Theoretic Analysis of Two Broadcasting Protocols in Mobile Ad Hoc Networks

Hao Zhang and Zhong-Ping Jiang Department of Electrical & Computer Engineering Polytechnic University Brooklyn NY 11201, USA Email: hzhang08@utopia.poly.edu zjiang@control.poly.edu

Abstract—Broadcasting provides important functionality for wireless ad hoc networks. It is used in route discovery procedure in on-demand routing protocols such as DSR and AODV. In order to save wireless network capacity and avoid packets collisions, numerous broadcasting algorithms have been proposed to reduce the number of transmitting nodes. It is important to analyze and compare the effectiveness of all kinds of protocols. All the previous work in this area is based on simulations and experiments. We are not aware of any published paper dealing with rigorous theoretical analysis. In this paper, we analyze two popular ad hoc broadcasting protocols. The results show the relation between the performance of each protocol and selection of parameters.

I. INTRODUCTION

Many on-demand protocols such as DSR (Dynamic Source Routing) and AODV (On Demand Distance Vector Routing) rely on broadcasting techniques to initiate the route discovery process. Nodes receiving route request messages simply transmit them in traditional broadcasting scheme called 'flooding'. Although this simple broadcasting technique is easy to realize, recent research has shown that it will cause *broadcast storm problem* [4] in ad hoc networks.

The broadcast storm problem is due to the excessive amount of retransmissions resulting in severe message collision and channel contention. Various methods to alleviate the broadcast storm problem are proposed in [2-11]. All of these methods are trying to reduce the number of retransmitting nodes so as to increase the network capacity and avoid message collisions. All these schemes are divided into several classes and compared by simulation in [1]. Here we would like to introduce these methods briefly according to how much topology information one node has to know in each protocol.

The counter-based algorithm in [4] blocks a node's re-transmission if the number of times it has received the same message during a random period of time exceeds some threshold. This scheme is simple and easy to apply, and it requires no topology information. Probabilistic schemes are also studied in [4]. Every node that receives the message will retransmit it with some probability p.

When p = 1 this method is equivalent to the traditional flooding scheme. This technique was explored in more detail in [5], which described the bimodal phenomenon based on percolation theory. A probability p_c may exist for an infinite network where most of the nodes will get the message if $p > p_c$ and few nodes will get the message if $p < p_c$. The authors of [5] proposed some gossip schemes and did experiments in collision-free ideal networks to show that using probability between 0.6 and 0.8, gossip algorithms result in significant performance improvement. The above protocols are very simple for application because they require no topology information. On the other hand, their performance may not be as good as those more complex algorithms.

The sender's neighbors are denoted by one-hop nodes, the neighbors of these one-hop nodes that are more than one hop from the sender are denoted by two-hop nodes. A distance-based algorithm was proposed in [4], where the minimum distance from all the senders to the receiver is used for re-transmission judgement. An area based algorithm was also proposed in [4]. In this scheme, the receiving node judges whether to retransmit or not by calculating the area uncovered by the sending nodes. In [2] the distance and the cover angle are employed in retransmission decision. Another scheme based on one-hop nodes location was proposed in [3], but it relies on the sending node to select the retransmitting nodes. The authors of [7] proposed dynamic probabilistic broadcasting protocols where p is calculated dynamically according to some local topology information such as distance, density, etc.. It achieves some improvements over the probabilistic methods with fixed p according to the simulations, but it still did not give any proof to show why one has chosen this value p.

Two-hop topology is required by a multipoint relaying algorithm in [8]. The information is obtained by periodically broadcasting 'Hello' messages among neighbor nodes. With this information, each node calculates the *local cover set* or *multipoint relay set* of the one-hop nodes. The nodes in this set retransmit the received message and others will be refrained from retransmission. Unlike the algorithm in [8], where the retransmitting nodes are appointed by the central sending node, individual receiving nodes are

This work has been partially supported by the National Science Foundation under grants ANI-0081527 and ECS-0093176.

responsible for retransmission decisions in [10]. These two-hop schemes are usually more effective than one-hop schemes because the heuristic algorithms usually can achieve near optimal local cover set. But the transmission of two-hop topology information may consume a higher percentage of capacity so as to harm the performance when network density is increased.

Two most important performance metrics to evaluate the efficiency of a broadcasting protocol are RE (Reachability) and SRB (Saved Rebroadcast). RE is defined in [4] as the number of mobile hosts receiving the broadcast message divided by the total number of mobile hosts that are reachable, directly or indirectly, from the source host. SRB is defined as 1 - k/n, where n is the number of hosts receiving the broadcast message, k is the number of hosts actually transmitted the message[4]. When REs are the same for two protocols, higher SRB means better performance. The effectiveness of all these algorithms are demonstrated in the original papers and they are compared later in [1] by ns-2 simulation. But there is no theoretic work showing how well each of these algorithms performs in different networks models. Without theoretic analysis, we cannot know RE and SRB quantitatively in each protocol. It is also difficult to choose parameters (such as threshold values in some schemes) properly to achieve optimal performance without formulas relating parameters to performance metrics. This paper analyzes two popular ad hoc broadcasting protocols and makes a first step towards this important issue.

II. THEORETIC ANALYSIS OF TWO BROADCASTING PROTOCOLS

Here we consider two protocols widely known as the distance-based algorithm and counter-based algorithm. SRB(n) denotes the Saved Rebroadcast for $n \times 1$ onedimensional network or $(n + 1) \times (n + 1)$ two-dimensional networks. It is assumed that the whole topology is known to our analysis, although the algorithms are decentralized and require no information or only local information. For the sake of simplicity, we begin with the situation where no packet collisions occur and all the nodes are equally distributed in the space. For the one-dimensional case, all n nodes are lined up and the distance between two adjacent nodes is 1. Here the transmission radius R is set to positive integers for simplicity.

In the counter-based algorithm, a node initiates a counter with a value of one and sets a RAD (which is randomly chosen between 0 and *Tmax* seconds) upon reception of a previously unseen packet[1]. During the RAD, the counter is incremented by one every time it receives a redundant message. If the counter is larger than or equal to a threshold $C \ (C \in Z^+, C \ge 2)$ when the RAD timer expires, the node should drop this message and do not retransmit it. Otherwise the message should be retransmitted.

In the distance-based algorithm, a node initiates a RAD

(which is randomly chosen between 0 and *Tmax* seconds) upon reception of a previously unseen packet. All the redundant packets will be cached during the RAD period. When the RAD expires, all sender locations are examined to see if the minimum distance between the receiver and the sender nodes is less than or equal to a threshold $d \in Z^+$. If it is true, this node should drop this message and do not retransmit it.

A. One-Dimensional Grid Network

For the one-dimensional case, the message is initiated from the original sender O and transmitted to the nodes located between 1 and +n. It is obvious that REs for these two schemes are 100%. We provide lower bounds and upper bounds for two kinds of SRB: SRB(n) and E(SRB(n)). SRB(n) stands for SRB of every broadcasting execution, while E(SRB(n)) stands for expectation of SRB for all broadcasting executions. Notice the bounds for SRB(n)are also the bounds for E(SRB(n)).

Theorem 1: If C = 2, $\lim_{n\to\infty} E[SRB(n)] = 1 - \frac{3}{2R+1}$ for the counter-based algorithm.



Fig. 1. Proof of Theorem 1

Proof: Among all the nodes that received a message after each transmission, only the node whose timer expires first can retransmit it and block other nodes from retransmission since the counter threshold C = 2. In the figure above, if node *i* transmitted the message, *i* additional nodes would receive it. Each of these *i* nodes has probability 1/i to transmit the message again and block other i - 1 nodes' transmissions. Let S_k be the number of new nodes that received the message at the *k*th transmission, where S_k can be any value in set $1, 2, \ldots, R$. Let a_k be a vector whose *i*th element $(1 \le i \le R)$ is the probability $P(S_k = i)$ (see Fig.1). Since a_{k+1} depends only on a_k , it is a Markov process. We have then

$$a_{k+1} = Pa_k \tag{1}$$

where P is the transition matrix and

We notice that $P^2 > 0$ (componentwise), so P is regular. Then the power P^k approaches a probability matrix Q when k goes to infinity, where $Qa_0 = u$ for any initial probability distribution vector a_0 and u is the unique stationary distribution[12]. We can compute the stationary distribution by:

$$Pu = u \tag{2}$$

The solution is:

$$\mathbf{u} = \begin{bmatrix} \frac{1}{1+2+\ldots+R} & \frac{2}{1+2+\ldots+R} & \cdots & \frac{R}{1+2+\ldots+R} \end{bmatrix}^T$$

Suppose totally x_n nodes transmit the message. Let $v_i(x_n)$ represents the fraction of the transmissions that the process can be expected to be in state s_i (i new news receive the message for one transmission). According to the Law of Large Numbers for regular Markov chains, $E[v_i(x_n)] \to u_i$ if $x_n \to \infty$. Additionally, $\sum_{i=1}^{R} iv_i(x_n) \frac{x_n}{n} = 1$. We have the following equation:

$$\lim_{n \to \infty} \sum_{i=1}^{R} i u_i E[\frac{x_n}{n}] = 1$$
(3)

From this equation and $SRB(n) = 1 - \frac{x_n}{n}$, we have $\lim_{n\to\infty} E[SRB(n)] = 1 - \frac{3}{2R+1}$. \blacksquare *Theorem 2:* For any 2 < C < R, $\liminf_{n\to\infty} SRB(n)$ $\geq 1 - \frac{C}{R+1}$, $\limsup_{n\to\infty} SRB(n) \leq 1 - \frac{C}{2R+C-1}$ for the counter based algorithm counter-based algorithm.

Proof: Suppose x nodes transmit the message in an arbitrarily selected R+1 consecutive nodes. x can only achieve its maximum when these nodes receive this message only from themselves (none of other nodes sent them the message). So $x \leq C$ and the lower bound is proved.

Now suppose k nodes transmitted the message in all the n nodes. Each of these k nodes (the original sender O is not included) should have received at least one copy of this message and each of other n-k nodes should have received at least C copies of this message. So the number of the total message copies should be at least (n-k)C + k. In each transmission, at most 2R nodes receive a copy, so we have then:

$$2Rk \ge (n-k)C + k \tag{4}$$

The ratio $\frac{k}{n} \geq \frac{C}{2R+C-1}$ and notice $SRB(n) = 1 - \frac{k}{n}$, the upper bound is proved.

Remark: From the above theorems we can see that smaller C/R leads to better performance. C = 2 is the optimal threshold value that can achieve the highest SRB for all R > 2.

Theorem 3: When $\frac{R-1}{2} \leq d \leq R-1$, $\lim_{n\to\infty} E[SRB(n)] = 1 - \frac{2}{R+d+1}$ for the distancebased algorithm.

Proof: If node *i* transmitted a message, then all the nodes in [i, i + d] will be blocked from retransmitting it. Any node i + d + j $(1 \le j \le R - d)$ can transmit it with probability 1/(R - d). But only one node can perform the transmission. This is because the maximum



Fig. 2. Proof of Theorem 3

distance between two nodes in [i + d + 1, i + R] is $R-1-d \leq R-1-\frac{R-1}{2} \leq d$. Let the state space s = (d+1, d+2, ..., R). The $(R-d) \times (R-d)$ transition matrix

$$\mathbf{P} = \begin{pmatrix} \frac{1}{R-d} & \cdots & \frac{1}{R-d} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \frac{1}{R-d} & \cdots & \frac{1}{R-d} \end{pmatrix}$$

The stationary distribution $u = [\frac{1}{R-d} \dots \frac{1}{R-d}]^T$. So $\lim_{n\to\infty} E[SRB(n)] = 1 - \frac{2}{R+d+1}$.

 $SRB(n) \ge 1 - \frac{1}{d+1}$, $\limsup_{n \to \infty} SRB(n) \le 1 - \frac{1}{2d+1}$ for the distance-based algorithm.

Proof: The smallest distance between two adjacent nodes which transmitted the same message (that means the nodes between them did not transmit the message) is d + 1. One the other hand, the largest distance between two adjacent nodes which transmitted the same message is 2d+1 (otherwise at least one node between them will not be blocked from transmission). So the inequalities follow readily.

Remark: The largest SRB is achieved when threshold d = R - 1 (d can not be larger than R, or RE is 0%.). The optimal value is 1 - 1/R, which is larger than the optimal value of the counter-based scheme $1 - \frac{3}{2R+1}$ for R > 1.

B. Two-Dimensional Grid Network

The structure of a two-dimensional grid network is as follows:

All the nodes are located at $(i, j)(-n/2 \le i, j \le +n/2)$ and the original message sender O is located at (0,0). This forms a grid network with $(n + 1)^2$ nodes (not including O). Suppose all the nodes received a message in each broadcasting execution, $SRB(n) = \frac{k}{(n+1)^2-1}$, where nodes received the message do not include the original sender and k denotes the number of transmitting nodes. We will use $SRB(n) = \frac{k}{n^2}$ for simplicity, and this approximation will not effect our results. We still use the coordinates to represent these nodes and assume R to be integers as in the one-dimensional case.

Theorem 5: RE of the distance-based algorithm is 100% for d < R.



Fig. 3. Proof of Theorem 5

Proof: We prove Theorem 5 by contradiction. Suppose there is a node i which did not receive a message. We will show that any other node j did not receive the message either. The circle centering at node i with radius R will be denoted by 'circle i'. According to the assumption, all the nodes within circle i did not transmit this message. At first we will show that any node whose location is no more than R - d from node i did not receive this message. If not, suppose node h received this message and $|h - i| \leq R - d$, where |h - i| denotes the distance from node h to i. There must exist some node p which sent node h this message and $|h - p| \leq d$. So we have $|i - p| \leq |i - h| + |h - p| \leq R - d + d = R$. This means that node i can receive this message from p, which leads to the contradiction with the assumption.

Next we return to prove that node $j \neq i$ did not receive the message. Suppose the coordinates of node i are (0, 0) and node j is located at (j_x, j_y) . Without loss of generality, assume $j_x \geq 0, j_y \geq 0$. Let q be the node locating at (R - d, 0). We have proved that q did not receive the message. By iteratively using this result for m times, where $(R - d)(m - 1) < j_x \leq (R - d)m$, we conclude that the node at $(j_x, 0)$ did not receive this message. Similarly, starting from $(j_x, 0)$, we obtain that node $j = (j_x, j_y)$ did not receive this message. Since node j is arbitrarily chosen, all the nodes in this two-dimensional grid must not have received the message, which contradicts the fact that at least the nodes within R distance from the original sender have received it. Therefore RE of the distance-based scheme is 100% if d < R.

Theorem 5 shows no lower bound is required for the threshold d to ensure 100% RE for the distance-based algorithm. But it is a different situation in the counterbased algorithm. A counter-example in the case of two-dimensional grid networks can be easily given as follows: take C = 2 in a 5×5 grid network with R = 5. Node O is the original message sender and node i_1, j_1, i_2, j_2 re-transmitted the message sequently. After these four transmissions, all the black nodes have received this message at least twice and are blocked from retransmission. Other white nodes did not receive the message and RE of this execution is less than 100%. The following theorem gives a sufficient condition which



Fig. 4. An example when C = 2 and RE < 100%

defines the range of parameter C when all the nodes can receive the message in every broadcasting execution.

Theorem 6: RE of the counter-based algorithm is 100% if C > 2R + 1.



Fig. 5. Proof of Theorem 6

Proof: Suppose at least one node did not receive the message as shown in Fig.3. White dots represent nodes that did not receive the message, black dots represent nodes that received the message. We can always find two adjacent nodes i, j, where node i is black and located at O_1 and node j is white and located at O_2 . Since node i received the message but did not transmit it (otherwise node j would have received it), node i must have at least received C copies of this message from some nodes locating inside circle O_1 but outside circle O_2 (do not include those nodes whose distances to O_2 are equal to R). There are 2R + 1 nodes in this area. To see this, notice that:

$$AC| = \sqrt{R^2 - |O_1 C|^2}$$
(5)

$$|BD| = \sqrt{R^2 - |O_2D|^2} \tag{6}$$

So

$$AB| = |CD| = |O_1O_2| = 1 \tag{7}$$

There is exactly one node between A and B (do not include B). The number of nodes that may send node i the message is at most 2R + 1. Therefore $C \le 2R + 1$ and the theorem follows.

Lemma 1: The number of nodes in a circle centering at (0,0) is less than $3.3R^2$ if R > 1.

Proof: Suppose $\lfloor a \rfloor$ denotes the largest integer that is not more than a. The number of nodes S(R) within a circle centering at (0,0) with radius R is calculated by:

$$S(R) = \sum_{i=-R}^{R} (2\lfloor \sqrt{R^2 - i^2} \rfloor + 1)$$

$$\leq \sum_{i=-R}^{R} (2\sqrt{R^2 - i^2} + 1)$$

$$= 4R + 1 + 4\sum_{i=1}^{R} \sqrt{R^2 - i^2}$$

$$\leq 4R + 1 + 4\sqrt{R\sum_{i=1}^{R} (R^2 - i^2)}$$

$$= 4R + 1 + \frac{4}{\sqrt{6}}R\sqrt{4R^2 - 3R - 1}$$

$$< 4R + 1 + \frac{8}{\sqrt{6}}R^2$$
(8)

Jensen's inequality is used above. Notice that

$$S(R) - 3.3R^2 < -0.03R^2 + 4R + 1 \tag{9}$$

Because two real-number roots of $-0.03R^2 + 4R + 1 = 0$ are less than 150, we have:

$$S(R) - 3.3R^2 < 0 \tag{10}$$

for all $R \ge 150$.

Additionally, we verified this lemma by a small computer program for 1 < R < 150 (It is easy to do so since all the coordinates are integers). So this lemma is proved for all integers R > 1.

Theorem 7: For any C > 2R + 1, $\liminf_{n \to \infty} SRB_n \ge 1 - \frac{2C}{R^2}$, $\limsup_{n \to \infty} SRB_n \le 1 - \frac{C}{3.3R^2 + C - 1}$ for the counter-based scheme.

Proof: According to Theorem 6, RE of the counterbased scheme is 100% for C > 2R + 1. The whole grid network can be covered by $\lceil \sqrt{2n}/R \rceil^2$ squares with diagonal length R. Here $\lceil a \rceil$ denotes the smallest integer that is not less than a. Similar to the one-dimensional case, at most C transmitting nodes can be in one such square. So the SRB_n satisfies:

$$SRB_{n} \ge 1 - \frac{\lceil \sqrt{2n/R} \rceil^{2}C}{n^{2}}$$

> $1 - \frac{(\sqrt{2n/R} + 1)^{2}C}{n^{2}}$ (11)

We have then

$$\liminf_{n \to \infty} SRB_n \ge 1 - \frac{2C}{R^2} \tag{12}$$

The lower bound is proved.

To prove the upper bound, assume k nodes transmitted the message. The number of the total received message copies should be at least $(n^2 - k)C + k$. According to Lemma 1, at most $3.3R^2$ nodes received a copy in each transmission. So we have

$$3.3R^2k > (n^2 - k)C + k \tag{13}$$

$$SRB_{n} = 1 - \frac{n}{n^{2}}$$

$$< 1 - \frac{C}{3.3R^{2} + C - 1}$$
(14)

Theorem 8: For any 0 < d < R, $\liminf_{n \to \infty} SRB_n \ge 1 - \frac{2}{d^2}$, $\limsup_{n \to \infty} SRB_n \le 1 - \frac{1}{(2d+1)^2}$ for the distance-based scheme.

Proof: The whole grid network can be covered by $\lceil \sqrt{2n/d} \rceil^2$ squares with diagonal length d. At most one node in each square can transmit the message. So the SRB_n satisfies:

$$SRB_{n} \ge 1 - \frac{\lceil \sqrt{2n/d} \rceil^{2}}{n^{2}}$$

> $1 - \frac{(\sqrt{2n/d} + 1)^{2}}{n^{2}}$ (15)

We have then

$$\liminf_{n \to \infty} SRB_n \ge 1 - \frac{2}{d^2} \tag{16}$$

The whole network can be partly covered by $\lfloor n/(2d+1) \rfloor^2$ $(2d+1) \times (2d+1)$ squares without overlapping. Each of these squares contains at least one transmitting node. So the SRB_n satisfies:

$$SRB_n < 1 - \frac{\lfloor n/(2d+1) \rfloor^2}{n^2} < 1 - \frac{(n/(2d+1)-1)^2}{n^2}$$
(17)

Therefore the upper bound is proved.

III. CONCLUSION

In this paper, we analyze two ad hoc broadcasting protocols in one-dimensional grid networks and two-dimensional grid networks. Sufficient conditions are given to achieve 100% RE for networks using each of these two algorithms. We can compute the exact limit $\lim_{n\to\infty} E[SRB(n)]$ in one-dimensional networks when C = 2 for the counterbased algorithm and $\frac{R-1}{2} \leq d \leq R-1$ for the distancebased algorithm. In other cases, lower bounds and upper bounds are obtained which indicate the relation between SRB and threshold values. A more complex problem is how to analyze these algorithms if nodes are randomly distributed in a unit-area disc. The results in [13] show that for $\pi R^2(n) = \frac{\log(n) + c(n)}{n}$ the network is asymptotically connected with probability one if and only if $c(n) \to \infty$. Here n denotes the total number of nodes in the unit-area disc. We are investigating the characteristics of RE and SRB when the whole network is asymptotically connected. Another important issue is energy conservation, which is also the main reason to reduce the number of rebroadcasting nodes. From the analysis, it is obvious that SRB will be large if R is chosen to be large (For example, choose R = n for the one-dimensional case). But increasing R does not necessarily result in performance improvement. This is because more energy will be consumed in each transmission in spite of less transmissions are needed to cover all the nodes. There exists a trade-off between the number of transmitting nodes and the energy consumed by each transmitting node. Generally the energy consumed by each node is considered as a nonlinear function of Rand there should exist optimal values R, C, d that achieve the minimum total energy consumed by all the nodes. We are working on this interesting problem and will report the results in our future work.

REFERENCES

- B.Williams and T.Camp, Comparison of broadcasting techniques for mobile ad hoc networks. Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MO-BIHOC), pp.194-205, 2002.
- [2] M.T.Sun, W.Feng, T.H.Lai, Location Aided Broadcast in Wireless Ad hoc Networks. Proc. IEEE WCNC 2002, pp.597-602, Orlando, FL, March 2002.
- [3] M.T.Sun, T.H.Lai, Computing Optimal Local Cover Set for Broadcast in Ad Hoc Networks. Proc. IEEE ICC 2002, pp.3291-3295, New York, NY, April 2002.
- [4] S.Ni, Y.Tseng, Y.Chen and J.Sheu, The broadcast storm problem in a mobile ad hoc network. Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM), pp.151-162, 1999.
- [5] Z.J.Haas, J.Y.Halpern and L.Li, Gossip-based ad hoc routing. IEEE INFOCOM, pp.1707-1716, June 2002.
- [6] Y.Sasson, D.Cavin and A.Schiper, Probabilistic Broadcast for Flooding in Wireless Mobile Ad hoc Networks. Technical Report, IC/2002/54.
- [7] J.Cartigny, D.Simplot, J.Carle, Stochastic Flooding Broadcast Protocols in Mobile Wireless Networks.Tech. Report LIFL Univ. Lille 1 2002-03, May 2002.
- [8] A.Qayyum, L.Viennot and A.Laouiti, Multipoint Relaying for Flooding Broadcast Messages in Mobile Wireless Networks. Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35'02), pp.3898-3907, 2002.
- [9] H.Lim and C.Kim, Multicast tree construction and flooding in wireless ad hoc networks. Proceedings of the ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM), pp.61-68, 2000.
- [10] J.Sucec and I.Marsic, An efficient distributed network-widebroadcast algorithm for mobile ad hoc networks. CAIP Technical Report 248 - Rutgers University, September 2000.
- [11] W.Peng and X.Lu, On the reduction of broadcast redundancy in mobile ad hoc networks. Proceedings of MOBIHOC, pp.129-130, 2000.
- [12] J.G.Kemeny, J.L.Snell, Finite Markov Chains. D. Van Nostrand Company, INC., pp.69-75, 1959.
- [13] P.Gupta and P.R.Kumar, Critical power for asymptotic connectivity. Proc. IEEE 37th Conf. on Decision and Control, pp.1106-1110, 1998.