COMPARISON OF REACTIVE ROUTING AND FLOODING IN WIRELESS SENSOR NETWORKS

Anne Saaranen
Nokia Research Center
P.O. Box 100
33721 Tampere
Finland

Carlos Pomalaza-Ráez
Centre for Wireless Communications
P.O. Box 4500
90014 University of Oulu
Finland

ABSTRACT

In recent years there have been numerous proposals for using reactive ad hoc routing protocols for wireless sensor networks. Reactive routing has been shown to perform fairly well in mobile ad hoc networks. However, since most of the proposed sensor networks have a very large number of nodes with an almost static topology, these protocols might not perform well. Moreover, changes in the sensor network topology are caused by nodes leaving the network once their batteries run out or after a malfunction rather than from movement. In this paper, through mathematical analysis, the use of reactive routing is compared with simple flooding of the sensed data throughout the network. The communication overhead of both approaches is evaluated, which gives a rough estimate of the energy consumption, an important design consideration in sensor networks.

I. INTRODUCTION

Currently, wireless sensor networks (WSNs) are the subject of numerous studies. In some of these studies WSNs are regarded as special cases of ad hoc networks. The most important characteristics of ad hoc networks are the lack of an underlying infrastructure, autonomous setup, and a dynamic topology. WSNs have the first two features, but their topology is not as dynamic as in traditional ad hoc networks. The sensor nodes are mostly stationary; changes in the topology result more from node failures than from movement.

Modified reactive ad hoc routing protocols have been proposed for sensor networks, however in [1, 2] it is shown that reactive protocols are at their best in a mobile environment. Therefore, it is quite possible that reactive routing might not perform well in sensor networks that have stationary nodes. In this environment proactive type routing protocols can give good throughput and generate less overhead.

In this study the worst case performance of reactive routing is compared with pure flooding of the sensed data. Flooding is not compared with proactive routing. The methodology for the latter type of comparison is different than for the case of reactive routing and thus it is a potential subject for future studies.

In the comparisons the overhead traffic generated by the two approaches is used as a metric. By overhead it is meant the amount of on demand routing packets or the extra transmissions of flooded data packets. These extra transmissions include those made by nodes not residing along the route from the originating sensor to the sink node. The impact of the data packet length and of the network size on the overhead traffic is also considered.

II. REACTIVE ROUTING

Reactive routing means that a route is searched only when needed. When a node has data packets addressed to a particular destination and there is no route to it, the node initiates a route discovery process. The node that initiates the route discovery is called an originator or a source node. Routes are typically searched using a route request message that is flooded throughout the entire network. The size of this message is relatively small, even though in some protocols the traversed route is aggregated to the route request message.

After the route request message has reached the destination node, which is the target of the route discovery, a route reply message is sent to the originator. The route reply message contains the IDs of all the nodes along the chosen route. After receiving this message the originator has a functional route to the destination and it can start the data transfer.

If a route breaks, a route error message is delivered from the node that detects the link failure to the originator node. When the originator receives this message, if necessary, it may initiate a new route discovery.

In the comparisons section the overhead generated by the two approaches is considered. In the case of reactive routing the overhead is all the traffic the routing protocol generates when finding or maintaining the routes. Because the cases in which the route breaks are excluded from these considerations, the overhead for reactive routing consists of flooded route requests and transmitted route reply messages.

III. COMPARISONS

In this section byte overhead is used as the evaluation metric. For a routing protocol byte overhead means the total number of bytes in the routing control messages needed to find a route to the sink. For flooding, byte overhead means the total number of bytes in the extra messages flooded throughout the network. In both cases the bytes in the data packets transmitted by nodes along the route from the originating node to the sink node are not counted as overhead.

First, the overhead of only one route leading from a sensor node to the sink is considered. Let N be the total number of nodes in the network. Let L be the length of the route from the originator node to the sink. The size of a routing protocol control message is denoted by S and the size of the data message by D. So, the amount of overhead (in bytes) generated by flooding the data packet throughout the entire network is

$$N \times D - L \times D = (N - L) \times D \tag{1}$$

The corresponding value for reactive routing is

$$N \times S + L \times S = (N + L) \times S \tag{2}$$

where $N \times S$ is the overhead generated by flooding the route requests and $L \times S$ is the overhead from the route reply messages. Using flooding is beneficial if the overhead of flooding is smaller than the overhead generated by reactive routing, i.e.

$$(N - L) \times D < (N + L) \times S \Rightarrow D < \frac{N + L}{N - L} \times S \tag{3}$$

This means that the size of a data message must be at most $\eta \equiv (N + L)/(N - L)$ times the size of a routing control message to justify usage of flooding instead of reactive route discovery. It is then sufficient to study the behavior of the coefficient η to find when flooding generates less overhead than a reactive routing protocol.

Figure 1 shows η as a function of the network size when the length of the route is 4 hops. It can be seen that for a linear topology (5 nodes in the network), the size of the data packet can be several times the size of the routing control message. However, for a more scattered network, where many of the nodes are not part of the route from the originator node to the sink node, the data packet size must be almost the same size as the routing control message.

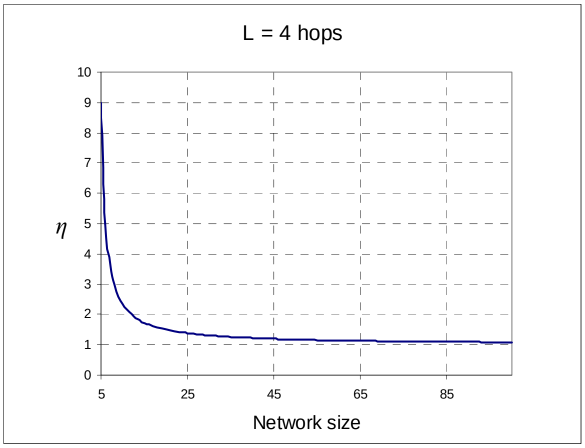

Figure 1. Ratio of data message length to routing control message length as a function of the network size.

Figure 2 and Figure 3 show how changing the route length affects the value of η. Figure 2 shows a segment of the curve shown in Figure 3. The results in Figure 2 confirm that for a large network, reactive routing generates less overhead if the route to the sink node is short, unless the data packet is almost as small as the routing control message. Figure 3 shows that in a distorted network topology, flooding is a feasible option even for large data messages. For example, in this network with 200 nodes in a nearly linear topology with a route of 160 hops, the data message can be 9 times the size of the routing control message. If the topology is even more linear with the sink node at one end and the sensor at the other end, routing only generates excess overhead without any real benefits. The overhead of sending the data packet to a few sensors not along the route is negligible. In this example with route length of 190 hops the data packet can be almost 40 times larger than the routing control message and flooding still generates less overhead than when using reactive routing.

The above observations illustrate the impact of the network size on an individual route.

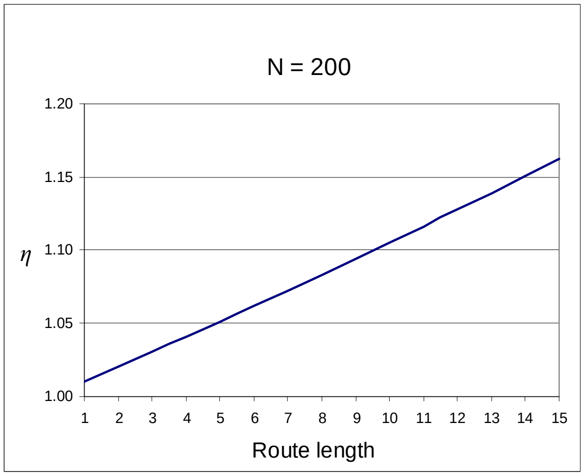

Figure 2. Ratio of data message length to routing control message length as a function of the route length.

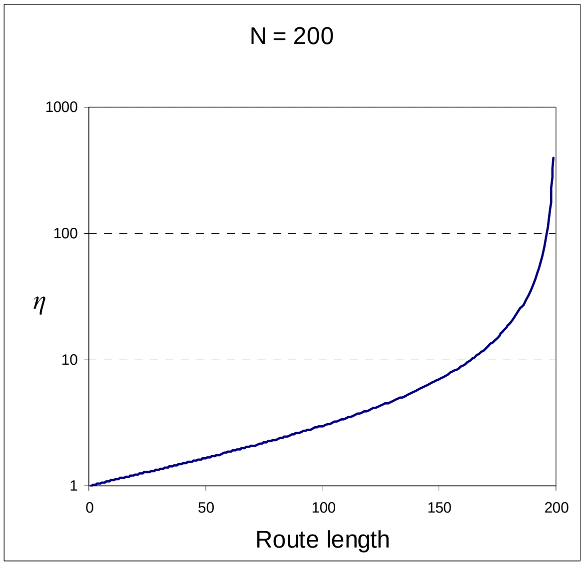

Figure 3. Ratio of data message length to routing control message length as a function of the route length.

Next, the maximum allowed relative size of the data message is evaluated, when the total network overhead is taken into account. For this case some assumptions are made about the routes' lengths. It is assumed that the network diameter is $\left\lceil \sqrt{N} \right\rceil$, where N is the total number of nodes in the network; and that the nodes are approximately uniformly distributed over the network area.

It is also assumed that the network area resembles a circle. The network area could also be assumed to be a square or some other geometrical form. The circle shape is selected here because the hop count from the sensors to the sink node is easier to estimate. To simplify the analysis the sink node is assumed to be in the center of the circle. If the sink node is placed on the border of the network area, the average hop count from the sensor nodes to the sink node increases and hence the data packet size can be larger than the maximum derived in the following analysis.

Figure 4 shows an example of a network area and the zones for different hop counts towards the sink in the center. The sensor nodes are assumed to all have the same transmission range, denoted by r. Since the network diameter is $\left| \sqrt{N} \right|$ the hop count from the sink node to the border of the network area is $\left\lceil \sqrt{N} \right\rceil / 2$. The route length, l, from a sensor to the sink node can then be $1, 2, \cdots, \left| \sqrt{N} \right| / 2$ hops.

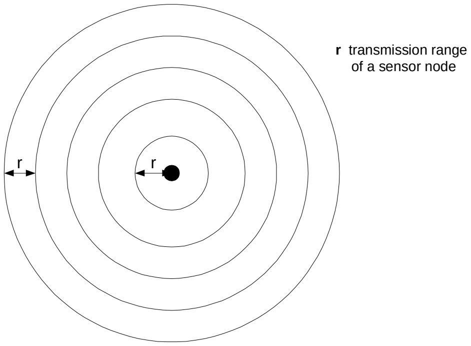

Figure 4. Network area and hop count from the sink node.

Next, the number of nodes that have the same number of hops to the sink node is estimated. This is done by comparing the area of the hop count zone to the total area of the network. The term hop count zone denotes the ring-shaped area that contains most of the nodes that are l hops away from the sink node. Figure 5 shows an instance of these hop count zones. Nodes inside the white area have the sink node within their transmission radius and hence are one hop away. The nodes in the yellow area are two hops away from the sink node; the nodes in the red area are three hops away, and so on. The area of the first hop zone is

$$A_r = \pi r^2$$

where r is the common transmission range of the sensor nodes. The area of the 2-hop zone is

$$A_{2r} = \pi(2r)^2 - A_r = \pi 4r^2 - \pi r^2 = 3\pi r^2$$

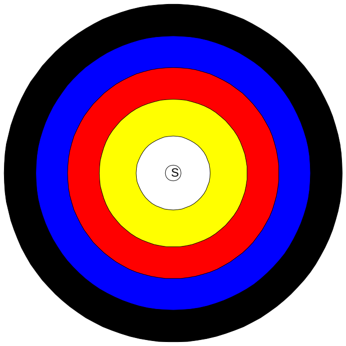

Figure 5. Hop count zones from the sink node, the darker the color the larger the hop count.

The equation for an l-hop zone area is,

$$\begin{aligned} A_{lr} &= \pi(lr)^2 - \pi[(l-1)r]^2 \\ &= \pi r^2 \left[l^2 - (l-1)^2 \right] = (2l-1)\pi r^2 = (2l-1)A_r \end{aligned} \qquad (3)$$

Since the diameter of the network area is the transmission range times the network diameter, $r \times \left| \sqrt{N} \right|$, the radius of the network area is $R = r \times \left\lceil \sqrt{N} \right\rceil / 2$. This means that the network area is:

$$A_R = \pi R^2 = \pi \left(\frac{r \left\lceil \sqrt{N} \right\rceil}{2} \right)^2 = \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 \pi r^2 = \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 A_r \qquad (4)$$

So, the fraction of nodes with l hops to the sink node is

$$\frac{A_{lr}}{A_R} = \frac{(2l-1) \times A_r}{\left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 \times A_r} = \frac{2l-1}{\left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2} = \frac{4(2l-1)}{\left\lceil \sqrt{N} \right\rceil^2} \qquad (5)$$

which means that there are

$$\frac{4(2l-1)}{\left\lceil \sqrt{N} \right\rceil^2} N = \frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} (2l-1) \qquad (6)$$

nodes having l hops to the sink node.

Now the overhead, if every node in the network sends one data message to the sink, can be estimated. Assuming that every node must find its own route i.e. every node must send a route request to the sink node. This gives an estimate for the worst case performance of the reactive routing approach. In practical implementations it is highly unlikely that there are many nodes that need a separate route discovery, since nodes can also learn routes by processing the route discovery requests of other sensor nodes.

The total overhead of flooding is the sum of the overhead generated by routes with length *l* hops, $l = 1, 2, \cdots, \left(\left\lceil \sqrt{N} \right\rceil \big/ 2\right)$. This overhead is then

$$\sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} (N-l) D \frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} (2l-1) = D \frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} (N-l)(2l-1)$$

This equation can be reduced to

$$D \frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} \left[\left(N + \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 + \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} - 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right] \tag{7}$$

The total overhead of reactive routing is

$$\sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} (N+l) S \frac{4 \times N}{\left\lceil \sqrt{N} \right\rceil^2} (2l-1) = S \frac{4 \times N}{\left\lceil \sqrt{N} \right\rceil^2} \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} (N+l)(2l-1)$$

This equation can be reduced to

$$S \frac{4 \times N}{\left\lceil \sqrt{N} \right\rceil^2} \left[\left(N - \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 - \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} + 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right] \tag{8}$$

Equations 7 and 8 can be used to compute the maximum size of the data packet for which flooding generates less overhead than reactive routing, e.g.

$$D < S \times \frac{\frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} \left[\left(N - \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 - \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} + 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right]}{\frac{4N}{\left\lceil \sqrt{N} \right\rceil^2} \left[\left(N + \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 + \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} - 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right]}$$

$$\Rightarrow D < S \frac{\left[\left(N - \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 - \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} + 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right]}{\left[\left(N + \frac{1}{2}\right) \left(\frac{\left\lceil \sqrt{N} \right\rceil}{2} \right)^2 + \frac{1}{2} \frac{\left\lceil \sqrt{N} \right\rceil}{2} - 2 \sum_{l=1}^{\frac{\lceil \sqrt{N} \rceil}{2}} l^2 \right]} \tag{9}$$

or $\Rightarrow D < S \times \eta$

For flooding to be a good choice the data packet size can be at most η times the size of the routing control message size. Figure 6 shows values of η as the function of the network size. Evidently, for most cases the data messages must be almost as small as the routing control messages to make flooding more efficient than reactive routing. Considering that it is not necessary for every node to find a route to the sink, the data messages must be even smaller to make flooding a better choice. The fact that sensor nodes using a reactive routing approach also store routes for future use also reduces the difference between these two approaches. In more sophisticated reactive routing protocols, nodes other than the sink can answer to route discoveries and thus the route request messages do not propagate any further. This optimization further reduces the overhead of the reactive routing approach. The impact of this optimization is difficult to evaluate because it is hard to estimate how far the route request message propagates before a node answers. This optimization may limit the route request only to the neighborhood of the initiator or the message may propagate throughout the entire network even when there are a few nodes nearby that know a route to the sink.

Note that flooding is a competitive choice in very small sensor networks, but in larger networks reactive routing is generally a better choice. The periodic peaks of the curve shown in Figure 6 are due to the quantization of the network diameter values.

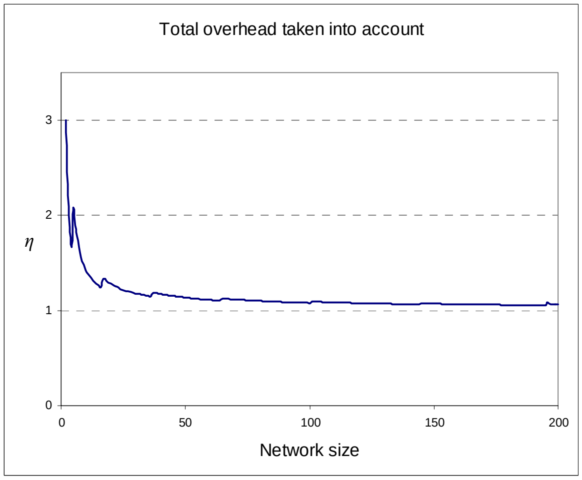

Figure 6. Ratio of data message length to routing control message length as a function of the network size.

IV. CONCLUSIONS

A comparison of reactive routing with simple flooding of the sensed data through a network has been carried out. This study concentrates on evaluating the overhead generated by these two methods and ignores other benefits of using simple flooding. These benefits such as simple implementation, small delay, and low processing requirements, were the initial reasons to carry out this study. Future work will try to evaluate other benefits and problems of both approaches.

For the network constraints considered, flooding is a better choice when the size of the data message is small, otherwise performing a route discovery process produces less overhead. The network topology is also a very important factor. If the sensors are uniformly scattered around the network area, reactive routing generates less overhead than flooding. However, if the sensors are in a "linear arrangement," flooding can be a good alternative for some configurations.

Both reactive routing and flooding of the data throughout the entire network can be used in the same network under an adaptive scheme. One such scheme could be the use of reactive routing for large data packets and flooding for small messages. However, if the route to the sink node is already known, it should also be used to send the small messages, flooding these throughout the entire network would be a waste of resources.

REFERENCES

[1] J. Broth, D. A. Maltz, D. B. Johnson, Y-C Hu, and J. Jetcheva, "A Performance Comparison of Multi-Hop Wireless Ad Hoc Networks Routing Protocols," Proceedings of ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98), October 1998.

[2] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, "Scenario-based Performance Analysis of Routing Protocols for Mobile Ad-hoc Networks," Proceedings of ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'99), August 1999.