A Comparison of Flooding and Random Routing in Mobile Ad Hoc Networks

Daejoong Kim, N.F. Maxemchuk

Columbia University

ABSTRACT

Mobile ad hoc networks usually use various forms of flooding to discover the location and route of a node or to diffuse a given message. Flooding is useful because it finds a good path to a destination or quickly diffuses the message all over the network. With the flooding, however, also comes the problem of propagating many unnecessary messages throughout the network. In this paper, first, we propose a flooding on the arms of hexagon that significantly reduces redundant messages. It is shown that the efficiency of the hexagon flooding reaches about 68% of the upper-bound efficiency of flooding. Secondly, we analyze random routing on the arms of hexagon that uses fewer messages to locate a destination than flooding. Assuming that a node has first-order neighbor knowledge, the average number of messages to locate a destination can be reduced to one-third of flooding. As the order of neighbor knowledge increases, random routing can further reduce unnecessary messages.

I. INTRODUCTION

Mobile ad hoc networks usually use various forms of flooding to discover the location and route of a node or to diffuse a given message all over the network. Flooding is useful because it finds a good route to a destination or quickly diffuses a given message, but at the same time it propagates many unnecessary messages all over the network.

In this paper, we present a hexagon flooding that significantly reduces redundant messages, and also show that most of the time random routing can be a more efficient routing algorithm than flooding to locate a destination in the network.

Recently, Vamsi K Paruchuri presented a similar flooding scheme, Optimal Flooding Protocol[1], which was also based on the hexagonal partitioning of the network area like the hexagon flooding described in this paper. He showed that his optimal flooding protocol could reduce a number of redundant messages, compared to pure flooding and GOSSIP[2]. However, the protocol still has some room for improvement. First, with the optimal flooding protocol, a source is located at the center of a hexagon. This makes routing in intermediate nodes more complex because they should apply a different routing rule if they receive the message directly from the source. Secondly, the protocol may not work in real mobile networks that are not partitioned into regular hexagons. If three neighbor vertex nodes cannot select the same next forwarding node in the designated area, then the protocol may repeat the flooding several times. Our stopping rules, which are described in the last part of this section, can prevent the repeated flooding.

In this paper, we consider a mobile ad hoc network that has a high density of nodes. Every node in the network is assumed to have the same transmission radius for convenient analysis. We assume that every node knows the geographical positions of itself and its neighbors within its transmission radius. It can be easily achieved using a global positioning system(GPS). We don't care about how to obtain and update neighbor knowledge, which is beyond the scope of this paper. Every message exchanged in the network is supposed to have a unique identifier consisting of the sender's identity and timestamp. It makes a node distinguish every message traveling in the network. Not to confuse the messages received, every node should remember the identities for at least T that is long enough to finish a flooding in the network.

Now, we describe a hexagon flooding. With the hexagon flooding, a network area is partitioned into hexagons. The transmission radius of a vertex node determines the length of the arms of the hexagon. The source and intermediate nodes that forward messages are at the vertexes of the hexagons. The messages emanated from the source are propagated along the arms of the hexagons. The hexagon flooding is based on the notion that an efficient way to cover an entire area with the minimum number of circles is to form hexagons by their intersection points. Any node located inside a hexagon is reachable from at least one of the vertex nodes of the hexagon, which are located at the intersection points. Therefore, a given message is delivered to every node in the hexagon if each vertex node broadcasts the message once.

For the details of the hexagon flooding, we need to explain both how to select the next forwarding nodes and when the messages stop traveling. The selection is made so that the network area can be partitioned into hexagons. The source emanates a message selecting three neighbor edge nodes which make an adjacent 120° angles. The selected nodes emanate the message selecting two neighbor edge nodes for the next forwarding nodes. The selected two nodes make a ±120° angle from the node from which the message was received. In this way, the messages are propagated until they cover the entire network area. To prevent repeated flooding in real networks, only the first received message is used to determine to forward the message. We summarize three selection rules and three stopping rules as follows.

Selection Rules

1. A source selects three edge nodes to make an adjacent 120° angles for the next forwarding nodes.
2. An intermediate node selects two edge nodes to make a 120° angle with the node from which it received the message.
3. If there is no edge node at the point calculated, select the nearest node to the point. The node should be located within the sector that makes a ±60° angle from the point.

Stopping Rules

1. Only selected nodes are allowed to re-broadcast a given message once.
2. If a selected node cannot find any next forwarding node by the selection rules, it does not re-broadcast the message.
3. If a node is not selected in the first received message, it does not re-broadcast the message.

The stopping rules make sure that flooding is over in a finite time because a node is not allowed to re-broadcast the same message more than once. The hexagon flooding works efficiently in high-density networks and it is scalable with respect to the number of mobile nodes in the region.

In section 2, we estimate the efficiency of hexagon flooding. It is compared to the upper-bound efficiency of the ideal optimal flooding that is assumed to minimize the number of redundant messages in mobile ad hoc networks. The efficiency of any other flooding scheme cannot exceed that of the ideal optimal flooding.

In section 3, we analyze random routing on the arms of hexagon. The analysis shows that most of the time random routing uses fewer messages to locate a destination than hexagon flooding. Random routing can further reduce unnecessary messages as the order of neighbor knowledge increases.

II. EFFICIENCY OF HEXAGON FLOODING

In mobile ad hoc networks, we define the efficiency of flooding as the ratio of the entire network area to the total areas that every broadcasted message covers. Therefore, assuming that every node has the same transmission radius, the efficiency η can be expressed as follows.

$$\eta = A_T / nA_R \qquad (1)$$

A_T is an entire area of the network, A_R an area that a node with a transmission radius R covers, and n the total number of broadcasts of the given message. The definition $1/\eta$ means the average number of messages received per node. The efficiency is 1 if every node in the network receives a given message once. The full efficiency can be achieved when a source that has a large transmission radius able to cover the entire network area, emanates a given message once. In real networks, however, the transmission radius is not large enough to cover the entire area. A lot of nodes with small radii should forward the message to cover the entire area. This results in some nodes in the overlapped area receiving redundant messages thereby decreasing the efficiency. We will use the efficiency factor η to compare the hexagon flooding with the ideal optimal flooding described in the next part.

Considering that a mobile ad hoc network has a high density of nodes, we can find a node at any strategic point. Let's suppose that we have an ideal optimal flooding algorithm. The algorithm may not be feasible, but we don't care about how to implement it. The algorithm is assumed to ensure optimal routing in the given network in order to minimize delivering redundant messages to every node in the network. In other words, the ideal optimal flooding is to find the minimal number of circles to cover the entire network area, keeping the distance between the centers of two nearest circles less than or at most equal to the radius R because they should be reachable to each other. In the mobile ad hoc network that has an entire area A_T and the same transmission radius R, the ideal optimal flooding could form the pattern as described in Fig. 1.

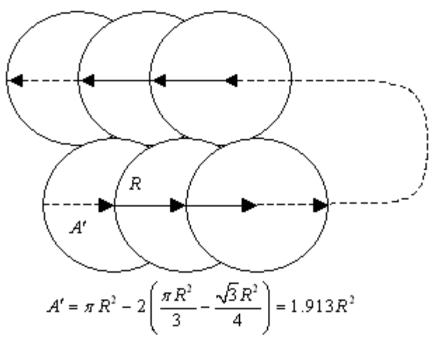

Fig. 1. Upper-bound efficiency of flooding

In Fig. 1, the circles are arranged to minimize the overlapped area, making the center of every circle reachable. The center of a circle corresponds to a selected forwarding node and the circle area means the reachable area from the node at the center. If every node receives a given message after n broadcasts, it means that the entire network area A_T contains n circles. Now, we can compute the efficiency of the ideal optimal flooding. In Fig. 1, we know that each circle extends the covered area by $A' = 1.913R^2$. Therefore, n times A' should be approximately equal to A_T. Since $n \times 1.913R^2 = A_T$, the efficiency is as follows.

$$\eta = \frac{A_T}{nA_R} = \frac{n \times 1.913R^2}{n \times \pi R^2} \approx 0.6 \qquad (2)$$

Expression (2) shows that the efficiency of the ideal optimal flooding does not depend on the transmission radius R or the number of transmissions n. Since $1/\eta$ means the average number of messages received per node, a node would receive average 1.6 messages in the case of using the flooding. The efficiency of any flooding scheme cannot exceed 0.6 in mobile ad hoc networks.

Actually, the average A' is smaller than $1.913R^2$ because the overlapped area increases when a given message goes around the border area of real networks. Besides, we ignored the overlapped area between the rows of circles in Fig. 1. Therefore, the efficiency number 0.6 is interpreted as the upper-bound efficiency of the ideal optimal flooding.

In this part, we analyze the hexagon flooding to estimate its efficiency. The left figure in Fig. 2 explains the hexagon flooding. The source A broadcasts a given message, selecting nodes B, C and D for the next forwarding nodes. The selected nodes B, C and D broadcast the message, selecting nodes F, G, H, I, J and E for the next forward nodes. Since every vertex node of a hexagon broadcasts the message once, all nodes inside the hexagon can receive the message at least once. In this way, the message is delivered to all nodes in the network. As in Fig. 2, the network area can be partitioned into regular hexagons if the network has a high density of nodes. The circles with their centers at each vertex of the hexagon make a regular pattern like the shaded triangle in Fig. 2. It means that we can estimate the average number of messages received per node by analyzing the shaded regular triangle. In the triangle, both areas A_1 and A_2 are reachable from three vertex nodes while the area A_3 is reachable from two vertex nodes. It means that a node located in A_1 or A_2 receives the message three times and a node in A_3 receives it twice. Therefore, the efficiency of the hexagon flooding can be computed as follows.

$$\eta = \frac{A_1 + A_2 + A_3}{3A_1 + 3A_2 + 2A_3} = \frac{0.43R^2}{1.04R^2} \approx 0.41 \qquad (3)$$

The efficiency of the hexagon flooding is about 68% of the upper-bound efficiency. It also means that a node receives on the average $2.4 (= 1/\eta)$ messages per node in the case of using the hexagon flooding. The expression (3) shows that the efficiency does not depend on the transmission radius R or the total number of transmitting nodes N. It makes the hexagon flooding scalable as the number of nodes increases in the region.

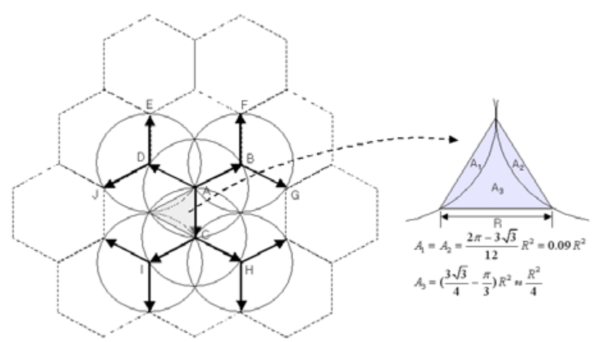

Fig. 2. Hexagon flooding

In the hexagon flooding, the total number of messages required to deliver a given message to all nodes in the network is N, which is the same as the total number of vertex nodes in the network. The same number of messages is also required to locate a destination whether it is close to the source or not. In next section, however, we show that random routing can significantly reduce the number of messages needed to locate a destination.

III. RANDOM ROUTING

In this section, we present a random routing on the arms of hexagon and analyze it to estimate the average number of messages needed to locate a destination. For the random routing, a source transmits a given message to a randomly selected edge node. The selected node checks if the destination is located within its transmission radius before forwarding the message. If not, the selected node forwards the message to the next node that is selected at random based on the Selection Rules 2 and 3. If the selected node cannot find any node according to the Selection Rules 2 and 3, then it may send the message back to the node from which it received the message. If the selected node finds the destination, it sends the message to the destination node and the message stops traveling.

It is assumed that a node has first-order neighbor knowledge, $k = 1$. For $k > 1$, we will give an intuitive explanation at the end of this section. For our analysis, let p be the probability that a given message, placed at random in the mobile ad hoc network, is forwarded to its destination on the next transmission. Then the probability that the message arrives at its destination on the n^{th} transmission is $p(1-p)^{n-1}$ which is a familiar geometric random variable. Therefore, in the network using random routing, the average number of transmissions from source to destination and variance σ^2 are readily computed.

$$m = 1/p\ ,\quad \sigma^2 = (1-p)/p^2 \tag{4}$$

Note that σ^2 is larger than m. Since $1-p \approx 1$, the variance is approximately the square of mean. This may be interpreted as saying that some messages may have very large hops to their destinations.

It remains to give estimates for the value of p. We can estimate p in mobile ad hoc networks in a similar way as Prosser did in wired networks[3]. In mobile ad hoc networks, the probability p depends on the location of a destination. We ignore the probability that the destination is at a vertex node because the number of vertex nodes is assumed to be trivial compared to the total number of nodes in the network. Assuming that the network has a high density of nodes, the network can be partitioned into regular hexagons as shown in Fig. 3. The destination can be found anywhere in the network with the same probability. Therefore we can estimate the desired probability p by analyzing the shaded regular triangle.

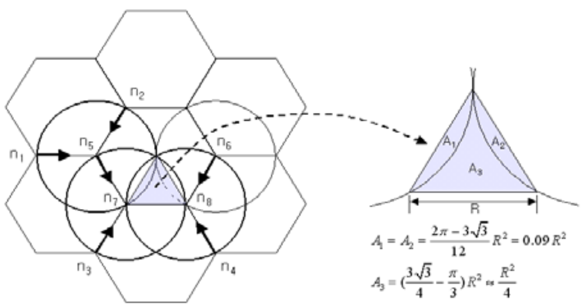

Fig. 3. Random routing

For our analysis, let a destination be located in a regular triangle as in Fig. 3. According to the number of reachable vertex nodes, the triangle is partitioned into three areas A_1, A_2 and A_3. Both areas A_1 and A_2 give higher probability p than A_3 because they have one more reachable vertex node. The message at a vertex node n_7 or n_8 definitely arrives at the destination on the next transmission, but the message at n_5 (or n_6) arrives only when the destination is in A_1 (or A_2). The set of vertex nodes from which the message is relayed to the destination on the next transmission, is called a destination zone. What we have to estimate is the probability that the message arrives at the destination zone on the next transmission.

The probability p also depends on the position of the source. Suppose that we have a message at n_3. The probability that the node n_3 selects n_7, which belongs to the destination zone, is one-third if n_3 is the source of the message, but one-half if the message comes from another node because n_3 is not supposed to select the node from which it received the message.

Now, we are ready to compute the desired probability p. Let $P(F_1)$ be the probability that a source is located at the destination zone. It is zero since the message is already in the destination zone. Let $P(F_2)$ be the probability that the source is located at first-order neighbor vertex nodes of the destination zone and $P(F_3)$ the probability that the source is one of the other nodes. $P(E)$ is the probability that the message arrives at its destination zone on the next transmission and $P(A_i)$ is the probability that the destination is in area A_i. First, we compute the conditional probabilities on the location of destination.

$$P(E|A_1) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + P(E|F_3)P(F_3)$$
$$= 0 \times \frac{3}{N} + \frac{1}{3} \times \frac{5}{N} + \frac{5}{N-3} \times \frac{1}{2} \times \frac{N-8}{N} = \frac{25N-150}{6N(N-3)} \approx \frac{4.17}{N}\bigg|_{\text{large } N}$$

$P(E|A_2) = P(E|A_1)$ by symmetry,

$$P(E|A_3) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + P(E|F_3)P(F_3)$$
$$= 0 \times \frac{2}{N} + \frac{1}{3} \times \frac{4}{N} + \frac{4}{N-2} \times \frac{1}{2} \times \frac{N-6}{N} = \frac{10N-44}{3N(N-2)} \approx \frac{3.33}{N}\bigg|_{\text{large } N}$$

For convenient computation, we assumed that the total number of vertex nodes in the network, N is large. The desired probability is as follows.

$$p = P(E|A_1)P(A_1) + P(E|A_2)P(A_2) + P(E|A_3)P(A_3)$$
$$= \frac{4.17}{N}\frac{0.09R^2}{0.43R^2} + \frac{4.17}{N}\frac{0.09R^2}{0.43R^2} + \frac{3.33}{N}\frac{0.25R^2}{0.43R^2} \approx \frac{3.68}{N}\bigg|_{\text{large } N} \tag{5}$$

From expressions (4) and (5), the average total number of transmissions from source to destination in random routing is $N/3.68$ and its variance is about $(N/3.68)^2$. Therefore, we can say that the random routing uses approximately one-third of the total number of messages that are used by flooding to locate a destination. However, the variance is very large in random routing while it is zero in flooding. About 97% of the time, random routing needs fewer copies of the messages than flooding.

Now, we consider random routing with $k > 1$. In expression (5), p is regarded as an average for $P(E|A_1)$, $P(E|A_2)$, and $P(E|A_3)$. Therefore, p can be represented as the following general expression.

$$p = \frac{H}{N} \times \frac{1}{3} + \frac{H}{(N-D)} \times \frac{1}{2} \times \frac{N-D-H}{N} = \frac{H}{N}\left(\frac{1}{3} + \frac{N-D-H}{2(N-D)}\right)$$
$$\approx \frac{5H}{6N}\bigg|_{\text{large } N} \tag{6}$$

D is the average number of vertex nodes in the destination zone and H is the average number of neighbor vertex nodes of the destination zone. In expression (6), we assumed that $N-D-H$ is approximately equal to $N-D$ for large N. Since the number of nodes on the k th hop is $3k$ in the hexagons topology, the average number of neighbors, H, has a value between $3k$ and $3(k+1)$ for the system that has the k th order neighbor knowledge. This means that H increases as k increases. Therefore, from both expressions (4) and (6) we can say that the average number of transmissions in random routing decreases as k increases.

In the random routing, the average number of transmissions also represents the delay from source to destination. If it takes a unit time per hop including processing and propagation delay, the average delay from source to destination is $N/3.68$ in the random routing while it is only a few hops in flooding.

IV. CONCLUSION

We presented a hexagon flooding scheme. Knowing that any flooding scheme can never outperform the ideal optimal flooding, our hexagon flooding is still quite efficient. Actually, the efficiency of hexagon flooding is about 68% of the upper-bound efficiency. We also presented random routing on the arms of hexagon. It turns out that the random routing uses on average one-third of the total number of message that are required in flooding to locate a destination. Additionally, as the order of neighbor knowledge increases, the random routing can further reduce the number of messages. If fast route discovery is not required, there is strong support for using random routing for energy efficient networks instead of flooding.

REFERENCES

[1] Vamsi K Paruchuri, Arjan Durresi, Durga S Dash, Raj Jain, "Optimal Flooding Protocol for Routing in Ad-hoc Networks", IEEE Wireless Communications and Networking Conference, March 16-23, 2003.

[2] Zygmunt Hass, Joseph Y. Halpern, Li Li, "Gossip based Ad Hoc Routing", IEEE INFOCOM, June 2002.

[3] Reese T. Prosser, "Routing Procedures in Communications Networks – Part I: Random Procedures", IRE Transactions on Communications Systems, vol. 10, no. 4, December 1962 pp. 322-329.