Topological Arrangement of Nodes in Wireless Networks Suitable for the Implementation of Network Coding

F.J. Böning, A.S.J. Helberg, M.J. Grobler
TeleNet Research Group
School for Electrical, Electronic and Computer Engineering
North-West University, Potchefstroom, South Africa
phone: +2718 299 1961; fax: +2718 299 1977
e-mail: 20072155@nwu.ac.za

Abstract—In this paper, we discuss the physical arrangement of wireless nodes to form topologies suitable for the implementation of Network Coding. We calculate the areas in which each node must be located for a specific Network coding suitable topology to be formed. We simulate these identified topologies in OPNET Modeler to analyze the practicality of implementing network coding in these topologies. We provide results indicating the trade-off between reduced network load and higher end-to-end delay. These results indicate that implementation of the simulated topologies will be valuable for future research.

Index Terms—Ad hoc Networks, Network Coding, Node Placement, Topology Boundaries

I. INTRODUCTION

Wireless ad hoc networks offer many advantages over traditional wired networks. This includes increased mobility, ease of use and lower installation complexity. A disadvantage however, is a lower achievable transmission rate as a result of limited bandwidth and interference. They also require complex routing protocols, which increases overhead. Therefore, efforts are being made to increase communication rates by using new concepts such as network coding. Very little research has been done on the implementation of network coding and the feasibility thereof. The work in this paper was initiated in reaction to the results obtained in [1]. The authors recommend the use of topological information to identify opportunities for the implementation of network coding in wireless networks. They do this by looking for “known” network coding suitable topologies within a larger network. In this paper, we investigate the physical limitations in terms of node placement of these “known” topologies through mathematical modeling. We also implement network coding in the OPNET simulation environment in these known topologies and compare the results to scenarios where network coding is not implemented.

The remainder of this paper is organized as follow: Section II gives an brief background on network coding and wireless networks. Section III details the objective of our study. Section IV describes the methodology followed. Section V includes the network coding suitable topologies used in our study. Section VI includes a description of our mathematical model. Section VII includes details on our simulations and the results obtained and Section VIII concludes the paper.

II. BACKGROUND

A. Network coding

1) Introduction: Network coding is a method implemented to utilize network connections more efficiently, by giving intermediate nodes the ability to forward linear combinations of the packets they receive [2]. An example of how this concept works, can be seen in Figure 1a [2]. Node A and node C need to exchange information through a intermediate node B. With the use of a traditional routing protocol, node A would sent a packet to node B, node C would sent a packet to node B and node B would sent each individual packet in two consecutive time slots to their respective locations. With network coding implemented, the third and fourth transmission can be combined into a single transmission which can then be decoded at the destinations. By sending a combined packet in a single time slot, throughput and therefore also the efficiency of the network is enhanced. The cost of having these intelligent nodes, capable of combining and decoding packets, is mainly an increase in the required processing power.

Network coding can be used in different applications for different purposes. It is crucial to remember that the transmission medium is shared (the air) and therefore, in order to prevent collisions, only one node can sent a packet at a given moment.

2) Unicast sessions: Another example, given in [3], of the implementation of network coding, is the application in a network consisting of more than three nodes were two unicast sessions exist simultaneously. Refer to figure 1b. As can be seen in the figure, node 1 needs to transmit a data frame to node 5 and node 5 needs to transmit a data frame to node 2. Due to range limitations of the nodes, transmission of the frames must be done through several intermediate nodes. The transmission paths of the two sessions overlap, creating an opportunity for the implementation of network coding as indicated in the figure.

3) Multicast sessions: Network coding can be implemented in existing wireless ad-hoc networks by finding and using
network coding suitable topologies. An example of such a network coding suitable topology is shown in figure 1c. In this example node A must send a data frame to node D and E, together with node B that must send a data frame to node D and E. Due to communication range limitations, all communication between node A and E as well as communication between node B and D, must be done through node C. Packets arriving at C from A and B, can be combined into a single network coded packet and sent to D and E simultaneously. Thus two multicast sessions, each requiring 2 transmissions, were combined to reduce the total number of transmissions from 4 to 3.

B. Wireless Networks

For the purpose of this paper, we considered and used the IEEE 802.11a/b/g standards. The following physical characteristics play a role in our simulations and calculations:

1) Distance vs. transmission rate: In wireless communication, transmission power is lost as the waves propagate through space. This is a result of a number of phenomena including free-space attenuation, multi-path interference, fading and shadowing. Reliable communication distance and transmission rate is influenced by all of these factors. An increase in attenuation causes reliable communication distance to decrease. When the communication speed is increased, the receiver sensitivity decreases and therefore also the communication distance [4].

2) Interference: The 2.4 GHz and 5 GHz open spectrums are divided into different communication channels which can be used for simultaneous communication between multiple radios. When implementing simultaneous communication methods, adjacent channel interference can occur [5]. Interference in wireless networks can also be caused by other equipment such as cordless or cellular phones as well as microwave ovens. Such sources of interference must be identified and, if possible, removed or the effects lessened for a wireless network to function effectively. Other phenomena such as multi-path interference, which is especially prominent when implementing wireless networks in buildings, must be considered.

3) The hidden node problem: The hidden node problem occurs when two transmitting nodes are close to each other, but can not receive the physical header of transmitted packets, therefore they are unaware of each other’s transmission. A receiver node can be located between these transmitting nodes receiving simultaneous and therefore interfering transmissions from the transmitting nodes. To avoid the occurrence of this situation, the IEEE 802.11 standard relies on a collision avoidance mechanism called the RTS/CTS mechanism [6].

III. Objective

The objective of our study was to define the physical dimensions and optimal node placement of various network topologies that are suitable for the implementation of linear network coding. These definitions are based on the distance at which communication can reliably be executed, the speed at which communication takes place and other aspects including interference. A mathematical model that can calculate these dimensions and node positions, was needed to ensure a theoretical reference, which simulations can be based on. The mathematical model must be able to describe practical situations accurately, therefore we used a practical path-loss model. The resulting topologies had to be simulated and analyzed to evaluate the practicality thereof and ensure that the topologies can be successfully implemented.

IV. Methodology

A. Network coding suitable topology identification

Grobler et al [1] proposed a method for identifying network coding suitable topologies in larger mesh networks. This method was used as the basis of the topology identification stage. This method comprised of the following five steps:

1) Select a Network Coding suitable topology of which the gain and capacity are known (a bow-tie or butterfly topology for instance).
2) Derive the connection matrix of the larger network from a suitable distance vector routing algorithm.
3) Search the larger network matrix for the known topology structure.
4) Implement Network Coding at the appropriate nodes.
5) Re-iterate steps (3) and (4) after a routing update.

B. Distance vs. transmission rate calculations and measurement

Reliable communication distance calculation methods were studied and implemented to determine the theoretical communication distance of each node. This was then used to define the distance at which reliable communication can take place at each different packet transfer speed setting. To simulate the achievable "real world" reliable transmission distance, different empirical and statistical models were analyzed.

C. Effects of interference and possible solutions

Interference from other electro-magnetic wave sources can cause packet loss and induce errors in transmitted data. In severe cases, communication can be lost completely. It is therefore important to be aware of possible interference sources and the impact they may have on communication.

D. Simulate topologies

In the simulation, packets were transferred from source nodes to both destination nodes and the "smart" nodes. At the "smart" nodes, packets were combined by using XOR operations and sent to destination nodes where decoding can then take place. OPNET modeler was used for all simulations. An in depth analysis of the simulated topologies, which implement network coding, was done in terms of load and delay. These results were then used to evaluate the feasibility of utilizing the network coding opportunities identified in larger mesh networks through the methods in [1].

V. NETWORK CODING SUITABLE TOPOLOGY IDENTIFICATION

Certain ad-hoc wireless network topologies are more suitable for the implementation of network coding than others. If a node knows where it is located within the network coding suitable topology, network coding can be implemented without the exchange of control packets. This property has a significant impact on the end-to-end delay of packets. Network coding suitable topologies that have been identified includes the linear topology, shown in figure 1a, the bow-tie topology, shown in figure 1c, the butterfly topology, shown in figure 2 and the hybrid butterfly topology, shown in figure 3. In this paper, we concentrated on the bow-tie, butterfly and hybrid butterfly topologies in our calculations and simulations.

VI. MATHEMATICAL MODEL

A. Introduction

We created a mathematical model to determine the theoretical node positions to be used in our simulations. The nodes were placed to cover the largest communication area possible. This was done by calculating the maximum distance between nodes to the extent that communicating nodes can still exchange information reliably with each other at a given communication distance using one of the IEEE 802.11 a/b/g communication specifications.

To describe practical situations accurately, we used the log-distance path loss model. This model uses empirical and statistical methods to describe attenuation experienced in indoor and outdoor environments. The log-distance path loss model for indoor applications is given by [7]:

\[
Pr(d) = Pr_0 - 10\alpha \log(d) + X_\sigma
\]

where
- \( Pr_0 \) is the signal strength 1 meter from the transmitter,
- \( \alpha \) is known as the path loss exponent,
- \( X_\sigma \) represents a Gaussian random variable with zero mean and standard deviation of \( \sigma \) dB
- \( Pr(d) \) represents the mean (expected) signal strength d meters from the transmitter

In [8], this model was empirically verified with the usage of four Cisco Aironet 1200 wireless access points, operating at 2.4 GHz (IEEE 802.11b). The receiver used was a Cisco Aironet 340 PCMCIA card. Experiments yielded \( \alpha = 4.02 \) and \( \sigma = 7.36 \) dB. The log-distance model is a simplistic model causing the attenuation induced by obstacles between the receiver and transmitter to be aggregated into \( \alpha \).

B. Mathematical model results

1) Stationary nodes: The output of the mathematical model for a bow-tie topology, is shown in figure 4a as an example. To produce this output, the model was given the following inputs:
- Use the bow-tie topology
- Use the IEEE 802.11b standard
- Communicate at 1 Mbps
- With nodes equally spaced from one another
Nodes A and D and nodes B and E must be in each other’s communication range respectively. Cross communication (between nodes A and E and nodes B and D respectively) uses the intermediate node C as a relay node. Observe the case for the transmission of information from the top nodes (nodes A and B) to the bottom nodes (nodes D and E) in a multicast fashion. Transmission of information is conducted through the intermediate node C. Node C (also referred to as the "smart" or "coding" node) can then monitor the packets sent to it and identify network coding opportunities. Coded packets can then be decoded at the destination nodes by using the packet first received from the applicable top node.

Any other identified topology, using any of the IEEE 802.11 a, b or g standards, communicating at any supported communication rate, can be given to the mathematical model as input. The model will then calculate the communication area of each node and display the nodes graphically as in figure 4a.

2) Repositioning of stationary nodes: Nodes are not always equally spaced form each other in reality, therefore we provided the mathematical model with the ability to reposition nodes within any topology. Consider the bottom right node in a bow-tie topology. This node must be able to communicate with the right top and the center (intermediate) node. The area in which this conditions are satisfied, is shown in figure 4b. When nodes are displaced from their default location, the communication circles are also displaced. New communication areas are then formed. When more than one node are displaced, the situation is least complicated when nodes can be moved around in. An example of the situation is shown in figure 4c. The bottom right node is displaced from its default location to another position within its communication area as was shown in figure 4b. This displacement caused the top right node’s communication area to change as shown in figure 4c. Now this top node can be moved around within the new formed communication area.

3) Hidden Nodes: To avoid the occurrence of hidden nodes, all the nodes in the identified topologies must be able to receive the physical header (PLCP header) of a packet being transmitted by all of the nodes individually. When using the IEEE 802.11 standard, the PCLP header must be transmitted at 1 or 2 Mbps while the remaining of the packet can be transmitted at higher transmission rates [9]. The consequence of this property is that the maximum reception range of the PLCP header can be larger than that of the data part.

When nodes are arranged in close proximity to each other in such a way that the transmission of packets can successfully occur at rates of 12 Mbps or above using the IEEE 802.11g standard and 6 Mbps or above using the IEEE 802.11a standard, the hidden node problem will not occur. This can be seen in figure 5, where each node in the topology transmit data at a high transmission rate. The packet header can be successfully transmitted to all the other nodes when sent at 1 Mbps or 2 Mbps as indicated by the larger two circles. However, if any of the topologies grow beyond this point and use one of the lower transmission rates including 11 Mbps and lower when using the IEEE 802.11b standard or 9 Mbps and lower using the IEEE 802.11g standard, hidden nodes may arise and the RTS/CTS mechanism must be enabled to avoid unnecessary collisions.
less database, two different wireless models can be selected. The more advanced model includes all the layers of the OSI protocol stack, while the basic model only includes the physical and data link layer with a packet source and sink combination. This project aims to implement linear network coding on a basic level, therefore the simpler of the two models was used. The main goal of our simulations was to evaluate the practicality of the network coding suitable topologies identified by our mathematical model. All the parameters produced by the mathematical model, including the placement of nodes for a specific transfer rate, were used to create the simulation environment in OPNET Modeler.

B. Discussion

A new “network coding layer” was created and placed on top of the LLC layer within the data link layer. The lower layers were modified to send the necessary destination and source addresses to the new network coding layer. The steps taken by a coding or intermediate node when the node receives a data packet, are shown below. The algorithm used in a decoding or destination node, is similar to the coding algorithm. The difference is that the coding operations are replaced with a decoding operation.

1) Packet arrived, proceed to step 2.
2) Is a network coding address specified for the direction of data flow?
3) If no, forward packet to original destination, else if yes go to step 4.
4) Insert packet into buffer and proceed to step 5.
5) Are there two packets in the buffer?
6) If no, go to step 10, else if yes go to step 7.
7) Is Network coding possible? (Packets are coded if their destination addresses are those of the two destination nodes respectively. The two addresses must differ from one another. The smaller packet is zero padded to the length of the larger packet. This operation is executed separately for each direction of data flow).
8) If yes, code packets, clear buffer and send coded packet to coding address (to destination nodes for decoding), else if no go to step 9.
9) Forward packet in buffer located at 2’nd position to original destination and proceed to step 10.
10) Move packet in buffer 1’st to 2’nd place downwards and proceed to step 11.
11) Wait for another incoming packet with the same direction of data flow and proceed to step 12.
12) Has flush timer expired?
13) If yes, Forward packet in buffer to original destination, else if no go to step 14.
14) Has another packet arrived?
15) If yes, go to step 4 else if no go to step 11.

C. Results

To simulate the simple scenario where voice or video data is transmitted through a network, the source nodes were set to transmit at a constant rate of 100 kb/s in the first set of simulations. The results of implementing network coding in the identified topologies compared to when no network coding techniques are used, are shown in table I. Since data are rarely transmitted at a constant rate in practical networks, we created a second simulation scenario with two variable rate source nodes, sending a data stream with varying load to evaluate the performance of the network coding algorithm under non-ideal conditions. Due to the random nature of the source data, a simulation time of 10 hours was chosen to enable a meaningful comparison between simulations. The results of the second set of simulations are shown in table II. For all the simulations, the transmission of data occurred only in one direction, from the two source nodes to the two destination nodes.

TABLE I

<table>
<thead>
<tr>
<th>Topology</th>
<th>Bow-Tie</th>
<th>Butterfly</th>
<th>Hybrid Butterfly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load:</td>
<td>24.97% decrease</td>
<td>22.11% decrease</td>
<td>22.54% decrease</td>
</tr>
<tr>
<td>Delay:</td>
<td>12.98% increase</td>
<td>10.09% increase</td>
<td>2.54% decrease</td>
</tr>
<tr>
<td>MAD:</td>
<td>12.98% increase</td>
<td>15.55% increase</td>
<td>35.78% increase</td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Topology</th>
<th>Bow-Tie</th>
<th>Butterfly</th>
<th>Hybrid Butterfly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load:</td>
<td>3.79% decrease</td>
<td>33.21% decrease</td>
<td>22.11% decrease</td>
</tr>
<tr>
<td>Delay:</td>
<td>2.75% increase</td>
<td>50.13% increase</td>
<td>33.67% increase</td>
</tr>
<tr>
<td>MAD:</td>
<td>3.24% increase</td>
<td>63.25% increase</td>
<td>42.78% increase</td>
</tr>
</tbody>
</table>

A summary of the results obtained are shown in figure 6. 1) Constant Bit Stream: The decrease of the global load is quite significant for all the topologies when a constant bit stream is used as a source. This is because of the constant and synchronized manner in which data arrives at the coding and destination nodes. The amount of transmissions are reduced, translating to a reduction in global network load for the same throughput. The decrease in global load is the most when network coding is implemented in the butterfly topology. This can be explained by taking into account that the total number of transmissions are reduced by two when network coding is implemented in the butterfly topology, while the total number of transmissions is only reduced by one in the bow-tie and hybrid butterfly topologies.

The global delay of packets successfully received by nodes and the media access delay of packets, were slightly higher with network coding enabled than with it disabled in the bow-tie topology. Data packets must wait in a buffer at the “smart” node for another network coding suitable packet to arrive before network coding can take place. The delay increase is however lower than what is expected because of the reduced number of packets that must be sent to the receiver nodes, which translates into reduced contention for the shared communication medium.

For the butterfly topology, the global delay of packets successfully received by nodes and the media access delay of packets, were lower with network coding enabled than with it disabled. This can be explained by taking into account that the total number of transmissions is reduced by two,
thereby also reducing the time packets wait to access the shared communication medium.

For the hybrid butterfly topology, the global delay of packets successfully received by nodes were slightly lower with network coding enabled than with it disabled, but the media access delay increase was much higher than what was experienced in the bow-tie topology. As with the bow-tie topology, the number of transmissions is reduced only by one, thus the delay decrease will not be as significant as in the butterfly topology.

2) Variable Bit stream: The same trends of a reduced global load and increased global and media access delays were seen when a variable bit stream was used in comparison to the usage of a constant bit stream as source. We can therefore derive that the usage of our network coding scheme can be of value when implemented in networks where very random bursts of data at different rates occur. The only drawback is that the usage of this scheme in such networks will induce extra delay, making the evaluated network coding scheme suitable for networks where the reduced transmissions and global load are more important than the delay induced on packets.

VIII. CONCLUSION

In this paper, we presented work regarding boundary identification, optimal node placement and the practicality of implementing network coding in the identified topologies in wireless ad-hoc networks. Results suggest that the implementation of network coding will decrease the overall network load with the cost of a higher network end-to-end delay of transmitted packets. We can therefore derive that this network coding scheme may produce better overall performance when implemented in sensor networks or highly congested Ad-Hoc networks. Future work will include the implementation of our network coding system on a 6 node test bed using Click Modular Router in the Linux environment.

REFERENCES


Johan-Henry Böning is currently pursuing a M.Eng degree in Computer Engineering at the North-West University. He received his B.Eng degree in Computer and Electronic Engineering (cum laude) in 2008 from the North-West University.