A NEW ROUTING PROTOCOL FOR THE RECONFIGURABLE WIRELESS NETWORKS

Zygmunt J. Haas
School of Electrical Engineering, Cornell University, Ithaca, NY, 14853
http://www.ee.cornell.edu/~haas/wnl.html

ABSTRACT: In this paper, we propose a new routing protocol, the *Zone Routing Protocol (ZRP),* for the Reconfigurable Wireless Networks, a large scale, highly mobile ad-hoc networking environment. The novelty of the ZRP protocol is that it is applicable to large *flat-routed* networks. Furthermore, through the use of the *zone radius* parameter, the scheme exhibits adjustable hybrid behavior of *proactive* and *reactive* routing schemes. We evaluate the performance of the protocol, showing the reduction in the number of control messages, as compared with other reactive schemes, such as flooding.

INTRODUCTION

Recently, there has been an increased interest in ad-hoc networking [1]. In general, ad-hoc networks are network architecture that can be rapidly deployed, without preexistence of any fixed infrastructure. A special case of ad-hoc networks, the *Reconfigurable Wireless Networks (RWN),* was previously introduced [2,3] to emphasize a number of special characteristics of the RWN communication environment:

- large network coverage; large network radius, r_{net},
- large number of network nodes, and
- large range of nodal velocities (from stationary to highly mobile).[1]

In particular, the topology of the RWN is quite frequently changing, while self-adapting to the connectivity and propagation conditions and to the traffic and mobility patterns. Examples of the use of the RWNs are:

- military (tactical) communication - for fast establishment of communication infrastructure during deployment of forces in a foreign (hostile) terrain
- rescue missions - for communication in areas without adequate wireless coverage
- national security - for communication in times of national crisis, when the existing communication infrastructure is non-operational due to a natural disasters or a global war
- law enforcement - similar to tactical communication
- commercial use - for setting up communication in exhibitions, conferences, or sale presentations
- education - for operation of virtual classrooms
- sensor networks - for communication between intelligent sensors (e.g., MEMS) mounted on mobile platforms.

Basically, there are two approaching in providing ad-hoc network connectivity: *flat-routed* or *hierarchical* network architectures. An example of a flat-routed network is shown in Figure 1 and of a two-tiered hierar-

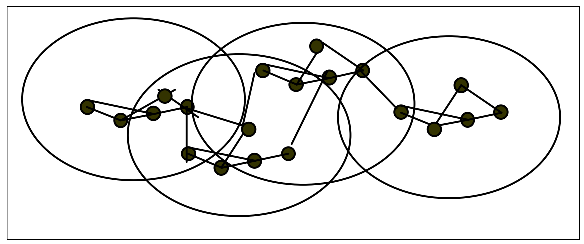

Figure 1: A flat-routed ad-hoc network

chical network in Figure 2. In flat-routed networks, all the nodes are "equal" and the packet routing is done based on peer-to-peer connections, restricted only by the propagation conditions. In hierarchical networks, there are at least two tiers; on the lower tier, nodes in geographical proximity create peer-to-peer networks. In each one of these lower-tier networks, at least one node is designated to serve as a "gateway" to the higher tier. These "gateway" nodes create the higher-tier network, which usually requires more powerful transmitters/receivers. Although routing between nodes that belong to the same lower-tier network is based on peer-to-peer routing, routing between nodes that belong to different lower-tier networks is through the gateway nodes.

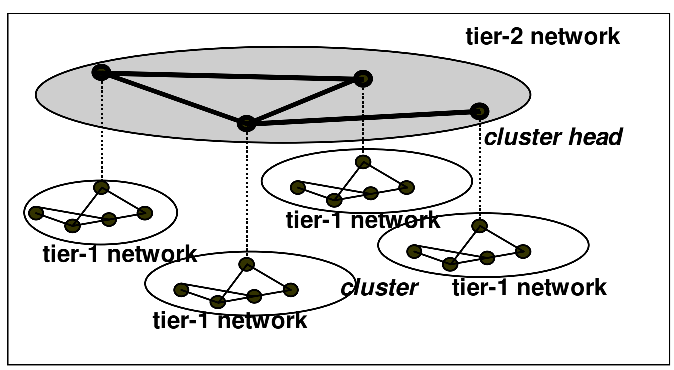

Figure 2: A two-tiered ad-hoc network

[1] For example, the maximal nodal velocity is such that the lifetime of a link can be between hundreds of milliseconds to few seconds only.

We will omit here the comparison of the two architectures. Nevertheless, we note that the flat-routed networks are more suitable for the highly versatile communication environment as the RWN-s. The reason is that the maintenance of the hierarchies (and the associated cluster heads) is too costly in network resources when the lifetime of the links is quite short. Thus, we chose to concentrate on the flat-routed network architecture in our study of the routing protocols for the RWN.

PREVIOUS AND RELATED WORK

The currently available routing protocols are inadequate for the RWN. The main problem is that they do not support either fast-changeable network architecture or that they do not scale well with the size of the network (number of nodes). Surprisingly, these shortcomings are present even in some routing protocols that were proposed for ad-hoc networks.

More specifically, the challenge stems from the fact that, on one hand, in-order to route packets in a network, the network topology needs to be known to the traversed nodes. On the other hand, in a RWN, this topology may change quite often. Also, the number of nodes may be very large. Thus, the cost of updates is quite high, in contradiction with the fact that updates are expensive in the wireless communication environment. Furthermore, as the number of network nodes may be large, the potential number of destinations is also large, requiring large and frequent exchange of data (e.g., routes, routes updates, or routing tables) between network nodes.

The wired Internet uses routing protocols based on topological broadcast, such as the OSPF [4]. These protocols are not suitable for the RWN due to the relatively large bandwidth required for update messages.

In the past, routing in multi-hop packet radio networks was based on shortest-path routing algorithms [5], such as Distributed Bellman-Ford (DBF) algorithm. These algorithms suffer from very slow convergence (the "counting to infinity" problem). Besides, DBF-like algorithms incur large update message penalty. Protocols that attempted to cure some of the shortcoming of DFB, such as Destination-Sequenced Distance-Vector Routing (DSDV) [6], were proposed and studied. Nevertheless, synchronization problems and extra processing overhead are common in these protocols. Other protocols that rely on the information from the predecessor of the shortest path solve the slow convergence problem of DBF (e.g., [7]). However, the processing requirements of these protocols may be quite high, because of the way they process the update messages.

Use of dynamic source routing protocol, which utilizes flooding to discover a route to a destination, is described in [8]. A number of optimization techniques, such as route caching are also presented that reduce the route determination/maintenance overhead. In a highly dynamic environment, such as the RWN is, this type of protocols lead to a large delay and the techniques to reduce overhead may not perform well.

A query-reply based routing protocol has been introduced recently in [9]. Practical implementation of this protocol in the RWN-s can lead, however, to high communication requirements.

A new distance-vector routing protocol for packet radio networks (WRP) is presented in [10]. Upon change in the network topology, WRP relies on communicating the change to its neighbors, which effectively propagates throughout the whole network. The salient advantage of WRP is the considerable reduction in the probability of loops in the calculated routes. The main disadvantage of WRP for the RWN is in the fact that routing nodes constantly maintain full routing information in each network node, which was obtained at relatively high cost in wireless resources

In [11], routing is based on temporary addresses assigned to nodes. These addresses are concatenation of the node's addresses on a physical and a virtual networks. However, routing requires full connectivity among all the physical network nodes. Furthermore, the routing may not be optimal, as it is based on addresses, which may not be related to the geographical locations, producing a long path for communication between two close-by nodes.

The above routing protocols can be classified either as *proactive* or as *reactive.* Proactive protocols attempt to continuously evaluate the routes within the network, so that when a packet needs to be forwarded, the route is already known and can be immediately used. Reactive protocols, on the other hand, invoke the route determination procedures on demand only. Thus, when a route is needed, some sort of global search procedure is employed.

The advantage of the proactive schemes is that, once a route is requested, there is little delay until route is determined. In reactive protocols, because route information may not be available at the time a routing request is received, the delay to determine a route can be quite significant. Because of this long delay, pure reactive routing protocols may not be applicable to real-time communication. However, pure proactive schemes are likewise not appropriate for the RWN environment, as they continuously use large portion of the network capacity to keep the routing information current. Since in an RWN nodes move quite fast, and as the changes may be more frequent than the routing requests, most of this routing information is never used! This results in an excessive waste of the network capacity. What is needed is a protocol that, on one hand, initiates the route-determination procedure on-demand, but with limited cost of the global search.

The introduced here routing protocol, which is based on the notion of *routing zones,* incurs very low overhead in route determination. It requires maintaining a small amount of routing information in each node. There is no overhead of wireless resources to maintain routing information of inactive routes. Moreover, it identifies multiple routes with no looping problems.

The *ZONE ROUTING PROTOCOL (ZRP)*

Our approach to routing in the RWN is based on the notion of a *routing zone,* which is defined for each node and includes the nodes whose distance (e.g., in hops) is at most some predefined number. This distance is referred to here as the *zone radius,* r_{zone}. Each node is required to know the topology of the network within its routing zone only and nodes are updated about topological changes only within their routing zone. Thus, even though a network can be quite large, the updates are only locally propagated. Since for radius greater than 1 the routing zones heavily overlap, the routing tends to be **extremely robust.** The routes within the network are specified as a sequence of nodes separated by approximately the zone radius.

We illustrate the *Route Discovery* protocol by an example in Figure 3. To allow source *S* to send a packet to destination *D,* a route from *S* to *D* needs to be determined. First, *S* verifies that *D* is not within its routing zone (to recall, each node knows all the nodes within its routing zone). Then, *S* sends a query to all the nodes on the periphery of its zone; i.e., *C, G,* and *H.*

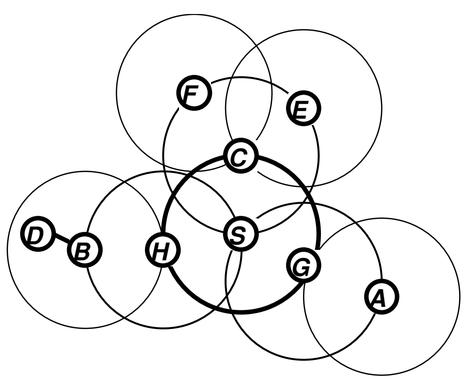

Figure 3: An example of Zone Routing

Now, in turn, each one of these nodes, after verifying that *D* is not in their routing zone, broadcast the query to their "peripheral" nodes. In particular, *H* sends the query to *B,* which recognizes *D* as being in its routing zone and responds to the query, indicating the forwarding path: *S-H-B-D.*

The mechanism by which *B* learns about the forwarding path is the *Route Accumulation.* Route Accumulation is a simple protocol by which each node that forwards the query appends its identity to the query message. In order to limit the message size and to bound the Route Discovery process, a *hop-count* is included within the query messages. The value of the hop-count in the initial query message is set to some maximal value, $hop_{\max}$. The value of the hop count is decreased by one, each time a query message is forwarded. When the hop-count reaches zero, the copy of the query message is discarded.

If the destination node is within maximum hop-count from the source node, the algorithm will discover at least one path between the two nodes, no matter what the value of the zone radius is. Because of space limitations, we omit the complete proof of the protocol.

The means by which each node learns about the topology of its zone is through any proactive algorithm. For example a "truncated" version of DSDV is possible, in which the reachability updates propagate only within distance limited by the zone radius.

Note that the ZRP requires only a relatively small number of query messages, as these messages are routed only to "peripheral" nodes, omitting all the nodes within the routing zones. As the zone radius is significantly smaller than the network radius, the cost of learning the zones' topologies is a very small fraction of the cost required by a global proactive mechanism. Furthermore, the amount of data stored at each node is similarly reduced. On the other hand, ZRP is much faster than a global reactive route discovery mechanisms, as the number of nodes queried in the process is on the order of $\left(\frac{r_{zone}}{r_{net}}\right)^2$ of the number of nodes queried by a global flooding process. Additionally, ZRP discovers multiple routes to the destination.

The Route Discovery process in ZRP can be made much more efficient in resources, at the expense of longer latency. Instead of querying simultaneously all the "peripheral" nodes at the boundary of the routing zone, these nodes can be queried either sequentially, one-by-one, or in groups. Thus, there is a tradeoff between the cost and latency of the ZRP Route Discovery protocol.

Finally, we note that the ZRP path, which consists of nodes spaced approximately by distance of zone radius, is more stable than a full path that includes all the nodes between the source and the destination. As the nodes move and links are frequently broken, the ZRP path is more stable than the full path.

The behavior of the ZRP can be adjusted by changing the value of r_{zone}. In particular, for large zone radius, the coverage area is a single zone and ZRP is a traditional proactive protocol. For small zone radius, the protocol is more reactive, and become pure flooding at zone radius of one.

Two other ZRP related protocols work in conjunction with the Route Discovery protocol: *Route Evolution* and *Routing Zone Update* protocols. The Route Evolution protocol changes the network-wide routes in response to changes in the connectivity status of the nodes on a path. The Route Zone Update protocol, which is based on the previously mentioned Route Accumulation protocol, allows each node to **efficiently** learn the complete topology of its zone. Finally, optimizations to the ZRP protocol can be incorporated as well, such as route caching and route eavesdropping. We omit here the description of these mechanisms.

NETWORK, TRAFFIC, AND MOBILITY MODELS

We assume that the mobile hosts are distributed randomly in a closed coverage area, as displayed in Figure 4, effectively creating a torus.

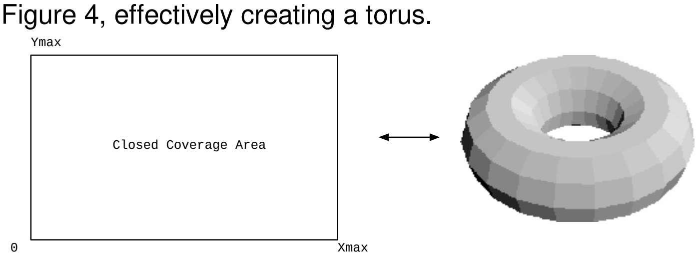

Figure 4: Coverage area mapped into a closed surface

Thus, for example, a mobile that "exits" the coverage area from the left side, appears as reentering the coverage area on the right side with the same velocity and the same direction. The distance between two mobiles located at (x_1, y_1) and (x_2, y_2) is, therefore:

$$D=\sqrt{\left[\min\left(|x_1-x_2|,|X_{\max}-x_1+x_2|\right)\right]^2+\left[\min\left(|y_1-y_2|,|Y_{\max}-y_1+y_2|\right)\right]^2}$$

The mobiles distributed in the coverage area can roam freely about it, according to the following mobility model. The movement of each mobile host in the coverage area is characterized by its velocity vector $\vec{v}=(v,\theta)$, where v is the mobile's speed and θ is its direction, measured with respect to the positive x-axis. The position of the mobile, (x, y) and its velocity $\vec{v}$ are updated periodically, every Δt time units as follows:

$$v(t+\Delta t)=\min\left[\max\left(v(t)+\Delta v,0\right),V_{MAX}\right],$$
$$\theta(t+\Delta t)=\theta(t)+\Delta\theta,$$
$$x(t+\Delta t)=x(t)+v(t)\cdot\cos\theta(t),$$
$$y(t+\Delta t)=y(t)+v(t)\cdot\sin\theta(t),$$

where V_{MAX} is the maximal mobile velocity (assumed here to be 65[mph]), Δv, the velocity change, is uniformly distributed within $\left(-A_{MAX}\cdot\Delta t, A_{MAX}\cdot\Delta t\right)$, A_{MAX} is the maximum acceleration/deceleration of the mobile (taken here to be $0.9\left[meter/\sec^2\right]$), and $\Delta\theta$, the change in the mobile's direction, is uniformly distributed in $(-\alpha\cdot\Delta t,\alpha\cdot\Delta t)$, where α is the maximal angular change of the mobile's direction per unit time (taken here to be $0.1745\left[radian/\sec\right]$).

Existence of a link between two mobiles separated by distance D is modeled as a stochastic random variable with one-sided normal probability density function; i.e., the probability of existence of a link is:

$$P(D)=\sqrt{\frac{2}{\pi\sigma^2}}\int_D^\infty \exp\left(-x^2\big/2\sigma^2\right)\cdot dx,$$

where we assumed in our evaluation that $\sigma=7.5$.

We assume a connectionless network layer, in which arriving packets carry the identification of the destination node. Packet destinations are chosen randomly among all the network nodes (other than the source itself).

THE SIMULATION

In our simulation, we assume that each node is equipped with a single half-duplex transceiver. Packets that cannot be sent (because, for example, the next node is busy) are buffered for future transmission. Our MAC is a simple *busy tone* based scheme: each "payload" channel is associated with a low-capacity control channel. When a node is busy, it emits a busy tone on its control channel. A node that wants to communicate with another node sends a request on the other node's control channel. The other nodes than acknowledges the connection and starts emitting the busy tone.

When a new packet is generated by a node, the routing protocol is initiated. The routing protocol takes into consideration only the existence of links between nodes and not the state of the nodes (i.e., whether the node is busy or idle). If a path between the source and the destination exists, the packet is queued for transmission along the discovered path in the *store-and-forward* manner. A packet is *blocked* if no such path exists. If any link on the path is broken due to mobility, the transmitted packet is considered *dropped* and lost.

We have compared our *Zone Routing Protocol (ZRP)* with a *flooding* algorithm and with the *Dynamic Destination-Sequence Distance-Vector (DSDV) Routing.* Here, we present the comparison of the number of required control messages for a route discovery in the ZRP protocol and the flooding protocol. We first describe the implementation of the flooding algorithm.

To determine a path in the flooding algorithm, a query message is broadcasted to all the source's neighbors. If the query was seen before by a receiving node, no action is taken. Otherwise, the query is rebroadcasted to all the neighbors (excluding the neighbor from whom the query was received), unless the node is the sought destination, in which case a reply is generated and sent back to the source. The process terminates by itself, as there is finite number of nodes in the network and once a query is seen by a node, no action is taken. Similarly to the ZRP, identity of the broadcasting node is added to the query packets. Thus, upon receiving the query, the destination node extracts the path to the source by reversing the order of the visited nodes identities in the query packet. Finally, to limit the scope of the flooding, a maximal *hop count* field is inserted into the original query packet and is decremented at each visited node. When it reaches zero, the query packet is discarded.

In simulating our *Zone Routing Protocol,* we assume that the MAC scheme provides immediate neighbor connectivity information. We assume that each node continuously learns the topology of its zone through a derivative of the Bellman-Ford algorithm. Finally, the *Route Accumulation* procedure is used to register the route in the query packet.

The graph in Figure 5 shows the number of the control packets for both, the ZRP and pure flooding, as a function of nodal transmission radius, $r_{transmit}$, for 10[km]X10[km] coverage area, networks of 10, 20, and 30 users, $r_{zone} = 2$, and $hop_{\max} = 5$. Clearly, the ZRP requires only a small fraction of the flooding control messages, especially for large $r_{transmit}$.

ACKNOWLEDGMENT

The author thanks Mr. Messay Amerga for implementing the simulation described in this paper and for collecting the numerical results.

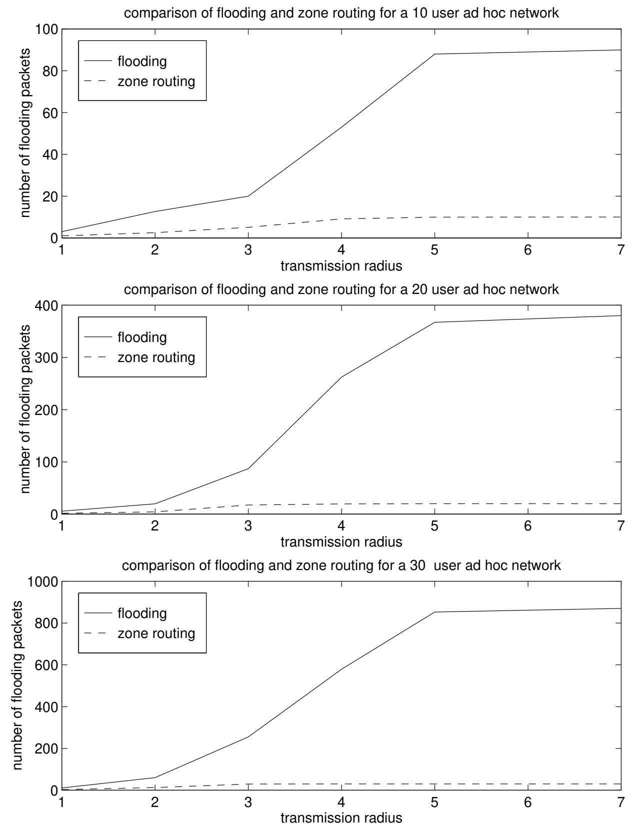

Figure 5: Comparison of the number of control messages

REFERENCES

[1] ARO/DARPA Workshop on Mobile Ad Hoc Networking, *University of Maryland, Institute for Systems Research,* March 14, 1997.

[2] Z. J. Haas, "The Relaying Capability of the Reconfigurable Wireless Networks," *VTC'97,* Phoenix, AZ, May 4-7, 1997.

[3] Z. J. Haas, "The MAC for the Reconfigurable Wireless Networks," *Sixth WINLAB Workshop on Third Generation Wireless Information Networks,* New Brunswick, NJ, March 20-21, 1997.

[4] J. Moy, "OSPF Version 2, *RFC 1583,* March 1994.

[5] B.M. Leiner, D.L. Nielson, and F.A. Tobagi, "Issues in Packet Radio Network Design," *Proceedings of the IEEE,* vol.75, January 1987.

[6] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers," *ACM SIGCOMM,* vol.24, no.4, Oct. 1994.

[7] C. Cheng, R. Reley, S.P.R. Kumar, and J.J. Garcia-Luna-Aceves, "A Loop-Free Extended Bellman-Ford Routing Protocol without Bouncing Effect," *ACM Computer Communications Review,* 19(4), 1989, pp.224-236.

[8] D. B. Johnson and D.A. Maltz, "Dynamic Source Routing in Ad-Hoc Wireless Networks," in *Mobile Computing,* T. Imielinski and H. Korth, editors, Kluwer, 1996.

[9] M. S. Corson and A. Ephremides, "A Distributed Routing Algorithm for Mobile Wireless Networks," *ACM J. of Wireless Networks,* Jan. 1995.

[10] S. Murthy and J.J. Garcia-Luna-Aceves, "A Routing Protocol for Packet Radio Networks," Proc. of ACM Mobile Computing and Networking Conference, *MOBICOM'95,* Nov. 14-15, 1995.

[11] J. Sharony, "A Mobile Radio Network Architecture with Dynamically Changing Topology Using Virtual Subnets, in Proc. of *ICC'96.*