Predictive Modeling of Network Wide Broadcasting Protocols for Mobile Ad Hoc Networks*†‡

Brad Williams Dinesh Mehta Tracy Camp

William Navidi

Department of Mathematical and Computer Sciences

Colorado School of Mines

Golden, Colorado 80401

April 20th, 2003

Abstract

Network wide broadcasting in Mobile Ad hoc NETworks (MANETs) provides important control and route establishment functionality for a number of unicast and multicast protocols. Simulation is frequently used to assess a broadcast protocol's performance because the complicated nature of MANETs often makes explicit modeling difficult if not impossible. With application of some simplifying assumptions, however, explicit modeling can be performed. We present models to predict retransmission frequency of three network wide broadcasting protocols using a probability based approach. Our models are able to predict protocol performance within 5% of simulation results when we only consider network layer functionality.

*This work supported in part by NSF Grants ANI-0073699 and ANI-0208352.

†Research group's URL is *http://toilers.mines.edu*.

‡Reference for this manuscript: Technical Report MCS-03-05, The Colorado School of Mines, April 2003.

1 Introduction

Mobile Ad hoc NETworks (MANETs) are wireless networks which are characterized by dynamic topologies and no fixed infrastructure. Each node in a MANET is a computer that may be required to act as both a host and a router and, as such, may be required to forward packets between nodes which cannot directly communicate with one another.

Network wide broadcasting, simply referred to as “broadcasting” for the remainder of this paper, is the process in which one node sends a packet to all other nodes in the MANET. Broadcasting may be used by a node to send data to all other nodes in the network *or* may be used by MANET unicast or multicast routing protocols to disseminate control information. For example, many unicast routing protocols such as Dynamic Source Routing (DSR) [1, 2], Ad hoc On demand Distance Vector (AODV) [3, 4], Zone Routing Protocol (ZRP) [5, 6, 7, 8], and Location Aided Routing (LAR) [9] use broadcasting or a derivation of it to establish routes.

All of these unicast routing protocols use a generally inefficient form of broadcasting called Simple Flooding. In Simple Flooding each node rebroadcasts each packet exactly one time. In dense networks, this wastes bandwidth and node resources. Recently, a number of research groups have proposed more efficient broadcasting protocols based on distributed and hierarchical methodologies. Many, if not all, existing distributed network wide broadcast protocols have been summarized and categorized in [10].

Categories proposed in [10] include Simple Flooding, Probability-Based Methods, Area-Based Methods and Neighbor Knowledge Methods. Simple Flooding is described in the proceeding paragraph. Nodes using Probability-Based protocols decide to rebroadcast according to a set probability or a simple conditional event which relates to the probability of reaching additional neighbors. Area-Based Methods require knowledge of sender node locations and estimate weather a rebroadcast would reach a significant amount of additional coverage area. Finally, Neighbor Knowledge Methods require the use of “Hello” type packets so that nodes have explicit data regarding their neighborhood topology; the nodes use the neighbor data in deciding whether to rebroadcast a packet. In this paper, we derive models for all protocols in the categories of Simple Flooding, Probability-Based Methods, and Area-Based Methods.

In [10], as well as in the original publications in which the protocols were proposed, a simulation evaluation was performed to gauge protocol applicability. Simulation comparison is particularly useful for comparing protocols under "real" network conditions in which interactions between various communication layers in a network topology affect protocol performance. However, simulator development and data collection may be prone to errors if the simulators are not carefully tested. Clearly, if explicit models that predict protocol performance are developed, they could be used to validate simulation results, or used in lieu of simulation. In this paper, we address this goal by proposing explicit models for five broadcasting protocols.

Unfortunately, explicit modeling is not necessarily an easy task. The complex interactions between nodes in MANETs and the complex interactions between communication layers within a single node are difficult to quantify. As such, we simplify the network to allow us to model protocols more easily. This means that our models may not be valid for all network conditions; however, they are appropriate for certain networks and the models do represent a foundation from which more robust representations can be built. In addition, results from the models can still be used to validate developed simulation.

In Section 2 we describe the protocols which we model. Section 3 contains a description of our simplifying assumptions, describes the simulation environment used to validate our models and provides a general model from which we derive specific protocol models. These specific protocol models are then derived and tested in Section 4. Finally, conclusions and future research are described in Section 5.

2 Protocol Details

In this section we describe the operation of the five broadcast protocols which we model analytically: Simple Flooding, the Probabilistic scheme, the Counter-Based scheme, the Distance-Based scheme and the Location-Based scheme. The following descriptions are taken from the review of broadcast protocols in [10]. In [10], the Probabilistic scheme and the Counter-Based scheme are described in the Probability Based Methods category; the Distance-Based scheme and the Location-Based scheme are described in the Area Based Methods category.

2.1 Simple Flooding

The algorithm for Simple Flooding [11, 12] starts with a source node broadcasting a packet to all neighbors. Each of those neighbors in turn rebroadcasts the packet the first time it receives the packet. Redundant packets are simply dropped. This behavior continues until all reachable network nodes have received and rebroadcast the packet once.

2.2 Probabilistic Scheme

The Probabilistic scheme from [13] is similar to Flooding, except that nodes only rebroadcast with a predetermined probability. In dense networks multiple nodes share similar transmission coverages. Thus, randomly having some nodes not rebroadcast saves node and network resources without harming delivery effectiveness. In sparse networks, there is much less shared coverage; thus, nodes will not receive all the broadcast packets with the Probabilistic scheme unless the probability parameter is high. When the probability is 100%, this scheme is identical to Simple Flooding.

2.3 Counter-Based Scheme

Ni et al [13] show an inverse relationship between the number of times a packet is received at a node and the probability of that node being able to reach additional area on a rebroadcast. This result is the basis of their Counter-Based scheme. Upon reception of a previously unseen packet, the node initiates a counter with a value of one and sets a Random Assessment Delay (RAD), which is a timer set to a random number of seconds in a given interval. During the RAD, the counter is incremented by one for each redundant packet received. If the counter is less than a threshold value when the RAD expires, the packet is rebroadcast. Otherwise, the packet is simply dropped.

2.4 Distance-Based Scheme

Nodes using the Distance-Based scheme [13] compare the distance between themselves and each neighbor node that has previously rebroadcast a given packet. Upon reception of a previously unseen packet, a RAD is initiated and redundant packets are cached. When the RAD expires, all source node locations are examined to see if any node is closer than a threshold distance value. If true, the node does not rebroadcast. This protocol requires knowledge of neighbor locations. Signal strength could be used to gauge the distance to the source of a received packet. Alternatively, if a Global Positioning System (GPS) is available, nodes could include their location information in each packet transmitted.

2.5 Location-Based Scheme

The Location-Based scheme [13] uses a more precise estimation of expected additional coverage in the decision to rebroadcast. In this method, each node must have the means to determine its own location, e.g. via GPS; rebroadcasting nodes add their locations to the header of the packet.

When a node initially receives a packet, it notes the location of the sender and calculates the additional coverage area obtainable were it to rebroadcast. If the additional area is less than a threshold value, the node will not rebroadcast, and all future receptions of the same packet will be ignored. Otherwise, the node assigns a RAD before delivery. If the node receives a redundant packet during the RAD, it recalculates the additional coverage area and compares that value to the threshold. The area calculation and threshold comparison occur with all redundant broadcasts received until the packet reaches either its scheduled send time or is dropped.

3 Model Basis

3.1 Simplifying Assumptions

Our goal is to predict the retransmission frequency of nodes, or percentage of nodes that rebroadcast any given broadcast packet, in networks in which node density is the only factor. Specifically, our approach is to sum the individual probability of each condition that would induce a node to retransmit a packet. Individual probabilities are determined assuming random events, including the random placement of nodes in the network. Therefore, our models are not appropriate for networks in which nodes are positioned in some systematic non-random manner. We only consider fully connected networks in which no partitions exist. Our mathematical models also assume a random traffic pattern, that is the broadcast packet source is random and uniformly chosen from the set of all network nodes.

As stated in Section 1, interactions between nodes and interactions between communication layers within a node are complex and difficult to model. For the purpose of this paper, we only consider the broadcast protocol interaction between nodes and disregard the underlying Medium Access Control (MAC) methodology as well as the dynamic nature of mobile ad hoc networks. Thus, we consider these protocols under idealized conditions of a Null MAC and a static topology.

While these assumptions limit the applicability of our models, our models are appropriate for certain viable networks. For example, consider a MANET with low data traffic, which has no congestion effects such as packet collisions, and slow moving nodes. In addition, our models represent a foundation for further model development. In review, our core simplifying assumptions are:

1. The nodes are randomly and uniformly distributed throughout the network area.
2. The network is fully connected (there are no partitions).
3. Source nodes of broadcast packets are randomly chosen from the set of all network nodes.
4. The MAC layer does not significantly interfere with network layer protocol performance. In other words, the network is not congested with packet traffic.
5. Nodes are static or move slowly.

3.2 Simulation Environment

In order to validate our models, we choose to compare model-predicted outcomes to simulation results. We chose simulation parameters to correlate with the simplifying assumptions listed in Section 3.1. We use NS-2 [14], a network simulator, to simulate node behavior for each protocol studied in this paper.

Network topologies for use in the simulation were generated by placing nodes in the 2-D simulation area via a random and uniform distribution. After each topology was generated, we analyzed it to ensure that the network was fully connected by constructing a multihop connectivity matrix from a simple adjacency matrix and verifying that a path existed between all node pairs. We do not allow the nodes in our simulation to move in order to ensure the network does not partition.

The packet traffic pattern in our simulation is random; i.e., for each packet transmitted, the source is chosen uniformly from the set of all network nodes. The interval between packet originations is two seconds. To ensure congestion does not affect the simulation results, we use a Null

# Nodes in 550m x 550m Area	# Nodes in 350m x 350m Area	Avg # of Neighbors
50	20	5
74	30	7.6
99	40	9.8
123	50	12.5
148	60	15.3
173	70	17.6
222	90	22.9
272	110	29.3

Table 1: Average Number of Neighbors for Different Numbers of Nodes in the 350m x 350m Center Area

MAC in our simulations. That is, instead of using the standard 802.11b MAC component provided in NS-2, we substitute it with a Null MAC object. The Null MAC object distributes outgoing packets to each neighbor's incoming packet queue; there are no delay or packet collision events. Packet payload size is set to 64 bytes and all node interface queues (IFQs) have a length of 50 packets. However, these parameters do not affect simulation outcomes due to the use of the Null MAC.

The network size studied was 350m x 350m. We note that a node at the perimeter of the network may have fewer than the average number of neighbors. Therefore, perimeter nodes will have a different probability of rebroadcasting than nodes at the center of the network. To avoid modeling this difference in behavior, we actually simulate a 550m x 550m network with a proportionally higher number of nodes and only collect statistics for nodes in the 350m x 350m center area. Authors of [15] use a similar technique.

In order to limit simulation length and iterative solution time for the models, the transmission range for all nodes was set at 100 meters. The number of nodes in the network is varied between 20 and 110. Table 1 shows the average number of neighbors for nodes in the center 350m x 350m

Simulation Parameter	Value
Simulator	NS-2 (1b7a)
Data Packet Size	64 bytes payload
Node Max. IFQ Length	50
Packet Traffic Pattern	Packet originated every two seconds
Network Area	350m x 350m
Number of Nodes	Varied
Node Tx Distance	100 meters
# of Trials	10
Confidence Interval	95%

Table 2: Simulation Parameters Common to All Studies

area for each scenario. Each simulation result presented is the mean and 95% confidence interval for 10 different topologies. Table 2 reviews our simulation parameters.

3.3 General Model

We model each broadcasting protocol with Equation 1, where R_{node} is the expected number of retransmissions (or rebroadcasts) per node per time period, S is the number of broadcast packets originated during the time period, and P is the probability that a given node will rebroadcast a given packet:

$$R_{node} = P * S. \tag{1}$$

We note that Equation 1 is valid without our assumptions specified in Section 3.1.

4 Model Presentation

Section 4.1 discusses Equation 1 in relation to Simple Flooding and the Probabilistic scheme. While the equation is trivial to define in these two cases, the exercise provides insight into our general model. Sections 4.2, 4.3, and 4.4 describe our work in mathematically predicting P, from Equation 1, as a function of neighbor density for the Counter-Based, Distance-Based and Location-Based models, respectively. Our approach is to first calculate the individual probabilities for cases which cause a node to rebroadcast. The summation of these individual probabilities is then P.

4.1 Models for the Simple Flooding and Probabilistic Schemes

In Simple Flooding, each node retransmits each broadcast packet once. In other words, in Equation 1, P = 1.0. Thus, the number of retransmissions per node per time period is

$$R_{node} = S.$$

Similarly, in the Probabilistic model, P is the global protocol parameter which each node uses in the decision to rebroadcast. Suppose this parameter is denoted by the constant K, where $0 < K < 1$. The number of retransmissions per node per time period for the Probabilistic model is then

$$R_{node} = K * S.$$

For Simple Flooding and the Probabilistic models, node behavior is independent of neighbor density. More efficient broadcasting schemes intrinsically modify their behavior as an inverse function of neighbor density. In sparsely populated networks, nodes using more efficient broadcasting schemes may retransmit all originated packets (i.e., P = 1.0); as neighbor density increases, P is reduced. The following three subsections present our models for three more efficient broadcasting protocols.

4.2 Model for the Counter-Based Scheme

In the Counter-Based scheme, a node v will only rebroadcast if it receives fewer than a threshold (T) number of redundant packets before its RAD expires. Each node keeps track of redundant packets with a counter; the counter is incremented by 1 for each redundant packet received before the RAD expires.

For a random node v to receive a redundant packet (and increment its counter) from a random node u, three events must occur:

A. Node u must be a neighbor of node v.

B. Node u must transmit the packet.

C. Node u must transmit the packet before v's RAD timer expires.

The probability Q that node v increments its counter (i.e., events A, B, and C have all occurred) is equal to $P(A \cap B \cap C)$. We define the individual probabilities, $P(A)$, $P(B|A)$, and $P(C|B \cap A)$ next.

To determine the probability that u and v are neighbors, we rely on our assumption that nodes are random and uniformly distributed in the network. If we assume that each node has a transmission radius of R and that all network nodes exist within an area A_{net}, then:

$$P(A) = \frac{\pi R^2}{A_{net}}. \tag{2}$$

In our Counter-Based model, we assume that each node has the same probability of rebroadcasting equal to P. Thus,

$$P(B|A) = P. \tag{3}$$

Finally, since each node assigns a random assessment delay from a uniform and random distribution, and since we assume that all of v's neighbors receive the packet at the same time, then each rebroadcasting neighbor has a 50% probability of rebroadcasting before v. Thus,

$$P(C|A \cap B) = \frac{1}{2}. \tag{4}$$

With Equations 2, 3, and 4, we determine the value for Q as:

$$Q = P(A \cap B \cap C) = \frac{\pi R^2 P}{2A_{net}}. \tag{5}$$

Equation 5 describes the probability that node v will increment its counter based on the actions of a single other node u. Our goal, however is to determine the probability P that node v rebroadcasts the packet. Recall that node v will rebroadcast if it receives less than the threshold T redundant packets before its RAD expires. Therefore, we can determine P by summing the probability that between 1 and $T-1$ packets are received by v such that v increments its counter in response to those receptions. The probability that exactly i packets cause node v to update its counter is given by the expression $\binom{N-2}{i} Q^i (1-Q)^{N-2-i}$. Note that $N-2$ is used rather than N because we exclude node v and the node from which it first receives the packet. Thus,

$$\begin{aligned} P &= \sum_{i=0}^{T-2} \binom{N-2}{i} Q^i (1-Q)^{N-2-i}. \\ &= \sum_{i=0}^{T-2} \binom{N-2}{i} \left(\frac{\pi R^2 P}{2A_{net}}\right)^i \left(1 - \frac{\pi R^2 P}{2A_{net}}\right)^{N-2-i}. \end{aligned} \tag{6}$$

We note that our summation range is from $i = 0$ to $i = T-2$. Since we only consider completely connected networks with no packets lost, node v is guaranteed to receive the packet at least one time. Therefore, we remove this event from our Counter-Based model and consider the probability that v receives between 0 and $T-2$ additional packets.

We used Matlab software to iteratively find the value of P satisfying Equation 6. In order to test our Counter-Based model and validate our simulation, we compared the Matlab results to a simulation of the Counter-Based scheme in a 350m x 350m network (with a 100m perimeter around the 350m x 350m network) containing a varying number of nodes. The simulation was performed with the parameters discussed in Section 3.2. The threshold value used by the Counter-Based scheme in both the model and the simulation was $T = 3$.

Figure 1 demonstrates that our mathematical Counter-Based model, given in Equation 6, compared to our simulation results, provides reasonably close estimates of the probability that a given

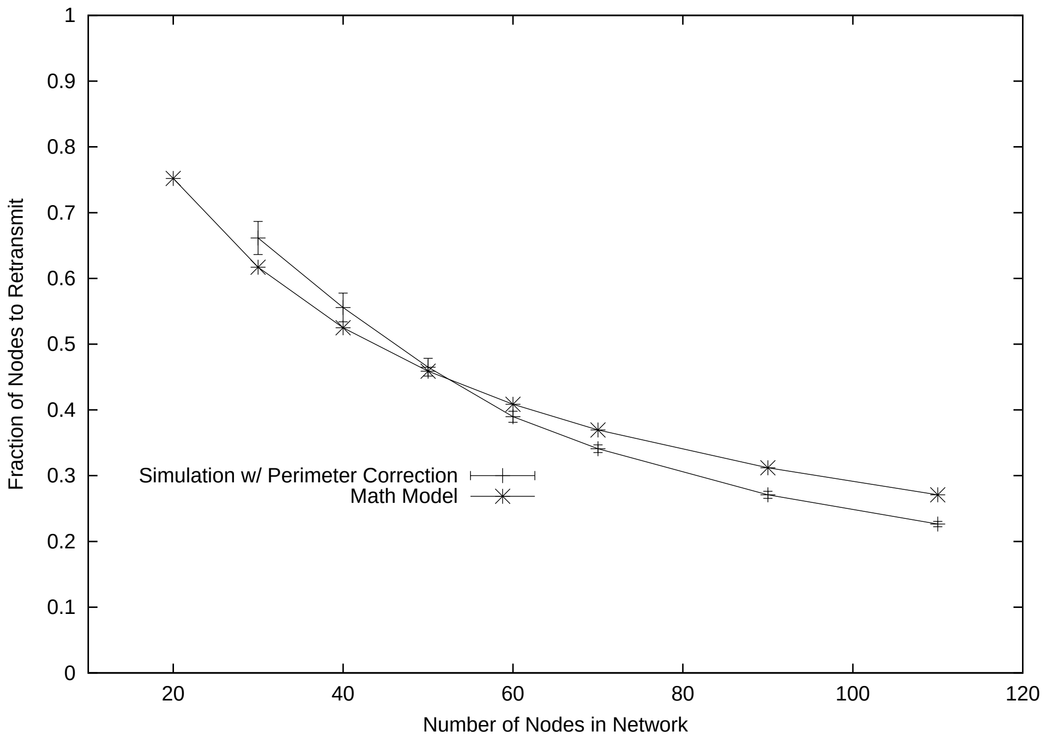

Figure 1: Math Model Prediction of Retransmission Probability for the Counter-Based Scheme

node rebroadcasts. Our Counter-Based model, compared to our simulation results, slightly under predicts the probability of retransmission in sparse networks and slightly over predicts the probability of retransmission in dense networks. We conclude that the average number of neighbors (which is dependent on the total number of network nodes, transmission range, and network area) as well as the threshold value strongly affect the retransmission frequency of nodes using the Counter-Based scheme. We believe the results from our Counter-Based model do not exactly match the results from our simulation due to the assumptions we had to make in the development of the model.

First, we assumed that each of v's neighbors received the packet at the same time as v; thus, v has a 50% chance of rebroadcasting before each neighbor. In our simulation, we expect a packet to be received by some neighbors before others which affects the overall retransmission frequency. Secondly, we assumed that each neighbor node has an independent probability (P) of rebroadcasting, which matches the overall probability of rebroadcasting in the network. In our simulation, we expect the probabilities of neighbors rebroadcasting to be dependent on one another and this dependency affects the overall retransmission probability.

4.3 Model for the Distance-Based Scheme

In the Distance-Based scheme, a node v will not rebroadcast if it receives its initial packet from a source node s that is within a threshold distance D. Consider an annulus area centered at v, with the radius of the inner circle equal to D and the radius of the outer circle equal to the transmission distance R of the node's radio. If node s is within the annulus area (i.e., outside the circle with radius D and inside the circle with radius R), then v will start a RAD and wait for redundant packets. During the RAD, if any redundant packet is received from a node u within a distance of D to v, then v will not rebroadcast.

We define the following events to determine the probability of v receiving a redundant packet from a random node u, such that u is within a distance of D to v, before its RAD expires:

A. Node u is within distance D of node v.

B. Node u transmits the packet.

C. Node u transmits the packet before v's RAD timer expires.

The probability that node v receives this kind of packet is equal to $P(A \cap B \cap C)$. We define the individual probabilities, $P(A)$, $P(B|A)$, and $P(C|B \cap A)$ next.

$$P(A) = \frac{\pi D^2}{A_{net}}. \tag{7}$$

As in the Counter-Based model, we assume that each node has the same probability of rebroadcasting equal to P. Thus,

$$P(B|A) = P. \tag{8}$$

As in the Counter-Based model, since each node assigns a random assessment delay from a uniform and random distribution, and since we assume that all of v's neighbors receive the packet at the same time, then each rebroadcasting neighbor has a 50% probability of rebroadcasting before v.

$$P(C|A \cap B) = \frac{1}{2}. \tag{9}$$

With Equations 7, 8, and 9, we determine the value for $P(A \cap B \cap C)$ as:

$$P(A \cap B \cap C) = \frac{\pi D^2 P}{2A_{net}}. \tag{10}$$

Equation 10 describes the probability that node v will receive a packet from a random node that will cause it to *not* rebroadcast. Our goal is to derive the probability P that node v rebroadcasts the packet. This happens with probability $((1 - P(A \cap B \cap C))^{N-2}$. Again, note that $N-2$ is used rather than N because we exclude both node v and the node, which is in the annulus area, that initiated v's RAD. In other words, none of the $N-2$ random nodes should prevent a rebroadcast by node v. We also require the source node s be outside the inner (threshold) circle. Since s is known to lie inside the outer circle, this has probability $(1 - D^2/R^2)$. Thus, the probability that a given node will rebroadcast is

$$P = \left(1 - \frac{D^2}{R^2}\right)\left(1 - \frac{\pi D^2 P}{2A_{net}}\right)^{N-2}. \tag{11}$$

We used a root finding function in Matlab to find the value of P satisfying Equation 11. We then compared the results of our mathematical Distance-Based model with a simulation of the Distance-Based scheme. We used the same simulation conditions discussed in Section 3.2. A threshold value of 35 meters was utilized for both the Distance-Based model and simulation.

Figure 2 illustrates that our Distance-Based model, compared to our simulation results, reasonably predicts retransmission frequency. Again, we conclude that the average number of neighbors and the threshold distance value strongly affect retransmission frequency. Our Distance-Based model is a better predictor for all node densities than the Counter-Based model. In fact, the confidence intervals from our simulations include the results from our mathematical Distance-Based

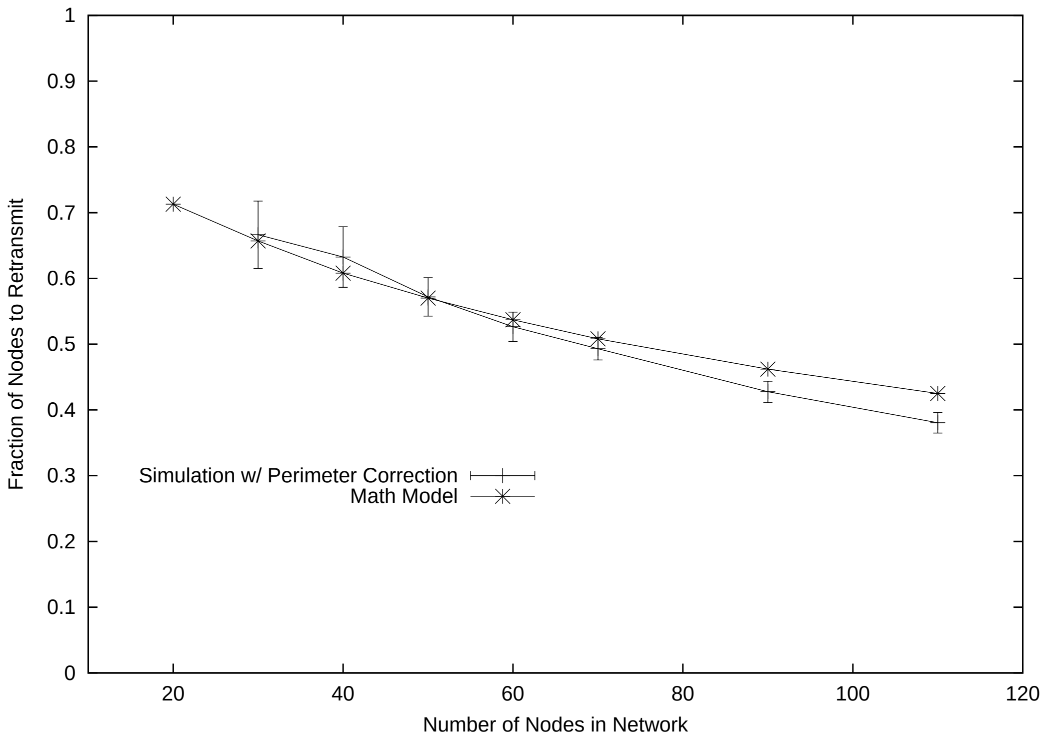

Figure 2: Math Model Prediction of Retransmission Probability for the Distance-Based Scheme

model for all cases except dense networks; in dense networks, our Distance-Based model slightly over predicts our simulation results.

4.4 Model for the Location-Based Scheme

Recall that a node using the Location-Based scheme attempts to quantify the expected additional area covered if it were to rebroadcast. The node uses knowledge of the locations of neighbors which have previously rebroadcast a packet in its decision to rebroadcast. The original presentation of the Location-Based scheme [13] lacks specific information about implementation details. Thus, we assume certain protocol details which we now discuss.

An exact calculation of additional reachable area is an algorithmically intensive one because it requires calculating areas of overlapping circles. Following the paper that first proposed the Location-Based scheme, we use an approximation method for this calculation. Our implementation

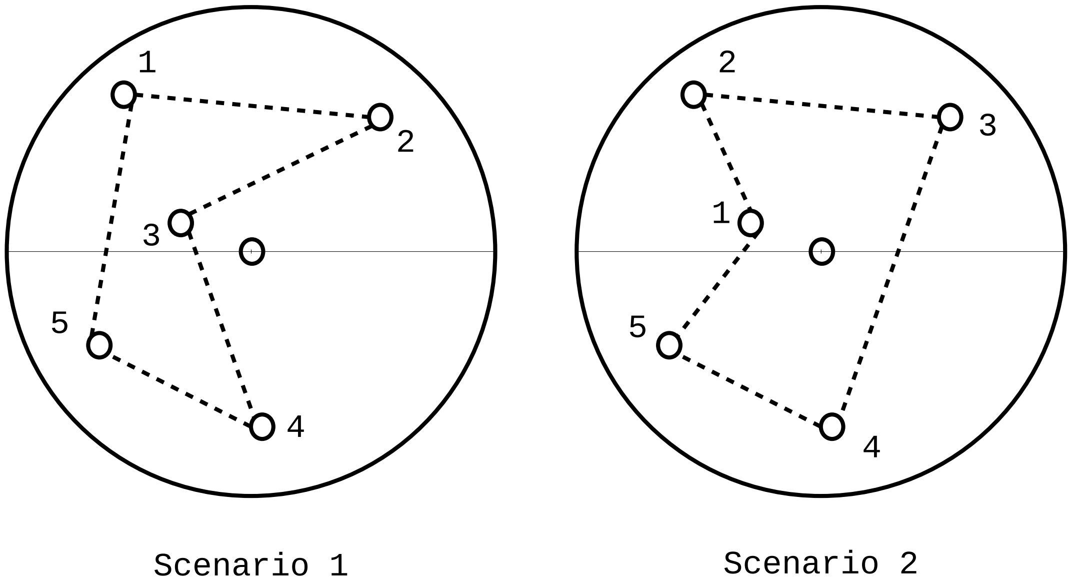

Figure 3: Scenarios Showing Ordering of Vertexes Based on Reception Order and Angularly from a Reference Vector

uses a two case approach. The node first checks if any sender of redundant packets received within the RAD time is less than a threshold distance away. If so, the node does not rebroadcast the packet. In this respect, the node is essentially using the Distance-Based scheme. Given the scenario where all redundant packets are received from neighbors that are farther away than the threshold distance, and given that the number of redundant packets is greater than or equal to three, our implementation uses a polygon approximation method to determine whether to rebroadcast. If the assessing node is inside the polygon formed with previously rebroadcasting neighbors as vertexes, then the assessing node is unlikely to reach sufficient additional area to justify a rebroadcast and the node simply drops the packet. Obviously, a polygon cannot be formed if less than three redundant packets are received. Therefore, if only one or two redundant packets are received within the RAD, then the node simply rebroadcasts the packet assuming the source nodes of those packets are farther away than the threshold distance.

We note that the ordering of the vertexes that is used to form the polygon is significant. If we connect the vertexes of the polygon in the order in which the assessing node receives the redundant

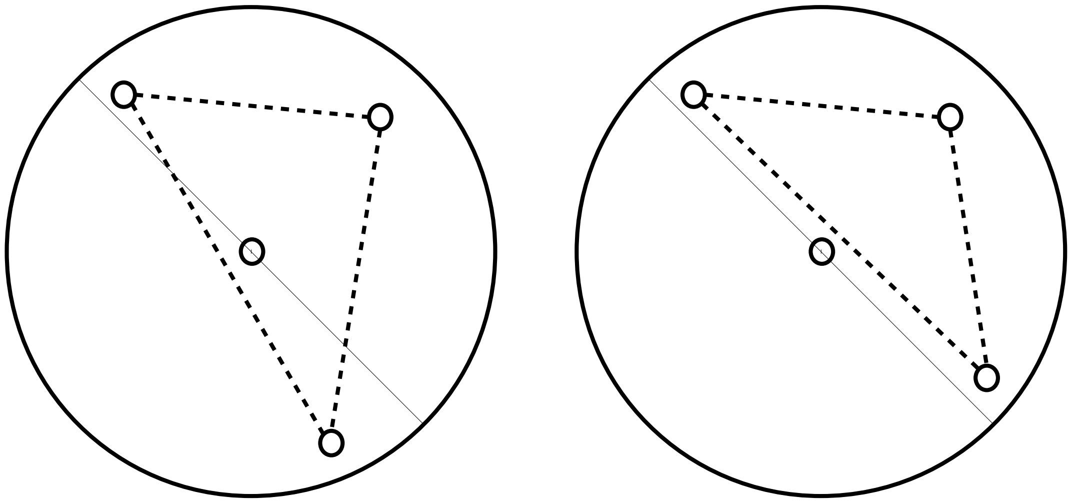

Figure 4: Scenarios Showing Assessing Node Both Inside and Outside of Polygon

packets, then polygons like that in Scenario 1 of Figure 3 may exist. In Scenario 1, the polygon does not contain the center of the circle (i.e., the assessing node). However, if the vertexes are ordered based on the angle formed from the center of the circle to the vertex and an arbitrary fixed reference vector emanating from the center of the circle, then the polygon may contain the assessing node. This example is illustrated in Scenario 2 of Figure 3. In this scenario, we used a clockwise ordering with the negative abscissa as the reference vector. We chose to order the vertexes angularly (as in Scenario 2) in our implementation of the Location-Based scheme because it simplifies the development of a predictive Location-Based model while maintaining the intent of the protocol, i.e., to assess additional available transmission area.

We next derive two propositions that will facilitate the development of an equation to estimate the retransmission probability in the Location-Based scheme. Figure 4 shows two scenarios in which three neighbors have previously rebroadcast a packet. In the first scenario, the assessing node is within the polygon formed by these neighbors. In the second scenario, the assessing node is not. These particular scenarios are chosen to illustrate Proposition 1.

Proposition 1 A set of points all occur within any semi-circular portion of the assessing node's transmission range *if and only if* the polygon formed by the angularly sorted set of points does *not* contain the center of the circle.

Proof: First, we show that if the polygon contains the center of the circle, then no empty semi-circular area exists. Suppose that this is false and there is an empty semi-circular area even though the polygon contains the center. Assume that the coordinate axes are oriented so that the center is at the origin and that the empty semi-circular area consists of the diameter that coincides with the *x*-axis and the portion of the circle below the *x*-axis. Then, all of the polygon vertexes lie strictly above the *x*-axis and a segment joining any pair of vertexes must lie strictly above the *x*-axis as well. It is a contradiction that such a polygon could contain the center of the circle because no polygon edges pass through or below the *x*-axis. Thus, if the polygon contains the center, no empty semi-circle portion of the circle can exist.

Next, we show that if there is no empty semi-circle, then the polygon formed by joining vertexes in an angularly ordered sequence must contain the center of the circle. Assume that the coordinate axes are oriented so that the center of the circle is at the origin *O*. A segment is said to cover an angle Θ if a semi-infinite ray from the origin at an angle of Θ with the *x*-axis encounters the segment. If the polygon does not contain the center, then there must be an angle α which is not covered by a segment (because polygons are formed by a sequence of angularly sorted vertexes). Consider the two polygon vertexes on each side of the radial line corresponding to α that are the closest angularly. Consider the sector containing α that is bounded by radial lines through the two polygon vertexes. This sector does not contain any polygon vertexes and must subtend an angle greater than 180 degrees (otherwise the segment would have covered α). This is a contradiction. Thus, if no empty semi-circular area exists, the polygon formed must contain the center of the circle. △

In Scenario 2 of Figure 4, all previously rebroadcasting neighbor nodes are contained within

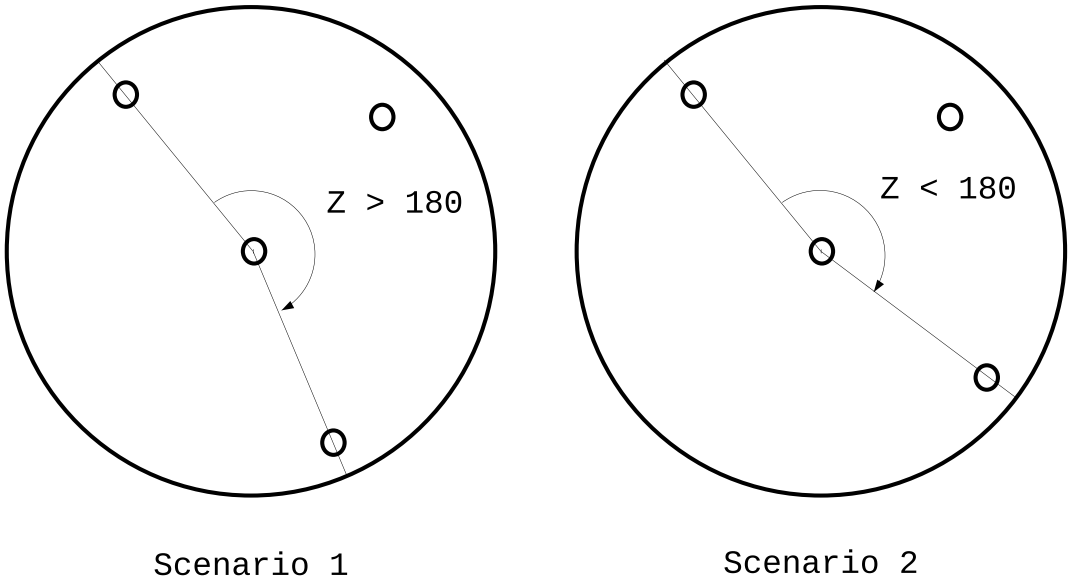

Figure 5: Vectors With Minimum Angle, Z, Such that All B Neighbors are Contained Within those Vectors

a semi-circular area of the assessing node's transmission distance. In other words, it is possible to draw two vectors from the center of the circle through two neighbor node locations such that all previously rebroadcasting nodes lie between those vectors and the internal angle, Z, formed by them is less than 180 degrees. Figure 5, illustrates this definition for the scenarios in Figure 4.

With Proposition 1, we now define Q, the probability the assessing node will rebroadcast (i.e., the assessing node is not in the polygon), as equal to the probability that all previously rebroadcasting nodes lie within a semi-circular area of the assessing node's transmission distance. We derive Q via Proposition 2 next.

Proposition 2 Given x points that are uniformly and randomly distributed within a circle, the probability Q that all x points lie inside any semi-circular portion of the circle is $x * 0.5^{(x-1)}$.

Proof: Given x points randomly and uniformly placed inside a circle, we define Q by assessing each point individually. Starting with point 1, the probability that point 2 is within a 180 degree ($Z <= 180$) clockwise angle is 0.5. This is also true for points 3,4,...x. Therefore, the probability that all $x-1$ points lie within a clockwise semi-circular area starting at point 1 is 0.5^{x-1}. We repeat this exercise starting with points 2,3,4,...x to account for the possibility that a clockwise semi-circular area starting at these points would contain all other points. The probability for each case is 0.5^{x-1}. Note that these cases are mutually exclusive and together give us the probability that all x points lie within any semicircular portion of a circle. Therefore the probability that all points lie within *some* semi-circular area is:

$$Q = x * 0.5^{x-1} \ . \ \triangle \tag{12}$$

Now we consider Equation 12 for the cases when $x = 1$ and $x = 2$. For both of these situations, Q resolves to 1.0. In terms of the semi-circular calculation, when there is only one or two points, then those points will always be within a semi-circular area. In terms of retransmission determination, the nodes that receive only one or two redundant broadcast packets from nodes more that a threshold distance away will always rebroadcast (the behavior of the Location-Based scheme which we are trying to model).

Let s be the node from which v first receives the packet. Let S_Y be the set of neighbors of v that broadcast after s and before v's RAD has expired. Let $Y = |S_Y|$. Then v will rebroadcast if:

A: The points in $S_Y \cup \{s\}$ are outside v's threshold area.

B: The ordered polygon formed by the points in $S_Y \cup \{s\}$ does not contain v.

First, suppose $Y = i$. Then, the probability these i points and s are within the annulus area (i.e., outside the circle with radius D and inside the circle with radius R) is

$$P(A|Y = i) = \left(1 - \frac{D^2}{R^2}\right)^{i+1} . \tag{13}$$

Second, based on Proposition 2,

$$P(B|Y = i) = (i+1)(0.5)^i. \tag{14}$$

To determine $Y = i$ (see the discussion of Equation 6 for details), we have

$$P(Y = i) = \binom{N-2}{i} \left(\frac{\pi R^2 P}{2A_{net}}\right)^i \left(1 - \frac{\pi R^2 P}{2A_{net}}\right)^{N-2-i}. \tag{15}$$

Thus,

$$\begin{aligned}
P &= P(v \text{ rebroadcasts}) \\
&= \sum_{i=0}^{N-2} P(Y=i)\, P(A \cap B | Y = i) \\
&= \sum_{i=0}^{N-2} P(Y=i)\, P(A|Y=i)\, P(B|Y=i) \\
&= \sum_{i=0}^{N-2} \binom{N-2}{i} \left(\frac{\pi R^2 P}{2A_{net}}\right)^i \left(1 - \frac{\pi R^2 P}{2A_{net}}\right)^{N-2-i} \left(1 - \frac{D^2}{R^2}\right)^{i+1} (i+1)(0.5)^i \\
&= \left(1 - \frac{D^2}{R^2}\right) \sum_{i=0}^{N-2} \binom{N-2}{i} \left(\frac{\pi (R^2 - D^2) P}{2A_{net}}\right)^i \left(1 - \frac{\pi R^2 P}{2A_{net}}\right)^{N-2-i} (i+1)(0.5)^i. \quad (16)
\end{aligned}$$

We used a root finding function in Matlab to find the value of P satisfying Equation 16. We then compared the results of our mathematical Location-Based model to a simulation of the Location-Based scheme based on the parameters from Section 3.2. A 35 meter threshold value was evaluated. Figure 6 shows that our Location-Based model under predicts the number of retransmissions at low node density and over predicts the number of retransmissions at high node density. We note that Equation 15 in the Location-Based model is quite similar to Equation 6 in the Counter-Based model. Thus, it is not surprising that the mathematical model performance versus simulation of the Location-Based scheme follows the pattern seen with the Counter-Based scheme. We also note that our Location-Based model predicts simulation results within 5% for each scenario.

We state previously that one source of inaccuracy in our models is that we do not account for a packet propagation model; instead, we assume each neighbor receives that packet at the same time as the source. In reality, there is a higher probability that a subset of neighbors on one side of the assessing node will rebroadcast before the assessing node; thus, a higher probability that these

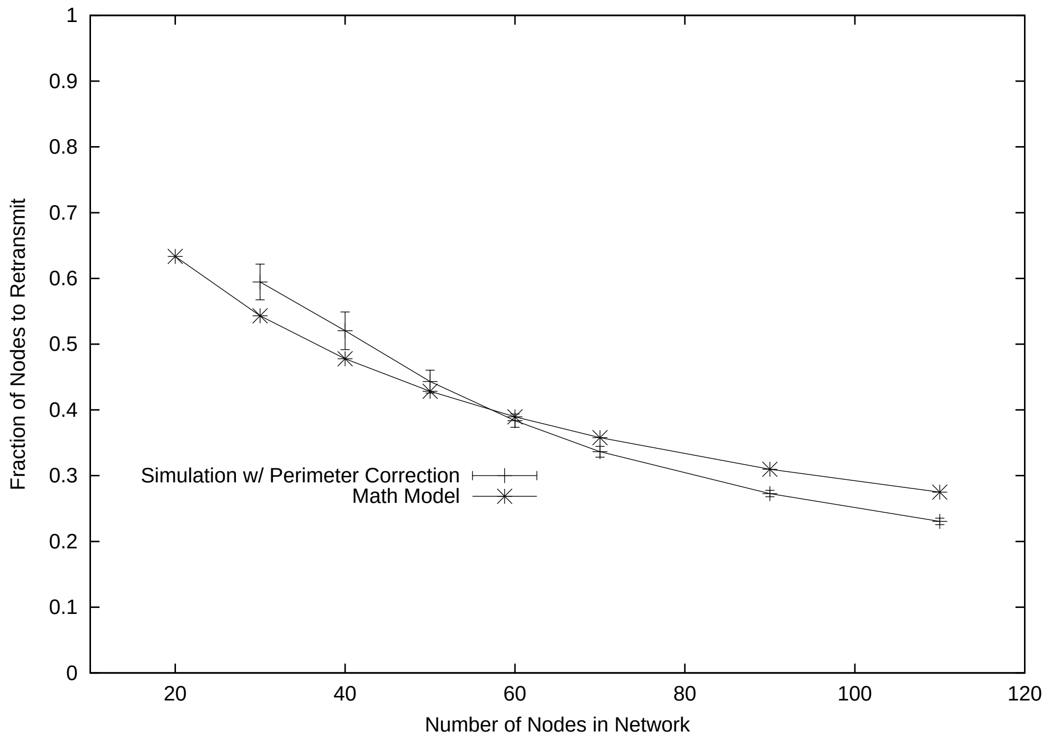

Figure 6: Math Model Prediction of Retransmission Probability for the Location-Based Scheme

nodes, when attached as a polygon, do not include the assessing node location as an interior point. Clearly, some correction should be made for a more probable propagation model.

5 Conclusions and Future Work

We proposed a generic procedure to produce models that can be used to explicitly predict the performance of network wide broadcast protocols. Specifically, we attempt to sum all probabilities that would induce a random node to rebroadcast a given packet and assume the behavior of all nodes mimics that one random node.

We developed mathematical models for the Counter-Based, Distance-Based, and Location-Based schemes and compared the results from our models to our simulation results. Each model predicts simulation results within 5% for all node densities. Unfortunately, our models tend to under predict retransmission frequency at low node densities (for both the Counter-Based and

Location-Based schemes) and over predict at high node densities (for all three schemes).

We believe that deviations between model prediction and simulator results is attributable to inaccurate assumptions. The assumption that each neighbor node receives a packet at the same time is a likely source of error in our models. An area of future research is to include a more accurate propagation model into the prediction scheme.

Another assumption that should be more closely evaluated is that each node in the network retransmits with the same independent probability. In all likelihood, there is a distribution of probabilities that, if included in our models, may improve the prediction accuracy of our models.

References

[1] D. Johnson, D. Maltz, Y.-C. Hu, and J. Jetcheva. The dynamic source routing protocol for mobile ad hoc networks (DSR). Internet Draft: draft-ietf-manet-dsr-07.txt, Feb. 2002.

[2] D. Johnson and D. Maltz. Dynamic source routing in ad hoc wireless networks. In T. Imelinsky and H. Korth, editors, *Mobile Computing*, pages 153–181. Kluwer Academic Publishers, 1996.

[3] C. Perkins and E.M. Royer. Ad-hoc on-demand distance vector routing. In *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, pages 90–100, 1999.

[4] C. Perkins, E. Royer, and S. Das. Ad hoc on demand distance vector (AODV) routing. Internet Draft: draft-ietf-manet-aodv-12.txt, November 2002.

[5] Z. Haas. A new routing protocol for reconfigurable wireless networks. In *Proceedings of the IEEE International Conference on Universal Personal Communications (ICUPC)*, Oct. 1997.

[6] Z. Haas, M. Pearlman, and P. Samar. The zone routing protocol (ZRP) for ad hoc networks. Internet Draft: draft-ietf-manet-zone-zrp-04.txt, July 2002.

[7] Z. Haas, M. Pearlman, and P. Samar. The interzone routing protocol (IERP) for ad hoc networks. Internet Draft: draft-ietf-manet-zone-ierp-02.txt, July 2002.

[8] Z. Haas, M. Pearlman, and P. Samar. The intrazone routing protocol (IARP) for ad hoc networks. Internet Draft: draft-ietf-manet-zone-iarp-02.txt, July 2002.

[9] Y. Ko and N.H. Vaidya. Location-aided routing (LAR) in mobile ad hoc networks. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM)*, pages 66–75, 1998.

[10] B. Williams and T. Camp. Comparison of broadcasting techniques for mobile ad hoc networks. In *Proceedings of the ACM Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC)*, pages 194–205, 2002.

[11] C. Ho, K. Obraczka, G. Tsudik, and K. Viswanath. Flooding for reliable multicast in multi-hop ad hoc networks. In *Proceedings of the International Workshop on Discrete Algorithms and methods for Mobile Computing and Communication (DIAL M)*, 1999.

[12] J. Jetcheva, Y. Hu, D. Maltz, and D. Johnson. A simple protocol for multicast and broadcast in mobile ad hoc networks. Internet Draft: draft-ietf-manet-simple-mbcast-01.txt (outdated), July 2001.

[13] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM)*, pages 151–162, 1999.

[14] K. Fall and K. Varadhan (Editors). UCB/LBNL/VINT network simulator - ns (version 2). http://www-mash.cs.berkeley.edu/ns/. Accessed November 15th, 2002, 1997.

[15] C. Bettstetter and O. Krause. On border effects in modeling and simulation of wireless ad hoc networks. In *Proceedings of the Third International Conference on Mobile and Wireless Communication Networks (MWCN'01)*, August 2001.