Effects of MAC Approaches on Non-Monotonic Saturation with COPE - A Simple Case Study

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Cloud, Jason, Linda Zeger, and Muriel Medard. “Effects of MAC Approaches on Non-monotonic Saturation with COPE - a Simple Case Study.” MILCOM 2011 Military Communications Conference. 747–753.</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1109/MILCOM.2011.6127766">http://dx.doi.org/10.1109/MILCOM.2011.6127766</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Institute of Electrical and Electronics Engineers (IEEE)</td>
</tr>
<tr>
<td>Version</td>
<td>Author's final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Dec 16 17:33:40 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/73927">http://hdl.handle.net/1721.1/73927</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution-Noncommercial-Share Alike 3.0</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">http://creativecommons.org/licenses/by-nc-sa/3.0/</a></td>
</tr>
</tbody>
</table>
Effects of MAC Approaches on Non-Monotonic Saturation with COPE - A Simple Case Study

Jason Cloud*†, Linda Zeger†, Muriel Médard*
*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA.
†MIT Lincoln Laboratory, Lexington, MA.
Email: {jcloud@, zeger@ll., medard@}mit.edu

Abstract— We construct a simple network model to provide insight into network design strategies. We show that the model can be used to address various approaches to network coding, MAC, and multi-packet reception so that their effects on network throughput can be evaluated. We consider several topology components which exhibit the same non-monotonic saturation behavior found within the Katti et. al. COPE experiments. We further show that fairness allocation by the MAC can seriously impact performance and cause this non-monotonic saturation. Using our model, we develop a MAC that provides monotonic saturation, higher saturation throughput gains and fairness among flows rather than nodes. The proposed model provides an estimate of the achievable gains for the cross-layer design of network coding, multi-packet reception, and MAC showing that super-additive throughput gains on the order of six times that of routing are possible.

I. INTRODUCTION

With the multitude of network technologies available that increase performance, it is difficult to efficiently integrate them into a coherent network. We develop a simple model that provides insight into cross-layer network design, and illustrate its use through examples involving the combination of the 802.11 medium access control (MAC), network coding (NC), and multi-packet reception (MPR). Our model not only provides strategies to integrate various technologies, but also predicts the achievable gains possible.

The development of the opportunistic inter-session NC scheme, COPE, by Katti et. al. [1] led to various models and analyses that attempted to explain COPE’s experimental results. Le et. al. [3] and Sengupta et. al. [4] developed models to describe these results, but only considered coding a maximum of two packets together at a time or did not address the interaction between NC and MAC fairness. As a result, their models provide throughput gains that are considerably smaller than the experimental results and do not explain the non-monotonic behavior shown in the upper half of Fig. 1.

Zhao and Médard [2] showed that the fairness imposed by the 802.11 MAC helps to explain this non-monotonic behavior. In addition, they demonstrated that the majority of the throughput gain achieved from COPE is a result of coding three or more unencoded, or native, packets together at time. Their analysis showed that these gains are not reflected in three node network models and that at least five nodes are required to accurately capture the effects of COPE. Fig. 1 shows the 802.11 MAC model and 5-node cross topology component from [2] (bottom). We start with these elements to develop a model that can be used for cross-layer design of various network technologies. The NC+MPR+MAC, NC+MAC, and Routing+MAC curves are derived from our model and provide an estimate of the achievable performance gains.

This work is sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, recommendations, and conclusions are those of the authors and are not necessarily endorsed by the United States Government. Specifically, this work was supported by Information Systems of ASD/R&E. Contributions of the Irwin Mark Jacobs and Joan Klein Jacobs Presidential Fellowship have also been critical to the success of this project.

Figure 1. Comparison of the empirical COPE performance data collected from a 20-node 802.11 wireless ad hoc network test bed (top), [1], and the resulting throughput using a model of the 802.11 MAC and 5-node cross topology component proposed by [2] (bottom). We start with these elements to develop a model that can be used for cross-layer design of various network technologies. The NC+MPR+MAC, NC+MAC, and Routing+MAC curves are derived from our model and provide an estimate of the achievable performance gains.
and other possible combinations of these 5-node topology components in order to help in our understanding of the effects of combining NC and various MAC implementations in larger networks.

Finally, we demonstrate the capability of the model to predict the gains and highlight design challenges from incorporating additional technologies, such as MPR, into the network. MPR is known to enhance wireless network performance and has been extensively researched with unencoded traffic [6, 7, 8, 9]; but little on the joint use of MPR, NC, and MAC design exists. Garcia-Luna-Aceves et. al. [10] compared the use of NC to MPR, but did not consider their combined use; and Rezae et. al. [11] provided an analysis of the combined use of NC and MPR in a fully connected network, but did not consider the effects of bottlenecks or multi-hop traffic. We show that our model provides an intuitive method for determining an estimate of the achievable gains from the combined use of MPR and NC in a congested, multi-hop network.

The remainder of the paper is organized as follows: Section II provides a detailed description of the network model; Section III provides an example of NC and MPR for 5-node network topology components using the existing 802.11 MAC; Section IV demonstrates the flexibility of using the model when considering the design of various network elements; and we conclude in Section V.

II. NETWORK MODELS AND PARAMETERS

Our main goal is to develop a simple model that gives insight into cross-layer design of wireless networks by using NC, various MAC approaches, and MPR as examples. To this end, we identify the fundamental behavior of each aspect of the network and model each element using simple, intuitive methods so that we can evaluate the potential throughput gains.

The performance of NC is modeled by considering the ability of a given node to combine multiple packets together as well as the primary implementation details of the particular NC scheme. We use COPE [11] as a case study. COPE uses the broadcast nature of the wireless channel to opportunistically code packets from different nodes together using a simple XOR operation. Any node that receives an encoded message is able to decode it using the unencoded, or native, packets captured from the wireless channel. We model COPE such that packet transmissions are never delayed. If a node does not have more than one packet to encode, it does not wait for another packet to arrive. Rather, it sends the packet unencoded at the first opportunity. In addition, all packets headed towards the same next-hop will not be encoded together because the next-hop would not be able to decode these coded packets due to the lack of enough degrees of freedom. We do not consider the complexity of the coding or decoding operations nor any other aspects of the NC implementation since their contributions to the overall network performance is small in relation to the specific implementation aspects mentioned above.

The MAC is modeled by identifying its primary behavior in the network and simplifications are made by assuming optimal performance from its secondary and tertiary behaviors. This gives us an intuitive approach in determining the potential throughput while ensuring that we understand the fundamental characteristics of the network. For instance, the 802.11 MAC employs a fairness mechanism that distributes channel resources equally among all competing nodes within a network. As we show in the following sections, this is the primary cause of the non-monotonic saturation behavior in the experimental throughput shown in the upper half of Fig. 1. Since the fairness mechanism is the major contributor to overall network performance, we assume optimal performance from all other aspects of the MAC. For example, the non-monotonic behavior 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, we do not consider the additional effects on overall throughput associated with the virtual 802.11 CS mechanisms (RTS/CTS) or other aspects of the MAC such as the potential of lost channel resources due to the MAC’s random back-off. Instead, we assume optimal performance from each of these secondary and tertiary behaviors. These assumptions, as a result, provide upper bounds to the achievable throughput in the various networks that employ the MAC.

We also show how additional techniques to increase network performance, such as MPR, can be similarly modeled. MPR can be implemented in a variety of methods from Code Division Multiple Access (CDMA) or multiple-input-multiple-output (MIMO) to orthogonal frequency division multiple access (OFDMA). In subsequent sections, we model MPR by allowing each node to successfully receive m multiple packets at the same time.

Finally, our model uses several basic canonical topology components that contain only five nodes where each node is both a source and a sink. Two of the possible components are shown in Fig. 2. These components are of interest because they form the primary structures in larger networks that create bottlenecks and congestion. By looking at the traffic that travels through the center node, these components help us model the performance gains of multi-hop traffic under both low and high loads. While all possible combinations of these basic canonical topology components should be evaluated, we only focus on two of the components, shown in Fig. 2, since the analysis of the others are redundant and provide little to the clarification of our approach.

Each component has specific constraints due to their structure and will effect the performance of the MAC, NC, and MPR in different ways. The center node $n_5$ in each component is fully connected regardless of the topology, and traffic flows originating from the center require only a single hop to reach their destination. Within the cross topology component, each traffic flow originating from a given node is terminated at the node directly opposite the center; and in the “X” topology component, all flows originating from a node in a given set terminates at a node in the opposite set. Therefore, each flow must pass through the center regardless of topology. For
example, nodes \(n_1, n_2,\) and \(n_5\) in the “X” topology component are fully connected and nodes \(n_3, n_4,\) and \(n_5\) are also fully connected; but \(n_1\) and \(n_2\) are not connected to \(n_3\) and \(n_4.\) All traffic between any node \(\{n_1, n_2\} \in X_1\) and any node \(\{n_3, n_4\} \in X_2\) must travel through the center.

In an effort to simplify the following explanation, we make several additional simplifications which can be easily incorporated into the model. We assume feedback is perfect, the load required for acknowledgments are contained as part of the initial transmission’s load, and the wireless channel is lossless. We also consider the additional constraint that each node is half-duplex, i.e. a node will overhear any transmission from its neighbors only if it is not transmitting. When considering the model for MPR, we allow \(m\) packets to be sent from different sources in a single time slot with the constraint that we try to maximize the number of neighbors a node can overhear within any given time slot. In essence, we duplicate CSMA/CA for each \(m = 2\) case in the sense that a node will transmit only if none of its neighbors are transmitting. For cases where \(m \geq 3,\) we pick the combination of transmitting nodes such that the average number of transmissions received by any given node within the network is maximized.

Sections [III and IV] provide both an analysis of the maximum achievable throughput and simulations over various values of the total offered load \(P\) to the network. The load \(P\) to the network from the set of source nodes \(i \in \mathcal{N},\) is defined as \(P = \sum_{i \in \mathcal{N}} \rho_i,\) where \(\mathcal{N}\) is the set of nodes in the topology component and \(\rho_i = k_i/100\) is node \(i\)’s individual load contribution, or the fraction of time required to send all of its \(k_i\) packets to the next-hop. We stochastically determine \(k_i\) using a binomial distribution given \(P\) with parameters \(n_i = K - \sum_{j=1}^{i-1} k_j\) and \(p_i = 1/(N-i+1), N = |\mathcal{N}|,\) in each iteration of our simulation and average these results for each total offered load evaluated. We then use the model described in this section to determine the total network throughput \(S\) which is equivalent to the total number of packets divided by number of time slots needed to send every packet to its intended destination.

III. MULTI-PACKET RECEPTION AND NETWORK CODING PERFORMANCE ANALYSIS

With each of the network topology components shown in Fig. 2, we provide an example of the performance analysis with and without the use of NC and MPR when the 802.11 MAC is used. We also consider both unicast and broadcast traffic where a unicast transmission is complete when all packets from each source successfully reaches their destination and a broadcast transmission is complete when all nodes have received every packet from all sources.

A. Cross Topology Component Analysis

Each node \(i \in [1, 5,]\) requires \(\rho_i\) of the time to send all of its packets one hop. The resulting total offered load to the network is then \(P = \sum_{i=1}^{5} \rho_i.\) We define the total network component load \(P_T = \rho_R + \rho_M\) as the time required to send all packets through the topology component where \(\rho_R\) is the required load to relay packets and \(\rho_M\) is the required load to send each native packet one hop. The load required to relay packets through the center is \(\rho_R = \frac{1}{c} \sum_{j=1}^{4} \rho_j\) where \(c\) is the number of packets that can be encoded together which is dependent on the number of neighbors a given node can overhear. In the case of the cross topology component and enough packets to code together, \(c = 4\) for \(m = \{1, 2\}\) and \(c = 2\) for all \(m \geq 3.\) The load needed to send each node’s unencoded packets one hop is defined as \(\rho_M \geq \frac{1}{m} \sum_{j=1}^{4} \rho_j + \rho_s.\) Because we setup the model so that we maximize the number of neighbors any given node can overhear when using MPR, this expression is met with equality if each \(\rho_j, j \in [1, 4,]\) is equal. Otherwise, the load \(\rho_M\) is lower bounded by this equation and is a function that is dependent on both the topology component’s configuration and each node’s load imbalance. Finally, let the fraction of allocated time slots a node receives as a result of the MAC be \(s_i.\)

The throughput \(S\) with unicast and broadcast traffic is shown as a function of \(P\) in Fig. 3 Each curve is obtained through simulation and is an average over the load distribution discussed in Section [II. Each star is obtained by analysis and depicts the maximum achievable throughput when the MPR and/or NC gain is maximized. When \(P_R < 1,\) each node is allocated enough time slots to send all of its packets, and the allocated load is \(s_j = \rho_j\) for \(j \in [1, 4]\) and \(s_5 = \rho_5 + \rho_R.\) The throughput \(S\) increases linearly as the network load increases, regardless of the use of MPR or NC, and reaches a maximum for each case when \(P_R = 1.\) The throughput then transitions into a saturated region for \(P_R > 1,\) where the allocated load for each node is \(s_j \leq \rho_j\) and \(s_5 \leq \rho_5 + \rho_R.\) When NC is not used, the throughput is \(S = s_5\) and when NC is used, the throughput will be a function of the number of packets that can be encoded together.

1) Routing (No Network Coding, \(m = 1\)): We use routing as the performance baseline. Consistent with the results found in [I] and the analysis performed in [II], the throughput increases linearly within the non-saturated region, \(P \in [0, 5/9].\) At \(P = 5/9,\) the throughput reaches a maximum of \(5/9\) (depicted by a star in Fig. 3). This occurs when each source’s load reaches \(\rho_i = 1/9\) for \(i \in [1, 5].\) The total load of the center node, as a consequence, is \(\rho_5 + \rho_R\) where \(\rho_R = \sum_{i=1}^{4} \rho_j = 4/9.\) Since \(P_R = 1, s_j = \rho_j\) and \(s_5 = \rho_5 + \rho_R.\)

The throughput saturates for \(P > 5/9.\) Initially, the 802.11 MAC allocates time slots to nodes requiring more resources.
The throughput is therefore the amount of time \( n_5 \) is able to transmit, \( s_5 = 1 - \sum_{i=1}^{4} s_i \), 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. \( s_i = \frac{1}{5} \) for \( i \in [1, 5] \). For large enough \( P \), the throughput saturates to the total amount of information that \( n_5 \) can transmit, i.e., \( S = s_5 = \frac{1}{5} \).

2) Network Coding Only (\( m = 1 \)): We now allow NC to be used by the center node. Each node transmits one at a time, allowing each node to receive four native packets or four degrees of freedom (three degrees of freedom through the use of opportunistic listening plus one degree of freedom from the packet originating at the 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.

From Fig. 3 when \( P \in [0, \frac{5}{9}] \), NC 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 \( P \in [\frac{5}{9}, \frac{5}{6}] \), NC is instrumental in achieving the throughput shown. The MAC does not limit channel resources until the maximum throughput of \( S = \frac{5}{6} \) is reached when \( P_T = \sum_{j=1}^{5} \rho_j + \frac{1}{2} \sum_{j=1}^{5} \rho_j = 1 \) where \( \rho_i = \frac{1}{6} \) for \( i \in [1, 5] \). At this maximum, the MAC ensures fairness among all competing nodes and the throughput saturates. As \( P \) increases, the gain provided by NC diminishes. The number of packets reaching \( n_5 \) from each edge node is limited by the MAC while packets introduced into the network component by \( n_5 \) are not. The coding gain, therefore, approaches zero as \( P \to \infty \).

3) Multi-Packet Reception of Order 2 and 4 (No Network Coding and \( m = \{2, 4\} \)): MPR is similar to the routing case described earlier except we now allow a maximum of \( m \) edge nodes 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 \( \frac{1}{2} \) that needed by routing while the center node cannot transmit multiple packets simultaneously and must transmit each received packet individually. Using CSMA, which restricts nodes opposite each other to transmit at the same time, the point at which the protocol saturates for symmetric source loads occurs when \( P_T = \sum_{j=1}^{4} \rho_j + \frac{1}{2} \sum_{j=1}^{4} \rho_j + \rho_5 = 1 \) where \( \rho_i = \frac{1}{7} \) for \( i \in [1, 5] \). This yields the maximum throughput of \( S = \frac{5}{7} \). 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.

The behavior for \( m = 4 \) is the same as that for \( m = 2 \) except the maximum of \( S = \frac{5}{6} \) occurs when \( P_T = \sum_{j=1}^{4} \rho_j + \frac{1}{4} \sum_{j=1}^{4} \rho_j + \rho_5 = 1 \) where \( \rho_i = \frac{1}{6} \). We allow all edge nodes to transmit their packets to \( n_5 \) simultaneously, requiring a total of \( \frac{1}{6} \) 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 transmission. As \( P \) increases, the MAC limits each node’s number of available time slots and \( S \) saturates to \( \frac{1}{5} \). Again, the gain in the saturated region for \( m = 4 \) is equal to the cases of \( m = 2 \) and routing.

4) Network Coding with Multi-Packet Reception of Order 2 and 4 (\( m = 2, 4 \)): For \( m = 2 \), the maximum throughput of \( S = \frac{5}{4} \) occurs when \( P_T = \frac{1}{4} \sum_{j=1}^{4} \rho_j + \frac{1}{2} \sum_{j=1}^{4} \rho_j + \rho_5 = 1 \) where \( \rho_i = \frac{1}{4} \) for \( i \in [1, 5] \). Each set of nodes, \( \{n_1, n_3\} \) and \( \{n_2, n_4\} \), uses \( \frac{1}{4} \) 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 NC throughput due to the 802.11 MAC. While the maximum achievable throughput is \( \frac{25}{16} \) times the NC without MPR throughput, the saturated gain for \( m = 2 \) is equal to the gain found when NC was used alone in this region.

The throughput using NC and \( m = 4 \) for unicast traffic is equivalent to NC and \( m = 2 \). All four edge nodes transmit to \( n_5 \) which then transmits two encoded packets in addition to its own; or we limit the number of simultaneous transmissions 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 \( S = \frac{5}{4} \) for the second which is consistent with the maximum unicast throughput. This difference indicates that increasing \( m \) when using NC may not be the optimal strategy. Although we do not show it here, the canonical topology components can be easily modified to include any number of nodes which would allow us to further look into the optimal strategy for broadcast traffic.

B. “X” Topology Component

The cross topology component gives insight into the performance of COPE and MPR in a dense network, and it represents the best case scenario when COPE is used since it maximizes the number of transmissions any given node receives. In order to understand the behavior of COPE and MPR in sparser networks, we limit the number of each node’s neighbors by analyzing the behavior of COPE and MPR in the “X” topology component shown Fig. 3(b). Fig. 4 shows both...
the maximum and average throughput resulting from the use of the “X” topology component. Within this section, we focus only on the cases involving NC since it can be easily verified that the routing and \( m = \{2, 4\} \) analysis for this topology component are the same as the cross topology component analysis.

1) Network Coding Only (\( m = 1 \)): Limiting the ability to overhear other edge nodes in the component results in the reduction in 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 \( P_T = 1/2 \sum_{j=1}^{4} \rho_j + \sum_{j=1}^{4} \rho_j = 1 \) where \( \rho_i = 1/7 \) for \( i \in [1, 5] \). 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. By comparing the results obtained from both the cross and “X” topology components, it is evident that the performance of COPE is highly dependent on the network structure. As the network becomes sparser, the gain from COPE is diminished.

2) Network Coding with Multi-Packet Reception of Order 2 and 4 (\( m = \{2, 4\} \)): For \( m = 2 \), the throughput increases linearly until it reaches its maximum at \( S = 1 \) when \( P_T = 1/2 \sum_{j=1}^{4} \rho_j + 1/2 \sum_{j=1}^{4} \rho_j + \rho_5 = 1 \) where \( \rho_i = 1/5 \) for \( i \in [1, 5] \). The throughput then saturates to the NC throughput for \( P_T > 1 \). For \( m = 4 \), the maximum unicast throughput is \( S = 5/4 \) and is achieved when \( P_T = 1/2 \sum_{j=1}^{4} \rho_j + 1/4 \sum_{j=1}^{4} \rho_j + \rho_5 = 1 \) where \( \rho_i = 1/4 \) for \( i \in [1, 5] \). The center node codes a maximum of two packets together from different edge node sets and transmits two encoded packets back to the set of edge nodes in addition to its own native packet. This gives each edge node enough degrees of freedom to complete all unicast transmissions. When considering broadcast traffic, each node still requires a maximum of one additional degree of freedom. Allowing \( n_5 \) to code all of the edge node’s native packets together and send one encoded transmission enables each node to extract the required degree of freedom and obtain the full set of messages from each source. The maximum throughput is therefore the same as the case for NC with \( m = 2 \) and is equal to \( S = 1 \). Similar to the cross topology component, the average throughput for both cases discussed in this section does not reach the maxima found because of the stochastic load distribution, which results in asymmetric traffic flows across the center node that limits the effectiveness of both COPE and the implementation of MPR that we chose. If each node has an equal amount of information to send, the maxima found in this section would be achieved.

The use of these components within our model allows us to determine the fundamental behavior of combining COPE and MPR in a larger network. For example, Fig. 5 shows a summary of our analysis by plotting the maximum unicast and broadcast throughput as a function of the MPR capability. It shows the super-additive throughput behavior when MPR is used in conjunction with NC by comparing this throughput with the throughput that would be obtained by adding the individual gains obtained using MPR and NC separately.

The use of the “X” topology component allows us to determine behavior that was otherwise masked in the cross component. It provided insight into the behavior of NC and MPR in sparser networks, methods in which to implement variants of COPE for broadcast traffic, and highlighted the super-additive gains from combining the two communication technologies. We can also look into variants of the cross and “X” components; but they are not discussed since they provide little clarification to the presentation of the model. In this example, however, these variants provide insight into the robustness of MPR and COPE’s throughput gains to topology changes.
IV. IMPROVING THE MAC FAIRNESS PROTOCOL

The use of our model highlighted a major drawback to the use of the 802.11 MAC in multi-hop networks. The non-monotonic throughput behavior that is evident in Fig. 3 and 4 significantly reduces throughput as the offered load to the network increases. We now show how the model can be used to redesign the MAC so that we eliminate this sub-optimal behavior in the presence of NC, MPR, and their combination. Furthermore, we use the model to develop a MAC approach that provides fairness to flows rather than to nodes.

It is obvious from Section III that if we want to eliminate the non-monotonic saturation throughput behavior, we want to allocate more channel resources to the center node than the edge nodes. We choose to allocate resources proportional to the amount of non-self-generated traffic flowing through each node when the network saturates. While allocating fewer resources to flows originating at the center and more resources to flows originated at edge nodes yields even higher throughput, this approach ensures that each flow of information is given the same priority. The center 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.

We design the revised MAC using a slight modification of the components found in Fig. 2. For the cross topology component, we let there be \(N-1\) 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. For the “X” topology component, we also let there be \(N-1\) 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 also connected to the center node. Within the cross topology component, each node communicates with the node directly opposite the center. In the “X” topology component, each node communicates with a node in a different set.

The allocated number of time slots each node receives so that the throughput \(S\) is maximized, subject to the flow constraints and \(\sum_{j=1}^{N-1} s_j/m + s_R = 1\), is divided into three cases where \(s_j\) 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. Similar to Section III, the throughput \(S = s_R\) when NC is not used. When NC is used, \(S\) is a function of the number of packets that can be coded together, which is dependent on the density of the network, the MPR coefficient \(m\), the use of CSMA, and the traffic type (unicast or broadcast). In order to simplify the explanation within this paper, we limit the examples we explore by considering only values of \(m = \{1, 2, 4\}\) and symmetric source loads. In the case of the “X” topology component, we also restrict our example to situations where the cardinality of each set \(X_1\) and \(X_2\) are equal. The cases include:

Cross Topology Component with Unicast Traffic or Broadcast Traffic: The cross topology component can be used to design part of the MAC for operation in dense networks.

Without NC, the center node requires a number of time slots equal to the number of source nodes \(N\). With NC, throughput is maximized by ensuring the center node codes the maximum number of native packets together. Implementing MPR for \(m = 4\) can potentially prevent each node from immediately decoding any coded message sent by the center. This is due to the potential of a given node transmitting at the same time as one of its neighbors. Generalizing for \(N\) and \(m\) as well as considering only integer numbers of time slots:

\[
S_j = \begin{cases} 
\frac{1}{\lfloor(N-1)/m\rfloor + 1} & \text{without NC} \\
\frac{1}{\lfloor(N-1)/m\rfloor + m} & \text{with NC}
\end{cases}
\]

and

\[
S_R = \begin{cases} 
\frac{N}{\lfloor(N-1)/m\rfloor + N} & \text{without NC} \\
\frac{N}{\lfloor(N-1)/m\rfloor + m} & \text{with NC}
\end{cases}
\]

“X” Topology Component: Using the “X” topology component helps gain insight into the design of the MAC for operation in sparser networks. From the 802.11 example in Section III-B, we determined that the throughput differs for both unicast and broadcast traffic. As a result, we define the fraction of time slots \(s_i^U\) allocated to each node for unicast traffic as:

\[
s_j = s_j^U = \begin{cases} 
\frac{1}{\lfloor(N-1)/m\rfloor + X_1} & \text{without NC} \\
\frac{1}{\lfloor(N-1)/m\rfloor + \max(|X_1|,|X_2|) + 1} & \text{with NC}
\end{cases}
\]

and

\[
s_R = s_R^U = \begin{cases} 
\frac{N}{\lfloor(N-1)/m\rfloor + N} & \text{without NC} \\
\frac{N}{\lfloor(N-1)/m\rfloor + \max(|X_1|,|X_2|) + 1} & \text{with NC}
\end{cases}
\]

When considering broadcast traffic, additional degrees of freedom may need to be sent by the center to complete the session. For \(m = \{1, 2\}\), no additional degrees of freedom are required by any node. For the case involving NC and \(m = 4\), each edge node will require one additional degree of freedom in order to decode all of the encoded packets sent by the center. As a result, the denominator in the NC equations of (3) and (4) is replaced by \(\lfloor(N-1)/4\rfloor + \max(X_1, X_2) + 2\), as well as the numerator in the NC case of (4) with \(\max(X_1, X_2) + 2\).

We apply the revised fairness protocol to both the 5-node cross and “X” topology components using the same methods described in Section III. We find that the throughput saturates at the maximum found in Section III for each topology component. Fig. 6 and 7 show both the unicast and broadcast throughput for the cross and “X” topology components, respectively, using our improved MAC approach. It is clear by comparing Figures 3 and 4 with Figures 6 and 7 respectively that the new MAC eliminates the non-monotonic saturation behavior. Furthermore, this comparison shows that the combination of NC, MPR with \(m = \{2, 4\}\), and the new MAC provides throughput gains on the order of six times that of routing alone with the 802.11 MAC in the saturated throughput regime.
V. CONCLUSION

We developed a simple, intuitive model by approximating key network elements with simple models of their underlying primary behavior. We then used the model to evaluate the performance of specific implementations of the 802.11 MAC with COPE and MPR in multi-hop networks. Gaining key insight into design strategies for combining the three technologies, each scenario presented gave a rough order of magnitude for the performance of implementing the MAC, NC, MPR, and their combination in larger networks. The model further shows that combining COPE with MPR results in super-additive throughput gains. We then demonstrated that the non-monotonic saturation experienced in [1] is explained by the sub-optimal behavior of the 802.11 MAC, and used our model to develop a MAC approach tailored toward the combined use of COPE and MPR that provides monotonic saturation behavior, as well as fairness to flows rather than nodes.

REFERENCES