the other hand, only transmits if none of the edge nodes are transmitting. This relaxation of CSMA to allow for simultaneous transmissions is referred to as MPR-adapted CSMA. The behavior for $m = 4$ with this relaxation is the same as that for $m = 2$ except the maximum of $S = s_5 = {}^{5}\!/_{6}$ occurs when:

$$p_T = l_R + l_M = \left(\sum_{i=1}^{4} l_i\right) + \left({}^{1}\!/_{m} \sum_{i=1}^{4} l_i + l_5\right) = 1, \quad l_i = {}^{1}\!/_{6} \tag{3.4}$$

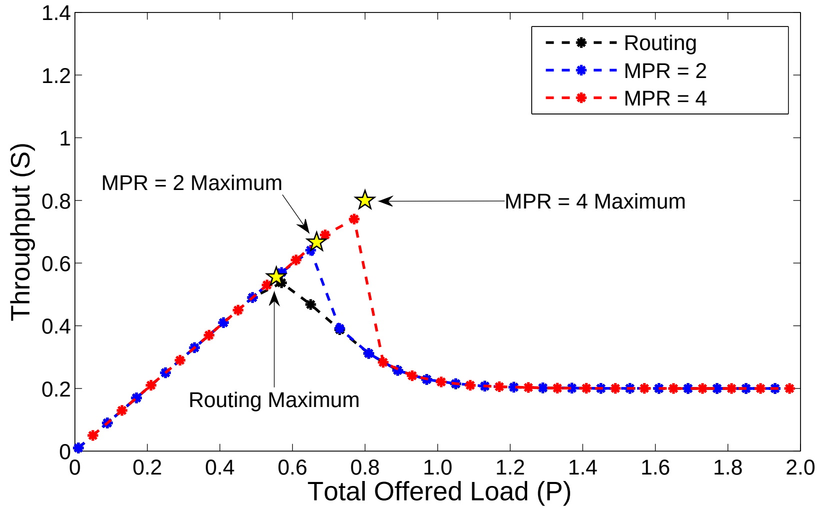

Figure 3.3: Unicast and broadcast throughput for the 5-node cross topology component with multi-packet reception.

for $i \in [1, 5]$ and $m = 4$.

All edge nodes transmit their packets to n_5 simultaneously, requiring a total of $1/6$ of the time-slots. Node n_5 then sends each node's packet individually, including its own, to the intended recipient requiring the remainder of the time-slots to finish each unicast/broadcast session. As P increases, the MAC limits each node's number of available time-slots and S saturates to $1/5$. Again, the gain in the saturated region for $m = 4$ is equal to the cases of $m = 2$ and routing. Figure 3.3 shows a summary of the throughput for both values of m. As with the network coding case, the $m = 2$ and $m = 4$ throughput curves do not reach the maxima indicated by the stars because the curves are averages over the stochastic packet distribution described in Chapter 2. The simulation results for multi-packet reception also indicate that the the gains shown in the figure are highly dependent on the asymmetry of each node's packet distribution. A more detailed discussion of the effects of asymmetry are again postponed to Chapter 4.

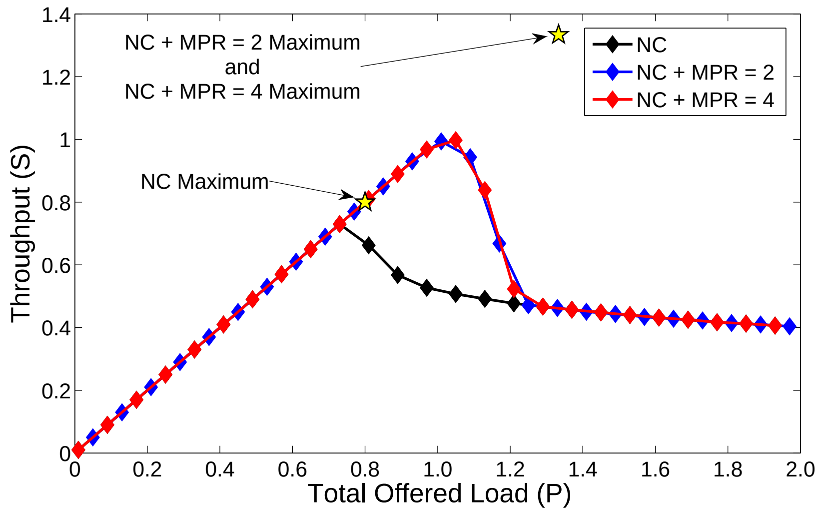

Figure 3.4: Unicast and broadcast throughput for the 5-node cross topology component with network coding and multi-packet reception.

3.1.4 Network Coding and Multi-Packet Reception ($m = \{2, 4\}$)

The case when multi-packet reception is combined with network coding results in further improvement in the throughput as seen in Fig. 3.4. Unlike the case where multi-packet reception was considered alone, the order in which each node transmits is crucial to achieving a throughput gain. This is due to the coding operations performed by n_5. Since each node relies on overhearing their neighbors to obtain the necessary degrees of freedom to decode any encoded messages sent by the center, multi-packet reception has the potential of preventing each node from overhearing its neighbor's transmissions. As a result, the use of an MPR-adaptive CSMA scheme is necessary to facilitate opportunistic listening and enable coding opportunities by n_5.

For $m = 2$, the maximum unicast and broadcast throughput of $S = {}^5\!/_4$ occurs for symmetric source loads when:

$$p_T = l_R + l_M = \left({}^1\!/_c \sum_{i=1}^{4} l_i \right) + \left({}^1\!/_m \sum_{i=1}^{4} l_i + l_5 \right) = 1, \quad l_i = {}^1\!/_4 \tag{3.5}$$

where $c = 4$ and $i \in [1, 5]$. Each set of nodes, $\{n_1, n_3\}$ and $\{n_2, n_4\}$, uses ${}^1\!/_4$ of the total number of time-slots to transmit to n_5 which then transmits a single encoded packet derived from all four node's native packets in addition to its own native packet. For $p_T > 1$, the throughput saturates to the saturated network coding throughput due to the 802.11 MAC. While the maximum achievable throughput is $^{25}/_{16}$ times the network coding without MPR throughput, the saturated gain for $m = 2$ is equal to the gain found when network coding was used alone in this region.

Relaxing the CSMA constraint results in the same throughput for unicast traffic. Broadcast traffic throughput, on the other hand, will be upper bounded by the unicast throughput and lower bounded by the $m = 2$ without network coding case. Furthermore, the broadcast throughput will be dependent on the mechanism of determining the order of transmissions, such as CSMA, round-robin, or other similar scheme, within the wireless channel.

The throughput using network coding and $m = 4$ for unicast traffic is equivalent to network coding and $m = 2$. All four edge nodes transmit to n_5 which then transmits two encoded packets in addition to its own; or the number of simultaneous transmissions are limited to two thus allowing n_5 to code everything together and send a single encoded packet to all of the edge nodes. Either strategy will achieve the same gain although the difference occurs when considering either unicast (former option) or broadcast (later option). The maximum throughput for broadcast traffic using the first method is $S = 1$, and $S = {}^{5}/_{4}$ for the second which is consistent with the maximum unicast throughput. This difference indicates that increasing m when using network coding may not be the optimal strategy.

The throughput curves in Figure 3.4 also show that the effects of unequal loads at each of the source nodes have a compounded effect on the throughput when network coding is combined with multi-packet reception. A more detailed discussion of the effects of asymmetry in the load distribution is again postponed to Chapter 4

3.2 "X" Topology Component Analysis

The cross topology component gives insight into the performance of opportunistic network coding and multi-packet reception in a dense network, and it represents the best case scenario when network coding is used since it maximizes the number of transmissions any given node receives. In order to understand the behavior of network coding and multi-packet reception in sparser networks, an analysis of the behavior of network coding and multi-packet reception is given in this section when restrictions are added on the number of each node's neighbors. This is accomplished through the use of the "X" topology component shown Fig. 2.4(b).

3.2.1 Routing and Multi-Packet Reception (No Network Coding and $m = \{1, 2, 4\}$)

Both routing and multi-packet reception with $m = \{2, 4\}$ in the "X" topology component result in the same throughput as that of the cross topology component. Since the center node transmits each packet received from the edge nodes unencoded, every node receives all of the needed degrees of freedom to complete all of the unicast sessions as well as the broadcast session. The throughput increases linearly with the total offered load until the maxima at $S = s_5 = 5/9$ for routing, $S = s_5 = 5/7$ for $m = 2$, and $S = s_5 = 5/6$ for $m = 4$ are reached. For $p_T > 1$, the throughput saturates to $S = 1/5$ for each case as a result of the fairness imposed by the 802.11 MAC. Figure 3.5 provides the simulation results for each of the cases. As with the cross topology component, the simulation results for multi-packet reception also indicate that the the gains are highly dependent on the asymmetry of each node's packet distribution; and a more detailed discussion of the effects of asymmetry are likewise postponed to Chapter 4.

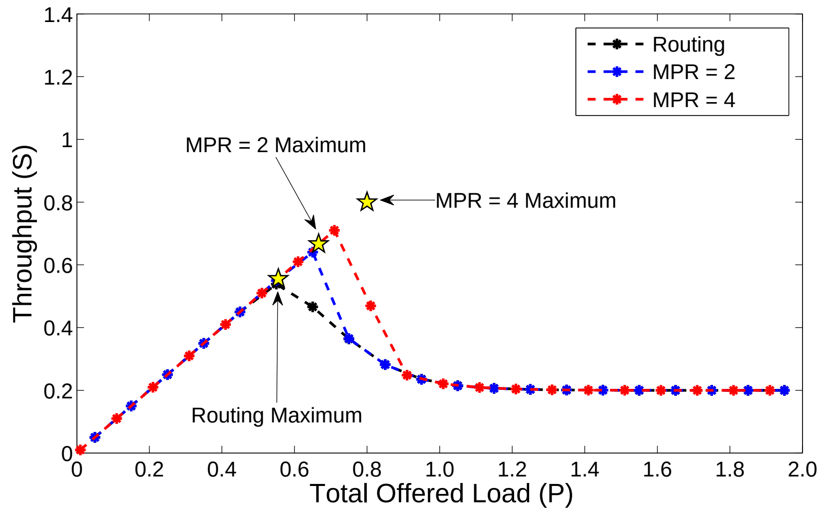

Figure 3.5: Unicast and broadcast throughput for the 5-node "X" topology component for routing and multi-packet reception.

3.2.2 Network Coding ($m = 1$)

Limiting the ability to overhear other edge nodes in the component results in the reduction of the number of possible packets that can be encoded together. Packets from different nodes within the same set, i.e., $\{n_1, n_2\} \in X_1$ and $\{n_3, n_4\} \in X_2$, cannot be encoded together because all flows transitioning between X_1 and X_2 are effectively headed towards the same next-hop. This forces n_5 to code only a subset of packets together which increases the number of transmissions the center node must make. For example, the center node must make a minimum of two transmissions for every four packets it receives from different edge nodes in order to ensure that each node has the necessary degrees of freedom to decode all of the packets.

Like the cross component's throughput, the throughput of the "X" topology component increases linearly until it reaches its maximum at $S = {}^5\!/_7$. Assuming symmetric source loads, this maximum occurs when:

$$p_T = l_R + l_M = \left({}^1\!/_c \sum_{i=1}^{4} l_i \right) + \left(\sum_{i=1}^{4} l_i + l_5 \right) = 1, \quad l_i = {}^1\!/_7 \tag{3.6}$$

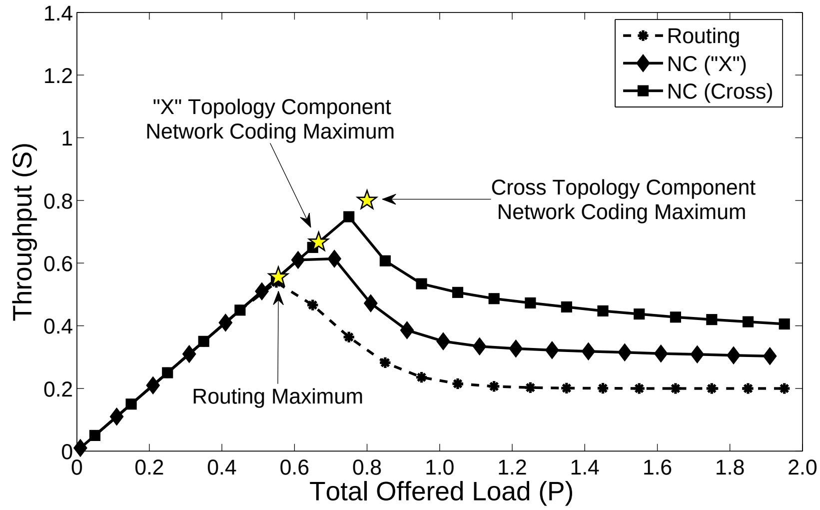

Figure 3.6: Comparison of the unicast and broadcast throughput for the 5-node Cross and "X" topology components with network coding.

for $i \in [1, 5]$ and $c = 2$. By comparing the results obtained from both the cross and "X" topology components, it is evident that the performance of opportunistic network coding is highly dependent on the network structure. The elimination of a single neighbor at each edge node results in approximately a 14% reduction in maximum throughput. This is a clear indication that as the network becomes sparser, the gain from network coding is diminished. The throughput saturates for $p_T > 1$ and the non-monotonic behavior in the saturated throughput regime is again due to the fairness aspect of the 802.11 MAC. Like the cross topology component, the gain provided by network coding approaches zero as both P and l_5 increase.

Figure 3.6 provides a comparison of the simulation results for implementation of network coding in both the cross and "X" topology components. Both the maximum and saturated throughput for the "X" topology component are significantly smaller than the throughput obtained using the cross topology component, and provides a clear indication that the use of opportunistic network coding is more beneficial in denser networks.

3.2.3 Network Coding and Multi-Packet Reception ($m = \{2, 4\}$)

The throughput for the case in which network coding is used in conjunction with multi-packet reception ($m = 2$) for the "X" topology component is similar to the cross topology component throughput. Continuing to restrict nodes within opposite sets to transmit at the same time, the throughput increases linearly until it reaches its maximum at $S = 1$ when:

$$p_T = l_R + l_M = \left({}^{1}\!/\!{}_{c} \sum_{i=1}^{4} l_i \right) + \left({}^{1}\!/\!{}_{m} \sum_{i=1}^{4} l_i + l_5 \right) = 1, \quad l_i = {}^{1}\!/\!{}_{5} \tag{3.7}$$

for $c = 2$ and $i \in [1, 5]$. The throughput for this case saturates to the network coding throughput found in the previous subsection for $p_T > 1$. The average and maximum throughput shown in Fig. 3.7 for $m = 2$ is achieved for both unicast and broadcast traffic when using CSMA to force nodes from different sets to transmit to n_5 at the same time. Removing this constraint results in the same throughput for unicast traffic. Broadcast traffic throughput will be upper bounded by the unicast throughput and lower bounded by the $m = 2$ without network coding case. Furthermore, the broadcast throughput will be dependent on the mechanism of determining the order of transmissions, such as CSMA, round-robin, or other similar scheme, within the wireless channel.

For $m = 4$, the maximum unicast throughput of $S = {}^{5}\!/\!{}_{4}$ is achieved when:

$$p_T = l_R + l_M = \left({}^{1}\!/\!{}_{c} \sum_{i=1}^{4} l_i \right) + \left({}^{1}\!/\!{}_{m} \sum_{i=1}^{4} l_i + l_5 \right) = 1, \quad l_i = {}^{1}\!/\!{}_{4} \tag{3.8}$$

for $c = 2$ and $i \in [1, 5]$. This maximum is obtained by allowing all four edge nodes to transmit to the center at the same time. The center node codes a maximum of two native packets together from different edge node sets and transmits two encoded packets back to the edge nodes, including its own native packets, in order to complete all unicast sessions. At the completion of all unicast sessions, each node still requires a maximum of one additional degree of freedom to complete the broadcast session. Allowing n_5 to code all of the native

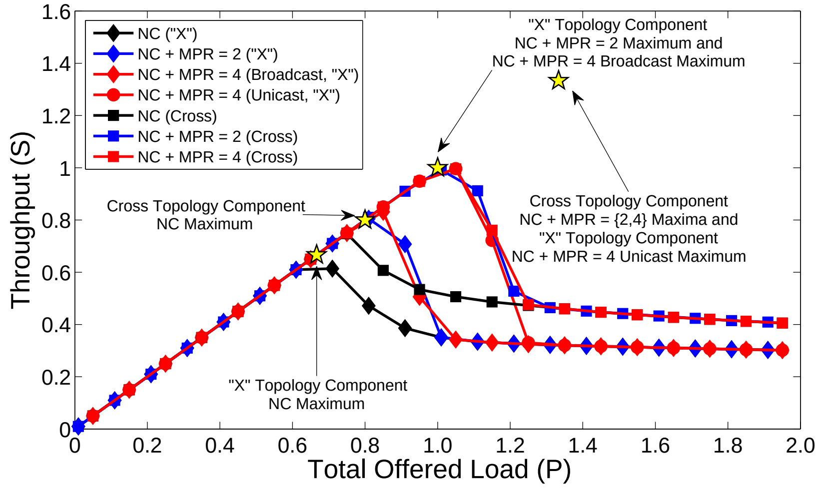

Figure 3.7: Comparison of the unicast and broadcast throughput for the 5-node cross and "X" topology components with network coding and multi-packet reception.

edge node packets together and send one additional encoded transmission enables each node to extract the required degree of freedom and obtain the full set of transmitted messages. The maximum throughput for this case is therefore the same as the case for network coding with $m = 2$ and is equal to $S = 1$.

Figure 3.7 provides a comparison of the throughput results found in this section with the throughput results found in the cross topology component case. In general, there is approximately a 20% decrease in network performance between the cross and "X" topology components. The only case unaffected by the modification of topologies is the network coding with $m = 4$ and unicast flows, which provides additional insight into the benefits of combining multi-packet reception and network coding. Similarly to the cross topology component, the average throughput for both cases discussed in this section do not reach the maxima found because of the stochastic load distribution shown in equation (2.1), which results in asymmetric traffic flows across the center node. If each node had an equal amount of information to send, then the maxima found in this section would be achieved.

3.3 Partial Topology Component Analysis

The use of the partial topology components found in Figure 2.4(c) and (d) provide additional information concerning the behavior of the gains obtained from network coding and multi-packet reception as the network topology changes. As the following sections will show, the large gains from network coding seen in dense networks are very fragile to topology changes while the gains in sparser networks are much more robust.

3.3.1 Partial Cross Topology Component

The partial cross topology component is shown in Fig. 2.4(c). Similar to the fully connected cross topology component, each node cannot overhear the node on the opposite side of the relay; but an additional constraint is imposed by removing one neighbor from two of the edge nodes. In general, the gains provided by network coding and multi-packet reception in this component are no better than the gains found using the "X" topology component. The analysis for routing and multi-packet reception alone are similar to those discussed in Section 3.1 and will not be readdressed here. Instead, only the cases involving network coding will be discussed.

Limiting the ability of just two nodes to overhear each other results in a dramatic decrease in the gain provided by opportunistic network coding. With network coding alone, transmission from each edge node to n_5 still takes a total of four time-slots; but the number of transmissions that n_5 must make is increased as a result of the deleted edge. Once each edge node has completed its transmission to n_5, the two edge nodes on either side of the deleted edge will require two additional degrees of freedom in order to decode any encoded messages sent from n_5 while the other edge nodes will require only one additional degree of freedom. The minimum number of transmissions from n_5 to provide the necessary degrees of freedom to each node is two and the resulting throughput is the same as the throughput found for the "X" topology component.

Adding multi-packet reception to the network coding case also results in a total network throughput that is equal to that found for the "X" topology component. When the order of transmission from each edge node to n_5 is enforced (i.e., CSMA), the maximum throughput for the case when $m = 2$ is $S = 1$ for both unicast and broadcast traffic. The throughput for $m = 4$ reaches it's maximum of $S = 5/4$ for unicast traffic and $S = 1$ for broadcast traffic.

It is surprising that the performance of the partial cross topology component is no better than the performance of the "X" topology component and that restricting the ability of a single pair of nodes to overhear each others transmissions results in a drastic decrease in network performance. This drastic reduction in performance shows that the network coding gains seen in the cross topology component are hard to achieve in a general network and rely heavily on the ability of every node to be able to overhear the majority, if not all, of the other nodes in the network.

3.3.2 Partial "X" Topology Component

Each node in the partial "X" topology component shown in Fig. 2.4(d) is connected to n_5 and only one pair of edge nodes can overhear each other. Similar to the cross, partial cross, and "X" topology components, the analysis and throughput for the cases involving routing and multi-packet reception only are the same for the partial "X" topology component. The added restriction from reducing the amount of overhearing performed by each edge node results in further decreases in network performance when using network coding.

In the partial "X" topology component, the throughput for unicast and broadcast traffic differ for each case that network coding is used. The center node can effectively code a maximum of two native packets together resulting in a maximum unicast throughput of $S = 5/7$ and a maximum broadcast throughput of $S = 5/8$. With broadcast traffic, n_5 sends two encoded messages, each containing two different sets of native packets, plus its own to the edge nodes. It must also send one additional coded message containing all four native edge node packets. This ensures each node has all of the degrees of freedom required.

	"X" Throughput	Partial "X" Throughput	Percent Difference
Routing	$5/9$	$5/9$	0%
Network Coding (Unicast)	$5/7$	$5/7$	0%
Network Coding (Broadcast)	$5/7$	$5/8$	12.5%
Multi-Packet Reception ($m = 2$)	$5/7$	$5/7$	0%
Multi-Packet Reception ($m = 4$)	$5/6$	$5/6$	0%
Network Coding with $m = 2$ (Unicast)	1	1	0%
Network Coding with $m = 2$ (Broadcast)	1	$5/6$	16.7%
Network Coding with $m = 4$ (Unicast)	$5/4$	$5/4$	0%
Network Coding with $m = 4$ (Broadcast)	$5/4$	$5/4$	0%

Table 3.1: Comparison of the maximum throughput obtained using the "X" and partial "X" topology components.

Combination of network coding and multi-packet reception shows that the achievable throughput remains dependent on the network topology for specific cases. For instance, the maximum throughput for the case where network coding and $m = 2$ is $S = 1$ for unicast traffic and $S = 5/6$ for broadcast traffic. This is a slight reduction in throughput from the other network topology components. On the other hand when $m = 4$, the maximum is the same as that found for the partial cross and "X" topology components. Since multi-packet reception restricts each node's ability to overhear other nodes' transmissions, the limitations imposed by the network topology do not impact the gains provided by the combined use of multi-packet reception ($m = 4$) and network coding.

Unlike the difference in performance between the cross and partial cross topology components, the difference in performance between the "X" and partial "X" topology components is relatively small. Table 3.1 provides a comparison of the performance of combining these technologies in both the "X" and partial "X" topology components. As indicated by the table, only two specific cases are effected by the additional constraint imposed by the partial "X" component. While the partial cross component showed that the gains seen in a dense network are somewhat fragile to topology changes, the partial "X" component shows that as the network becomes less dense, the gains obtained from network coding and multi-packet reception are more robust.

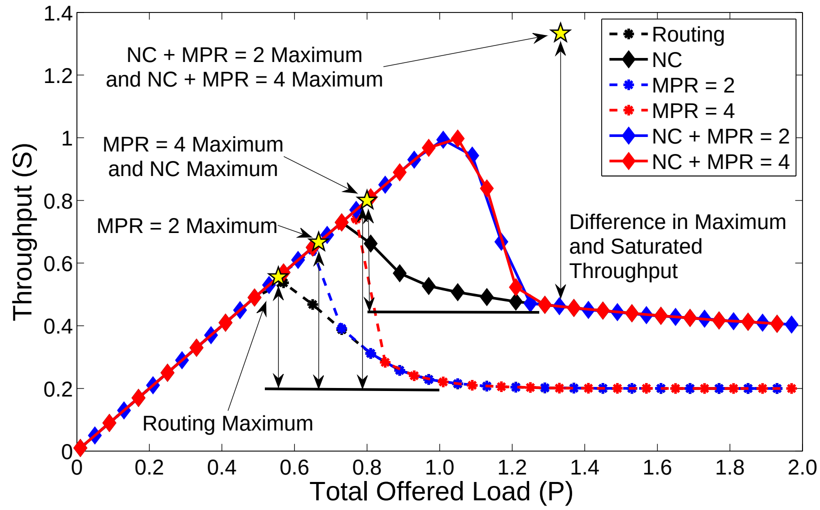

Figure 3.8: Maximum and average unicast and broadcast throughput for the 5-node cross topology component. Each vertical double arrow shows the difference in the maximum and saturated throughput due to MAC fairness.

3.4 Summary

The use of each component in the model allows for the fundamental behavior of combining network coding and multi-packet reception in a larger network to be determined. Figures 3.8 and 3.9 provide a summary of the analysis provided in this chapter by plotting the throughput as a function of the total offered load for both the cross and "X" topology components. It is clear from these figures that the fundamental behavior of the 802.11 MAC when the channel is congested leads to a sub-optimal saturation behavior. Chapter 4 will address this through the design of an alternate MAC that provides monotonic saturation behavior.

One of the major results that may not have been apparent from earlier sections is that the model identified that the combination of network coding and multi-packet reception produces *super-additive* gains. For example, the maximum throughput obtained from network coding and $m = 2$ in the cross topology component is $5/4$. This is a gain of 2.5 over the maximum routing throughput of $S = 1/2$. On the other hand, the gain from the sum of the maximum throughput gains obtained using network coding alone and multi-packet reception ($m = 2$) alone is 2.1. Figure 3.10 provides a comparison of the maximum throughput obtained using

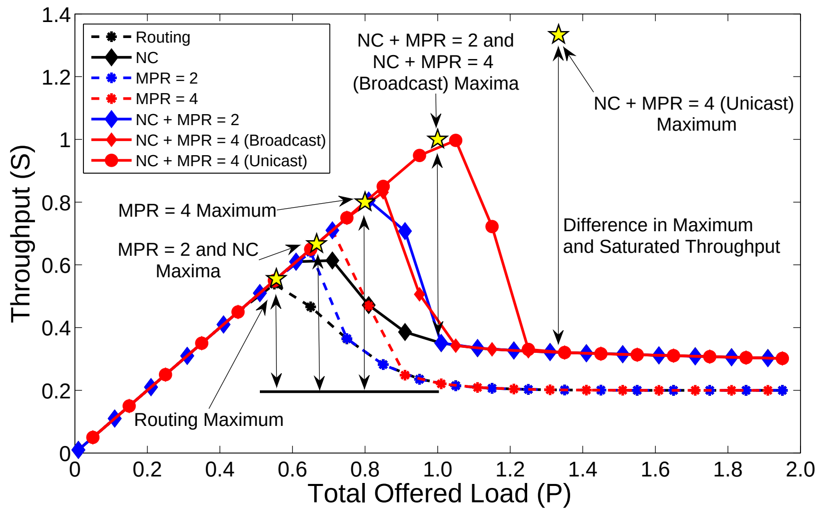

Figure 3.9: Maximum and average unicast and broadcast throughput for the 5-node "X" topology component. Each vertical double arrow shows the difference in the maximum and saturated throughput due to MAC fairness.

the "X" topology component and provides additional clarification on the nature of the super-additive gains.

The use of the "X", partial cross, and partial "X" topology components further show behaviors that were otherwise masked in the cross component. Each component provides insight into the behavior of network coding and multi-packet reception in sparser networks, methods in which to implement variants of opportunistic network coding for broadcast traffic, and highlighted the robustness of the gains to topology changes.

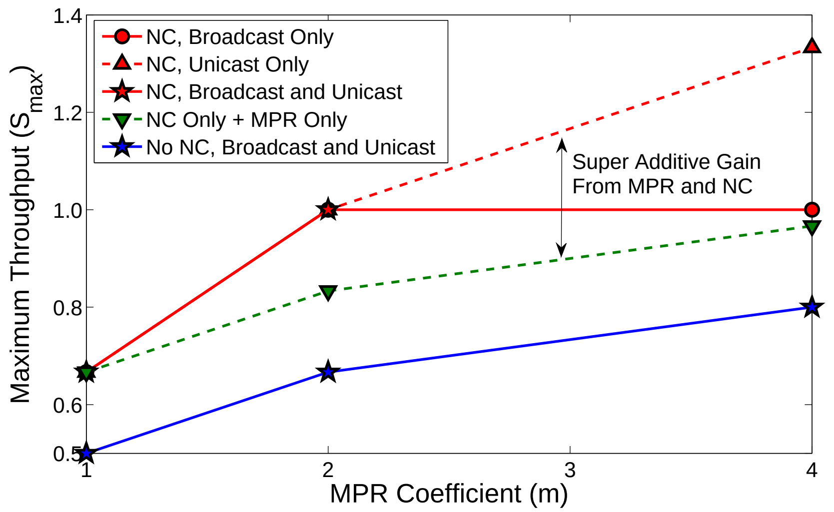

Figure 3.10: Maximum throughput of the 5-node "X" topology component as function of the MPR capability. Super-additive gains are achieved when using network coding in conjunction with multi-packet reception.

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Chapter 4

Cross-Layer Design with Network Coding, Multi-Packet Reception, and MAC

Chapter 3 showed how the model developed in Chapter 2 could be used to determine the potential gains from combining multiple communication technologies. This chapter will further the work shown in previous chapters by first using the model to design a cross-layer approach to eliminate the non-monotonic saturation behavior resulting from the the fairness aspects of the 802.11 MAC and then extend the model to determine the behavior of the network under different conditions. The performance of the new cross-layer approach will first be evaluated using the methods proposed in Chapter 3, then evaluated under conditions involving asymmetric traffic across the bottleneck node, and finally evaluated when multi-packet reception is implemented at only a subset of nodes in the network.

4.1 Improving the MAC Fairness Protocol

The decision to use a specific MAC can have serious implications on the overall network performance. This was the case in Chapter 3 when the 802.11 MAC was used in combination with network coding and multi-packet reception. Under high offered loads, the 802.11 MAC was the primary reason for the non-monotonic throughput saturation shown throughout the last chapter. This leads to the obvious question: how would you design a MAC given that you are using network coding and multi-packet reception? In addition, can the fairness imposed by the MAC be tailored to specific applications? This section will focus on answering these questions in an arbitrary sense with and without network coding as well as with and without multi-packet reception.

While several approaches to improve fairness among flows in 802.11 networks have been suggested, none have considered the combined use of multi-packet reception and network coding. As a result, the approach presented in this section will optimize the throughput subject to multi-packet reception, network coding, the topology configuration, and provides fairness to *flows* rather than to *nodes*. The basic idea behind the improved protocol approach is to allocate resources proportional to the number of different flows passing through a given node when the network saturates. While allocating more resources to flows originating at the center and less resources to flows originated at edge nodes would yield even higher throughput, the developed policy ensures that each flow of information is given the same priority. The center, or relay, node will be allocated more resources than each edge node in order to relay information; but it must also limit the amount of self-generated traffic so that it equals the average per-node non-self-generated traffic being relayed.

This approach can be implemented multiple ways within a network. For instance, the basic 802.11 MAC architecture can be used with slight modifications to allow the contention window range to be dynamically determined based on the flows crossing through a given node in addition to some queuing policies that provide fairness to each flow. Another approach, suggested by Zhao and Médard [2], is to prioritize encoded packets in the queue over unencoded packets. Regardless of the specific implementation details, the model will allow for the determination of the specific behavior that is needed to achieve the goals suggested above.

The new MAC will be designed using the generalized network topology components depicted in Figure 2.5 where each component consists of N transmitting nodes. The extended cross topology component used assumes that each node, except the center, which is connected to every other node, is connected to $N-2$ other edge nodes. All unicast flows originating from any given node terminates at the node in which it is not connected. Nodes in the extended "X" topology component are split into two sets. Each set of nodes is fully connected and every node is also connected to the center node. All unicast flows generated in the network traverse sets forcing each flow to converge at the center.

The goal of the MAC is to allocate the number of time-slots each node receives so that the throughput S is maximized subject to providing fairness to *flows* rather than *nodes* and the flow constraint:

$$\sum_{j=1}^{N-1} {}^{s_j}/_{m} + s_R \leq 1 \tag{4.1}$$

where the center node is considered to be the Nth node, s_j is the fraction of time-slots allocated to each edge node, and s_R is the fraction of time-slots allocated to the center node. The allocation of time-slots is divided into several cases based on the topology component configuration, the type of traffic (unicast or broadcast), and the specific network technologies implemented. Similar to the previous chapter, the throughput $S = s_R$ when network coding is not used and S is a function of the number of packets that can be effectively coded together, which is dependent on the multi-packet reception coefficient m, the use of CSMA, topology, and the traffic type (unicast or broadcast), when network coding is used.

4.1.1 MAC Design for the Cross Topology Component with Unicast or Broadcast Traffic

Before beginning, this section and the next assumes that a single packet is distributed to every node in order to simplify the notation and explanation. Extending the results presented in both sections to allow for arbitrary packet distributions is a fairly straight forward exercise and is not explicitly discussed below. Without network coding, the center node requires a number of time-slots equal to the number of transmitting source nodes N. With network coding, throughput is maximized by ensuring the center node codes the maximum number of native packets together.

As seen in the previous chapter, implementation of multi-packet reception can potentially prevent each node from immediately decoding any encoded message sent by the center. This is especially true if a CSMA like scheme (i.e., a node will not transmit if it knows another node is already transmitting) is not used, since then nodes with the ability to overhear each other may transmit at the same time. For example when $m = 2$ and CSMA is not used, the center node needs to send two encoded packets. Each packet should contain linear combinations of all of the edge node packets. In addition, each packet should be combined in a different manner so that each encoded packet is linearly independent from the other encoded packet. This ensures that each edge node has the necessary degrees of freedom to decode each packet since there is large probability that at least one node transmitted at the same time as one of its neighbors. As opposed to the last chapter where network coding operations were performed over a field size of 2, this implementation requires field sizes slightly larger than two to ensure linear independence between encoded packets.

Generalizing for N and m, as well as considering only integer numbers of time-slots, the proportion of channel resources distributed to each node is:

$$s_j = \begin{cases} \frac{1}{\lceil (N-1)/m \rceil + N} & \text{without NC} \\ \frac{1}{\lceil (N-1)/m \rceil + m_c + 1} & \text{with NC} \end{cases} \tag{4.2}$$

and

$$s_R \leq \begin{cases} \frac{N}{\lceil (N-1)/m \rceil + N} & \text{without NC} \\ \frac{m_c+1}{\lceil (N-1)/m \rceil + m_c + 1} & \text{with NC} \end{cases} \tag{4.3}$$

where m_c is defined for the various cases outlined in Table 4.1. The implementation of

	MPR-Adapted CSMA Used	MPR-Adapted CSMA Not Used
$m = 1$	$m_c = 1$	$m_c = 1$
$m = 2$	$m_c = 1$	$m_c = 2$
$m = 4$	$m_c = 3$	$m_c = 3$

Table 4.1: Definition of m_C in equations (4.2) and (4.3) for $m = \{1, 2, 4\}$, as well as with or without the use of an MPR-adapted CSMA scheme.

CSMA with multi-packet reception requires a relaxation of the constraint that only a single node can transmit at a time. With multi-packet reception, up to $m - 1$ connected nodes are allowed to transmit at the same time in the cross topology component. This relaxation is to referred to as MPR-adapted CSMA in the table and throughout this chapter.

Equation (4.3) is met with equality if MPR-adapted CSMA is used for $m = 2$ as well as for all cases when $m = \{1, 4\}$. Equation (4.3) may be met with inequality when MPR-adapted CSMA is not used for $m = 2$. This is due to a non-zero probability that any given node may miss a packet from a node in which it can overhear due to the half-duplex constraint.

It is important to note the difference between the use of an MPR-adapted CSMA scheme and the non-use of CSMA. While using a scheme such as MPR-adapted CSMA results in a significant throughput gain for small N, its use becomes insignificant as N grows. In addition, equations (4.2) and (4.3) apply to both unicast and broadcast flows. Since each node requires at most two degrees of freedom per round of transmission from every node, treating every session as if it were a broadcast session is the most efficient way of completing every unicast session.

4.1.2 MAC Design for the “X” Topology Component with Unicast and Broadcast Traffic

The optimal allocation of resources to maximize the throughput subject to the equal flow constraint and equation (4.1) varies based on the network topology. The previous subsection provided an approach for a dense network while this subsection provides an approach for a sparser network. For the "X" topology component, the fraction of time-slots allocated to each node varies based on the type of traffic being considered. For instance, the fraction of time-slots s^U allocated to each node for unicast traffic when using MPR-adapted CSMA to only allow $\lceil m/2 \rceil$ nodes within the same set the capability to transmit simultaneously is:

$$s_j = s_j^U = \begin{cases} \frac{1}{\lceil (N-1)/m \rceil + N} & \text{without NC} \\ \frac{1}{\lceil (N-1)/m \rceil + \max(|X_1|,|X_2|)+1} & \text{with NC} \end{cases} \tag{4.4}$$

and

$$s_R = s_R^U = \begin{cases} \frac{N}{\lceil (N-1)/m \rceil + N} & \text{without NC} \\ \frac{\max(|X_1|,|X_2|)+1}{\lceil (N-1)/m \rceil + \max(|X_1|,|X_2|)+1} & \text{with NC.} \end{cases} \tag{4.5}$$

When considering broadcast traffic, additional degrees of freedom must be sent by the center to complete each session. Without network coding, equations (4.4) and (4.5) still hold. With network coding, there is a possibility that each destination node will require a maximum of one additional degree of freedom per node for $m = 2$ or three degrees of freedom per node for $m = 4$ when either $\mid X_1 \mid \geq m$ or $\mid X_2 \mid \geq m$ and the order of node transmission is not enforced (i.e., MPR-adapted CSMA is not used). Providing these additional degrees of freedom can be accomplished by the center node sending at most three additional coded packets, where each coded packet contains a different linear combination of all of the native edge node packets. As with the cross topology component, a field size larger than 2 is necessary to ensure that each encoded packet is linearly independent. The fraction of time-slots each node receives for broadcast traffic, s^B, with network coding is then bounded by:

$$s_j^U \geq s_j^B \geq \frac{1}{\lceil (N-1)/m \rceil + \max(|X_1|,|X_2|)+m} \tag{4.6}$$

and

$$s_R^U \leq s_R^B \leq \frac{\max(|X_1|,|X_2|)+m}{\lceil (N-1)/m \rceil + \max(|X_1|,|X_2|)+m}, \tag{4.7}$$

where each edge node receives the fraction of time-slots corresponding to s_j^B and the center node receives the fraction of time-slots corresponding the s_R^B.

The fraction of time-slots allocated to each edge node for broadcast traffic increases when the difference between the number of simultaneous transmissions in each set, X_1 and X_2, is zero. In contrast, the fraction of time-slots allocated to each edge node decreases when there are more simultaneous transmissions in one set than in the other. For example, if two nodes in X_1 and two nodes in X_2 transmit at the same time when $m = 4$, then the center node must send three coded packets to ensure that each edge node has the necessary degrees of freedom to decode each packet. On the other hand, if four nodes transmit simultaneously in X_1 and zero in X_2, then the center must send four packets (either unencoded or encoded) to ensure each node has the necessary degrees of freedom to complete the broadcast session. Since the center must use more channel resources to relay the necessary degrees of freedom to each node, the fraction of time-slots allocated to each edge node must be decreased.

4.1.3 Throughput Performance of the Improved MAC

Now that the behavior of the MAC needed to provide both fairness to flows as well as non-monotonic saturation when the channel becomes congested has been determined, the techniques used in Chapter 3 can be reapplied to verify the throughput performance for the combination of network coding, multi-packet reception, and improved MAC. Using the model developed in Chapter 2, the throughput S can be calculated for each of the regimes $p_T < 1$, $p_T = 1$, and $p_T > 1$. Since the analysis is the same as that found in the previous chapter, it will not be addressed here again. Instead, a summary of the performance of the improved MAC is shown in Figures 4.1 and 4.2 for the cross and "X" topology components respectively. These figures verify that the throughput saturates at the *maxima* found in

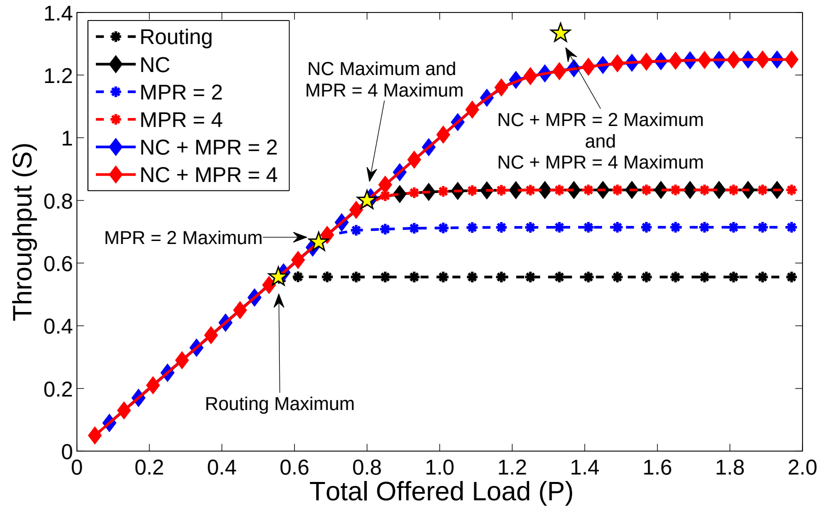

Figure 4.1: Maximum and average broadcast and unicast throughput for the 5-node cross topology component with the improved MAC.

Chapter 3 for each topology component. As the network saturates, the improved fairness protocol limits each node's access to the channel. When each node's load is greater than the limit imposed by the protocol, the total throughput saturates at the maxima indicated by the stars.

The simulation results in Figures 4.1 and 4.2 show that the maxima may not be reached. This is due to asymmetry resulting from the stochastic distribution of packets. As the network initially saturates, some nodes will have higher loads than others. This results in a lower throughput than when the load distribution across all nodes is symmetric. As P increases, the average throughput for each case asymptotically approaches the maxima.

The gains associated with the modification of the fairness protocol are listed in Table 4.2. While these gains may not be achievable, in their entirety, due to the assumptions made by the model, they do provide an indication of the potential benefits of incorporating network coding, multi-packet reception, and a MAC that provides monotonic saturation. First, the table highlights the benefit of using the new MAC in a multi-hop network by showing that the saturated throughput from using the new MAC to only route packets through the network is 2.8 times the saturated throughput when the 802.11 MAC was used. Second, the table

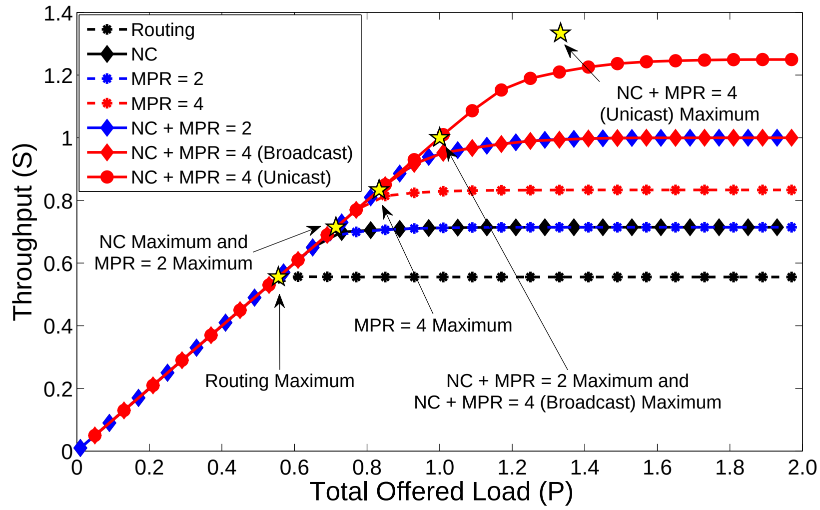

Figure 4.2: Maximum and average broadcast and unicast throughput for the 5-node "X" topology component with the improved MAC.

further illustrates that the gains from combining network coding and multi-packet reception are super-additive. For example, the sum of the individual gains from network coding and multi-packet reception with $m = 2$ in the "X" topology component is only 73% of the combined gain from use of NC + MPR ($m = 2$) together when the MAC gain is removed (e.g., ${}^{2(3.6-2.8)}/_{(5.0-2.8)} = 0.73$). This super-additive behavior manifests itself in most of the cases shown in the table. The only instances where the gains are not super-additive is in the "X", partial cross, and partial "X" topology components for broadcast traffic and NC + MPR = 4. In each case, the sum of the individual gains from network coding and multi-packet reception are equal to the gains obtained from the combined use of network coding and multi-packet reception. This leads to the conjecture that the gain from the combination of network coding and multi-packet reception are lower bounded by the sum of the individual gains (i.e., $G_{NC+MPR} \geq G_{NC} + G_{MPR}$) in an arbitrary network.

Case	(a) Cross Topology		(b) "X" Topology		(c) Partial Cross Topology		(d) Partial "X" Topology	
	Unicast	Broadcast	Unicast	Broadcast	Unicast	Broadcast	Unicast	Broadcast
Routing	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
Network Coding (NC)	4.2	4.2	3.6	3.6	3.6	3.6	3.6	3.1
MPR ($m = 2$)	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
MPR ($m = 4$)	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2
NC + MPR ($m = 2$)	6.3	6.3	5.0	5.0	5.0	5.0	5.0	4.2
NC + MPR ($m = 4$)	6.3	6.3	6.3	5.0	6.3	5.0	6.3	5.0

Table 4.2: Gains in the saturated network throughput of the improved MAC relative to the saturated routing throughput using the 802.11 MAC.

4.2 Per-Node Throughput and Delay Performance of the Improved MAC

The gain provided by the use of multi-packet reception and network coding is dependent on the number of transmitting nodes N within the topology component. While the gain manifests itself in the throughput of each canonical topology component, the major benefit is realized in the delay, or time it takes to complete all flows. For purposes of illustration, the analysis is restricted to the cases in which the MAC uses MPR-adapted CSMA (i.e., simultaneous transmissions from connected nodes is minimized), traffic is symmetric across each topology component, and the improved fairness protocol is used. Combining equations (2.7) and (4.2) through (4.5), relaxing the integer constraints, and assuming an equal number of nodes in each set within the "X" topology component, the limiting throughput for each canonical topology component as $N \to \infty$ is:

$$\lim_{N\to\infty} S_{Cross} = \begin{cases} \frac{m}{m+1} & \text{without NC} \\ m & \text{with NC} \end{cases} \tag{4.8}$$

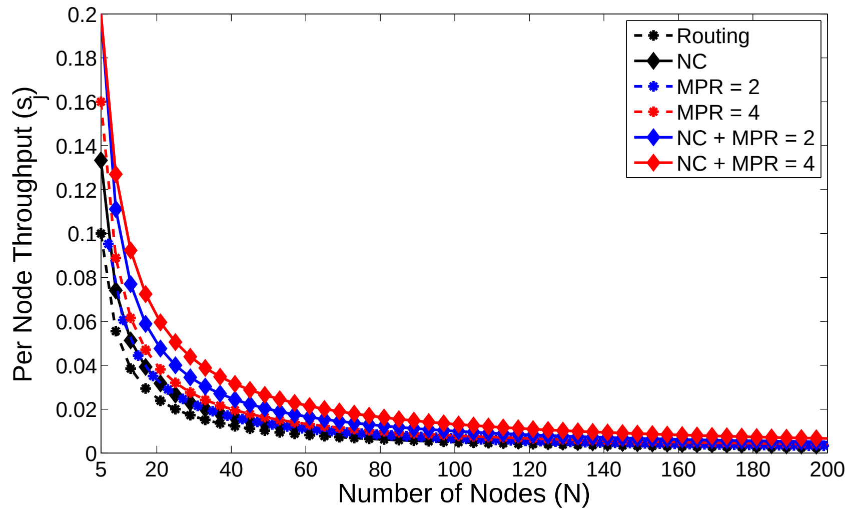

Figure 4.3: Throughput per node of the "X" topology component for large N using the improved MAC.

$$\lim_{N\to\infty} S_X = \begin{cases} \frac{m}{m+1} & \text{without NC} \\ \\ \frac{2m}{m+2} & \text{with NC.} \end{cases} \tag{4.9}$$

It is clear from the above equations that the gain has a dependency on the connectivity of the network. As the network becomes more connected, the interaction between network coding and multi-packet reception combine to create super-additive gains.

Considering the per-node throughput $S_{Node} = s_j$ for $j \in [1, N]$, equations (4.2) through (4.5) show that the throughput for both the original 802.11 MAC and improved MAC scales on the order of $1/N$. Fig. 4.3 shows the $1/N$ per node throughput behavior for the "X" topology component, using the improved MAC, as a function of the number of nodes. As expected, the throughput per node asymptotically approaches zero as N grows. While there are gains from multi-packet reception and network coding for moderately sized networks (i.e., $N = [5, 100]$), the throughput gains are limited for larger ones.

On the other hand, there are significant gains from multi-packet reception and network coding, while using the improved MAC, when considering the *delay*, or total time to complete all sessions. Fig. 4.4 shows the total time to complete all flows within the extended "X"

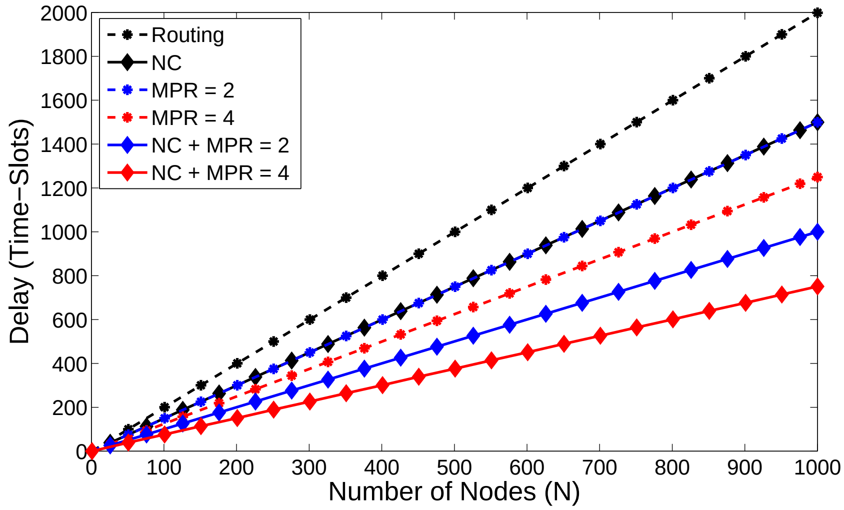

Figure 4.4: Time to complete all flows if each source has only a single packet to send using the improved MAC.

topology component, as N grows assuming that a single packet is distributed to each node. It can be easily verified from Fig. 4.4 that the delay gains for multi-packet reception with $m = 2$ or $m = 4$ and network coding cases are approximately 2 and $8/3$ respectively.

4.3 Asymmetric Traffic Performance of Multi-Packet Reception and Network Coding

The performance of opportunistic network coding and multi-packet reception in networks with bottlenecks is highly dependent on the symmetry of traffic across the bottleneck. Situations where the traffic is approximately symmetric, or equal, across the bottleneck maximizes the performance gains provided by both network coding and multi-packet reception as shown by the stars in Figures 3.8, 3.9, 4.1, and 4.2. The curves in each figure represent an averaging over instantaneous asymmetries in the traffic. This section provides insight into the effects of asymmetric traffic while using the improved fairness protocol. For the purposes of analyzing the effects of asymmetric traffic, the "X" topology component is used as the primary topology component in the analysis, since its limitations from the reduced number of nodes any given edge node can overhear compounds the effects of asymmetric traffic on network throughput.

Two different asymmetry scenarios are addressed. The first addresses the effects of asymmetry with a MAC that does *not* limit the number of nodes that transmit in either set X_1 or X_2 as defined in the previous sections. The MAC allows nodes within the same set to take advantage of multi-packet reception and does not restrict multiple nodes from sending to the relay in a given time-slot. If only nodes within the same set have data to send, the MAC allows for up to m nodes to send their respective packets to the relay. Figure 4.5 shows the throughput for an "X" topology component as the asymmetry between traffic from different sets is increased. The value plotted on the x-axis is the asymmetry ratio and is defined as:

$$\nu = \frac{\sum_{i \in X_2} k_i}{\sum_{j \in X_1} k_j}, \tag{4.10}$$

where k_i and k_j are the number of packets that each node $i \in X_2$ and $j \in X_1$, respectively, needs to send to a given node on the opposite side of the relay. In this scenario, the effectiveness of multi-packet reception is not diminished since it can be fully utilized regardless of where the traffic is originating. This results in a constant throughput independent of ν. On the other hand, the effectiveness of network coding decreases as ν increases. Figure 4.5 shows that the throughput for cases involving network coding is maximized for perfectly symmetric traffic flows, indicated by the stars in Figures 4.1 and 4.2 when $\nu = 1$, and saturates to the multi-packet reception only throughput for very large ν. However, Fig. 4.5 shows that network coding still provides significant gains for asymmetry ratios of less than five.

The second asymmetry scenario involves the use of a MAC that limits the transmission of nodes from the same set (i.e., MPR-adapted CSMA is used). In this scenario, it is assumed that nodes within the same set do not transmit at the same time unless the degree of multi-

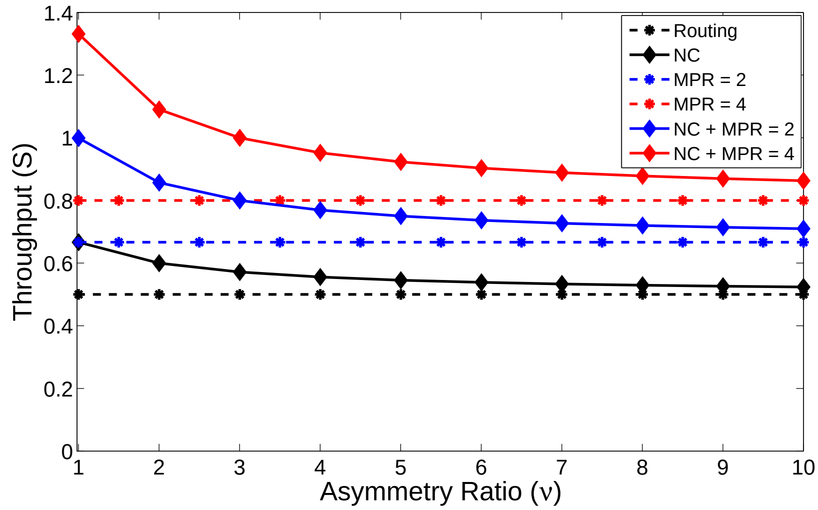

Figure 4.5: Throughput of an "X" topology component as a function of the asymmetry ratio with an offered load of 1 when CSMA is not used to limit transmission order.

packet reception, m, requires that they do so. When $m = 2$, only a single node in each set will transmit in a time-slot. As traffic becomes more asymmetric, one set of nodes will eventually run out of data and the other set will be forced to continue sending data to the relay one node at a time. Similarly, the center node will run out of data from different sets to code together when network coding is used alone; and as a result, network coding and multi-packet reception with $m = 2$ is the same for this scenario. For $m = 4$, two nodes from the same set will transmit in the same time-slot since the topology component contains only two sets of nodes. Extending this concept to large m requires that $^{m}/_{2}$ nodes on a given side of the relay be allowed to transmit in the same time-slot.

Figure 4.6 shows how asymmetric traffic from different sets affects network throughput when network coding is used. Unlike the case where the order is not enforced, the throughput when using multi-packet reception for some m, in addition to network coding, is reduced for large ν. For the case when $m = 2$ with or without network coding, the throughput is maximized when traffic is perfectly symmetric but saturates to the throughput obtained for the routing case. This saturation can be seen by limiting all traffic to be originated from a single set of nodes. The MAC will restrict transmission from each node to the center,

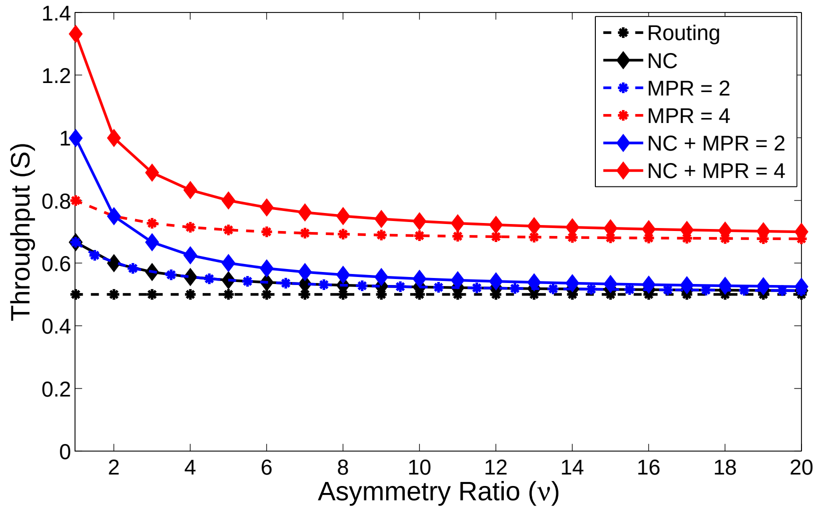

Figure 4.6: Throughput of an "X" topology component as a function of the asymmetry ratio with an offered load of 1 when MPR-adapted CSMA is used to limit transmission order.

which eliminates the gain resulting from the use of multi-packet reception; and the center node must send each relayed packet unencoded to the next hop, which eliminates the gain resulting from network coding. The case for $m = 4$ is similar to that of the $m = 2$ case; but only a maximum of two nodes in a set are allowed to transmit in the same time slot. The throughput will saturate for large ν to $2/3$, which is achieved for $m = 2$.

While implementation of a MAC that allows for the full employment of multi-packet may be more difficult, the potential throughput gains achieved when heavily asymmetric data is flowing through a bottleneck are significant. In any case, Figures 4.5 and 4.6 show that there is still significant gain from multi-packet reception, network coding, and their combination when traffic is substantially asymmetric ($\nu \leq 10$) with either implementation of the MAC. It is important to note that, in the presence of erasures, even with asymmetric traffic, the potential gains are significant. While network coding may not necessarily increase throughput under highly asymmetric loads, the network coding gain will manifest itself when recovering from packet erasures. Care must be taken to ensure that the MAC, multi-packet reception, and network coding design and implementation are done in conjunction with each other to realize the achievable gains.

4.4 Effects of Limiting Multi-Packet Reception to Only the Bottleneck Node

Since implementing multi-packet reception in a system may be a difficult and costly upgrade depending on the method of implementation, it is important to look at the throughput gains if strategic nodes for implementing multi-packet reception are targeted and the rest are left without the capability. The limitation of not having multi-packet reception at each edge node, as would be expected, reduces the effectiveness of opportunistic network coding and limits the total number of packets that the center node can code together.

In order to determine the throughput, an equal number of packets are deterministically distributed to each node and the throughput is calculated using equation (2.7) as the number of nodes N increases towards infinity. It is further assumed that each node has the ability to capture a packet. That is, if multiple transmissions occur in a given time-slot, a node without multi-packet reception will receive one transmission without error and treat the remaining transmissions as noise. If capture is not feasible, the network coding with multi-packet reception gain will equal the network coding alone gain for topology components such as the cross. The network coding with multi-packet reception gain for less connected topologies, in contrast, will be higher depending on the implementation of the MAC since the topology limitations decrease the probability of two node's transmissions conflicting.

The number of additional packets that the center node must send when each edge node does not have multi-packet reception is dependent on m. Limiting multi-packet reception to the center node essentially splits a component into m disjoint sets where all edge nodes in a set are fully connected and each node is connected to the center. An $m = 2$ will result in two disjoint sets that requires the center node to send $\lceil (N-1)/2 \rceil + 1$ degrees of freedom to each edge node in order to complete all unicast and broadcast sessions. The first term in this equation is the number of transmissions needed to relay all traffic from the set of edge nodes and the second is the number of transmissions needed to send the center's own traffic. In the case of the "X" topology component, the division has already been performed as a result of the topology configuration, so the throughput is the same as that found in Section 4.1.3. The throughput for the cross topology component becomes the same as that of the "X" topology component as a result of the limited implementation of multi-packet reception. An $m = 4$ results in four disjoint sets that requires the center node to send $\lceil (N-1)/2 \rceil + 1$ degrees of freedom to the set of edge nodes to complete all unicast sessions and $\lceil 3(N-1)/4 \rceil + 1$ degrees of freedom to each edge node to complete all broadcast sessions. Within both components, the result of increasing m is offset by the requirement that the center node send additional degrees of freedom. The broadcast throughput for both components becomes upper bounded by the throughput of the "X" topology component when using both network coding and $m = 2$; and the unicast throughput for network coding with $m = 4$ is upper bounded by $4/3$ for both topology components.

The cases where network coding is not used are unaffected by limiting multi-packet reception to the center node only. Since the center must forward all packets individually, it inherently communicates all of the necessary degrees of freedom to each of the edge nodes. These results are displayed in Figure 4.7 for the cross topology component and Figure 4.8 for the "X" topology component as N increases towards infinity, the loads are symmetric across all nodes, and the MAC is fully saturated. These figures show that there are significant throughput gains when considering network coding with multi-packet reception within topologies similar to the "X" topology component, but little for topologies that are more connected or for larger values of m. While the figures assume that $N \to \infty$, implementing multi-packet reception at each node, including the edge nodes, is the enabler that provides the very large super-additive gains obtained through the use of opportunistic listening for smaller values of N as well. Regardless of N, implementing multi-packet reception at only the bottle-neck node yields very limited gains in general.

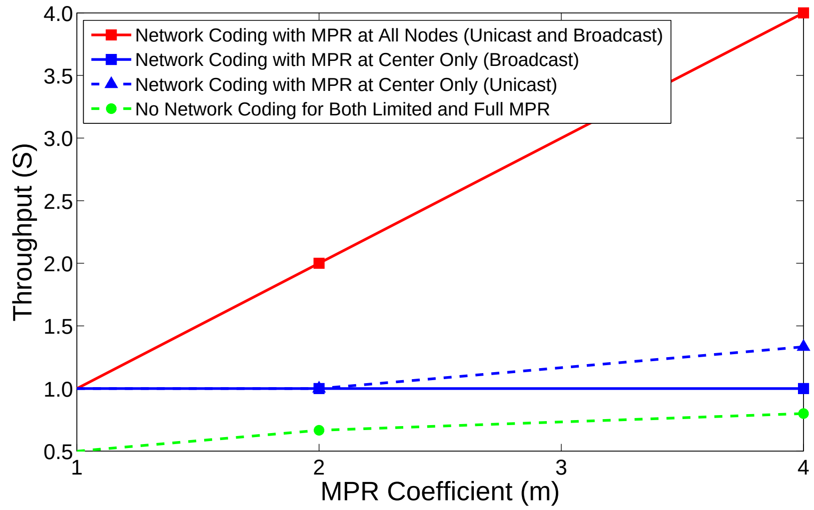

Figure 4.7: Unicast and broadcast throughput as $N \to \infty$ for the cross topology component with MPR implemented at each node and MPR implemented at only the center node under fully saturated and symmetric loading.

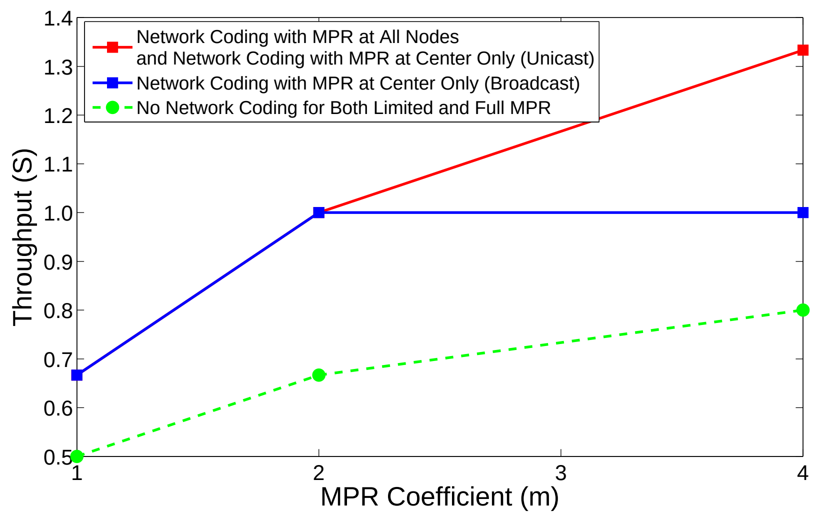

Figure 4.8: Unicast and broadcast throughput as $N \to \infty$ for the "X" topology component with MPR implemented at each node and MPR implemented at only the center node under fully saturated and symmetric loading.

4.5 Summary

This chapter focused on using the model developed in Chapter 2 to design a cross-layer solution that takes into account network coding and multi-packet reception. The proposed solution is a new MAC approach that takes into account multiple communication technologies, provides fairness to *flows* rather than *nodes*, and exhibits monotonic saturation behaviors. The model, in addition to the new MAC approach, was then used to look at several different measures of network performance such as total network throughput, per-node throughput, and delay. Section 4.1.3 further emphasized the super-additive gains that are achievable with the joint use of network coding and multi-packet reception and showed that the gains from the combined use of these two technologies are lower bounded by the sum of the gains provided by their individual implementations. Section 4.2 showed that the model can be used to determine the gains in per-node throughput and the delay to complete each session. While the per-node throughput decreases on the order of $1/N$, substantial per-node throughput gains are achievable with the combined use of network coding and multi-packet reception in smaller networks. In addition, the combined use of network coding and multi-packet reception results in gains of up to $8/3$ over routing alone when considering the time it takes to complete all unicast and broadcast sessions. While network coding and multi-packet reception may not have a considerable impact on per-node throughput in larger networks, there are significant gains in the delays experienced when using these two technologies in combination.

Finally, the effects of asymmetric traffic across the network's bottleneck and limited implementation of multi-packet reception were determined by using modifications of the topology components presented in previous chapters. Section 4.3 showed that the gains achieved from network coding and multi-packet reception are maximized for fully symmetric traffic across the bottleneck node and seriously impacted under highly asymmetric traffic conditions. Section 4.4 showed that selective implementation of multi-packet reception at strategic nodes in the network has little impact in sparser networks, but significant reductions in the achievable throughput should be expected in denser networks.

Chapter 5

Conclusions and Future Work

This thesis focused on developing a simple, intuitive approach to cross-layer wireless network design and provided an estimate for the achievable gains from implementing several different network technologies in multi-hop wireless networks. Specifically, the performance of various MAC approaches, opportunistic network coding, multi-packet reception, and their combination was determined. The benefits and drawbacks of implementing each technology was explored, and a new MAC approach was suggested to improve overall network performance.

In order to understand the design trade-offs, implementation details, and predict the performance of combining multiple communication technologies, a model was developed in Chapter 2 for various network elements that focused on the primary behaviors that are major contributors to network performance while idealizing the secondary and tertiary behaviors that have less of an effect on performance. This allowed a tractable model to be produced. This model could be easily modified to include any number of different network and communication technologies. Specifically, element models were proposed for three specific network technologies: opportunistic network coding, medium access control, and multi-packet reception. Opportunistic network coding was modeled by using COPE as a baseline and identifying the primary contributors to network performance provided by this scheme. A model for the 802.11 MAC was developed that focused on the fairness aspects of the protocol, which were shown in Chapter 3, as well as in [2], to be a primary contributor to the non-monotonic throughput saturation experienced in the COPE experiments conducted by Katti *et. al.* Multi-packet reception was modeled by taking a high level approximation of OFDMA, SDMA, CDMA, etc., by allowing for multiple packets to be received by any node in a given time-slot without collision or loss. Furthermore, a set of key canonical topology components that simulate multi-hop traffic and bottlenecks in larger networks were identified. Each specific topology component was shown to provide insight into the design implications and effects of implementing network coding, MAC, multi-packet reception, and their combination. Finally, a general framework to characterize the performance of each technology through both analysis and simulation was developed. This framework focused on characterizing network performance from a total network throughput perspective, but Chapter 4 showed how the framework could be extended to characterize performance in terms of delay and per-node throughput.

Chapter 3 provided a detailed explanation of the analysis for evaluating the performance of network coding, the 802.11 MAC, multi-packet reception, and their combination using the model proposed in Chapter 2. First, a direct comparison of the total network throughput obtained using the model and the experimental COPE data was made. This comparison validated the model by providing a direct correlation between the throughput behavior under varying offered loads with and without the use of network coding. In addition, the performance of network coding and multi-packet reception with the use of the 802.11 MAC were evaluated both separately and in combination with each other, using each of the topology components shown in Figure 2.4. The cross topology component was used to show the maximum achievable gains obtained with opportunistic network coding by simulating a dense network, the "X" topology component provided gains respective of a sparser network, and both the partial cross and partial "X" topology components showed the robustness of these gains to topology changes. In each case, it was shown that the 802.11 MAC leads to non-monotonic saturation behavior under high offered loads to the network by distributing channel resources equally among the competing nodes.

A new MAC approach that eliminates the non-monotonic saturation seen using the 802.11 MAC and provides fairness to *flows* as opposed to *nodes* was proposed in Chapter 4. The proposed MAC was designed using a cross-layer approach which takes into account multi-packet reception, network coding, the type of traffic (unicast or multicast/broadcast), and the local network topology configuration. Performing the same analysis and simulations outlined in Chapter 3, it was determined that the MAC provides monotonic saturation at the maximum achievable throughput seen in the previous chapter. Furthermore, the model helped determine that using network coding and multi-packet reception in combination provides a total network throughput that is up to 6.3 times that of using routing alone with the 802.11 MAC. The analysis and simulations also highlighted the fact that the combination of network coding and multi-packet reception provides *super-additive* additive gains. The analysis of the performance of the new MAC approach using the different topology components further aided in determining specific design details that need to be taken into account when implementing both network coding and multi-packet reception in a network. It was shown that, under the half-duplex constraint, multi-packet reception has the potential of reducing the effectiveness of network coding and a field size larger than two may be required when performing the linear network coding operations for multicast/broadcast traffic.

Extensions to the topology components aided in determining the per-node throughput, delay gains, and the effects of asymmetric traffic across the network's bottleneck. While the per-node gains provided by combining network coding and multi-packet reception scale on the order of $1/N$, significant gains are still obtainable in small to medium size networks (e.g., $N = [5, 100]$). These extended topology components further showed that while there are significant gains in total throughput, the time it takes to complete each unicast or multicast/broadcast session when using network coding and multi-packet reception is up to $8/3$ less than the time it takes when routing each packet individually through the network. Furthermore, asymmetric traffic across the bottleneck node was shown to significantly decrease the effectiveness of network coding and multi-packet reception. While not much can be done to eliminate the affects of asymmetric traffic on the throughput gains provided by network coding, the use of a MAC scheme that does not enforce transmission order counteracts the affects of asymmetric traffic with multi-packet reception. Specifically, it was shown that the use of a CSMA like scheme (i.e., a node will not transmit unless all of its neighbors are idle) under highly asymmetric traffic flows will significantly reduce the gains provided by multi-packet reception and that removing this constrain allows for the full potential of multi-packet reception to be reached. Finally, the proposed model was used to evaluate the performance of implementing multi-packet reception at a limited number of nodes. In denser networks, limited implementation of multi-packet reception significantly reduced the total network throughput; while for sparser networks, the difference between implementation of multi-packet reception at every node and at only strategic nodes was shown to be less pronounced.

Possible directions for future work include further validation of the proposed model and the inclusion of packet losses or errors into the model. Recent work by Sundararajan *et. al.* [24], Traskov *et. al.* [25], and Heindlmaier *et. al.* [26] using a graph theoretic approach for finding the optimal throughput for multicast traffic in an arbitrary network has shown significant promise. Extending these results to include multi-packet reception can help in determining the difference between the optimal throughput and the throughput obtained using the model proposed in this thesis. Additionally, experiments using network coding and multi-packet reception can provide additional data to further validate the model. Furthermore, the model developed in this thesis assumed a loss-less channel. Incorporating packet erasures and determining their effects on network performance is also a possible future avenue of research.

In summary, a cross-layer design is required to maximize the gains obtained with the joint implementation of network coding, multi-packet reception, and medium access control. A tractable model that matches experimental results with fidelity was the foundation of both the analysis and simulations presented in this thesis, as well as the design of a new medium access control approach that provides monotonic saturation throughput behavior. The model highlighted the super-additive throughput gains obtainable through the joint use of multi-packet reception and network coding, as well as provided key insights into the design and implementation details needed to maximize network performance.

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Bibliography

[1] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the Air: Practical Wireless Network Coding," *IEEE/ACM Transactions on Networking*, vol. 16, no. 3, pp. 497 - 510, June 2008.

[2] F. Zhao and M. Médard, "On Analyzing and Improving COPE Performance," in *Information Theory and Applications Workshop (ITA)*, February 2010, pp. 1 - 6.

[3] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204 - 1216, July 2000.

[4] S.-Y. Li, R. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371 - 381, February 2003.

[5] R. Koetter and M. Médard, "An Algebraic Approach to Network Coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782 - 795, October 2003.

[6] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A Random Linear Network Coding Approach to Multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413 - 4430, October 2006.

[7] D. Lun, N. Ratnakar, R. Koetter, M. Médard, E. Ahmed, and H. Lee, "Achieving Minimum-Cost Multicast: A Decentralized Approach Based on Network Coding," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 3, March 2005, pp. 1607 - 1617.

[8] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of Linear Coding in Network Information Flow," in *International Symposium on Information Theory, ISIT 2005. Proceedings.*, September 2005, pp. 264 - 267.

[9] Y. Wu, P. Chou, and S. Kung, "Information Exchange in Wireless Networks with Network Coding and Physical-layer Broadcast," *Conference on Information Sciences and Systems*, 2005.

[10] Z. Li and B. Li, "Network Coding: The Case for Multiple Unicast Sessions," in *42nd Annual Allerton Conference, Communication, Control and Computing*, 2004.

[11] X. Bao and J. Li, "Adaptive Network Coded Cooperation (ANCC) for Wireless Relay Networks: Matching Code-on-Graph with Network-on-Graph," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 574 - 583, February 2008.

[12] S. Sengupta, S. Rayanchu, and S. Banerjee, "An Analysis of Wireless Network Coding for Unicast Sessions: The Case for Coding-Aware Routing," *INFOCOM*, 2007.

[13] J. Le, J. Lui, and D. M. Chiu, "How Many Packets Can We Encode? - An Analysis of Practical Wireless Network Coding," *INFOCOM*, 2008.

[14] H. Seferoglu and A. Markopoulou, "Network Coding-Aware Queue Management for Unicast Flows over Coded Wireless Networks," in *2010 IEEE International Symposium on Network Coding (NetCod)*, June 2010, pp. 1 - 6.

[15] J. J. Garcia-Luna-Aceves, H. R. Sadjadpour, and Z. Wang, "Challenges: Towards Truly Scalable Ad Hoc Networks," in *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2007.

[16] L. Tong, Q. Zhao, and G. Mergen, "Multipacket Reception in Random Access Wireless Networks: From Signal Processing to Optimal Medium Access Control," *Communications Magazine, IEEE*, vol. 39, no. 11, pp. 108 - 112, November 2001.

[17] Q. Zhao and L. Tong, "A Multiqueue Service Room MAC Protocol for Wireless Networks with Multipacket Reception," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 125 - 137, February 2003.

[18] S. Ghez, S. Verdú, and S. Schwartz, "Stability Properties of Slotted Aloha with Multipacket Reception Capability," *IEEE Transactions on Automatic Control*, vol. 33, no. 7, pp. 640 - 649, July 1988.

[19] G. Celik, G. Zussman, W. Khan, and E. Modiano, "MAC for Networks with Multipacket Reception Capability and Spatially Distributed Nodes," *IEEE Transactions on Mobile Computing*, vol. 9, no. 2, pp. 226 - 240, February 2010.

[20] J. J. Garcia-Luna-Aceves, H. R. Sadjadpour, and Z. Wang, "Extending the Capacity of Ad Hoc Networks Beyond Network Coding," in *Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing (IWCMC)*, 2007.

[21] A. Rezaee, L. Zeger, and M. Médard, "Multi-Packet Reception and Network Coding," in *MILITARY COMMUNICATIONS CONFERENCE (MILCOM)*, November 2010, pp. 1393 - 1398.

[22] "IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pp. C1 - 1184, December 2007.

[23] A. Duda, "Understanding the Performance of 802.11 Networks," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2008, pp. 1 - 6.

[24] J. K. Sundararajan, M. Médard, R. Koetter, and E. Erez, "A Systematic Approach to Network Coding Problems Using Conflict Graphs," in *invited paper, UCSD Workshop on Information Theory and its Applications*, February 2006.

[25] D. Traskov, M. Heindlmaier, M. Médard, R. Koetter, and D. Lun, "Scheduling for Network Coded Multicast: A Conflict Graph Formulation," in *IEEE GLOBECOM Workshops*, December 2008, pp. 1 - 5.

[26] M. Heindlmaier, D. Traskov, R. Kotter, and M. Médard, "Scheduling for Network Coded Multicast: A Distributed Approach," in *IEEE GLOBECOM Workshops*, December 2009, pp. 1 - 6.