8

Evaluation of the Ad-Hoc Connectivity with the Zone Routing Protocols

Zygmunt J. Haas and Marc R. Pearlman
School of Electrical Engineering, Cornell University,
323 Rhodes Hall, Ithaca, NY, 14853
Tel: (607) 255-3454, fax: (607) 255-9072, e-mail: *haas@ee.cornell.edu*
URL: *http://www.ee.cornell.edu/~haas/wnl.html*

Abstract

*In this paper, we evaluate the novel routing protocol for a special class of ad-hoc networks, termed by us the **Reconfigurable Wireless Networks (RWNs).** The main features of the RWNs are: the increased mobility of the network nodes, the large number of nodes, and the large network span. We argue that the current routing protocols do not provide a satisfactory solution for routing in this type of an environment. We propose a scheme, coined the **Zone Routing Protocol (ZRP),** which dynamically adjusts itself to the operational conditions by sizing a single network parameter - the **Zone Radius.** More specifically, the ZRP reduces the cost of frequent updates of the constantly changing network topology by limiting the scope of the updates to the immediate neighborhood of the change – the Zone Radius. We study the performance of the scheme, evaluating the average number of control messages required to discover a route within the network. Furthermore, we compare the scheme's performance, on one hand, with reactive flooding-based schemes, and, on the other hand, with proactive distance-vector schemes.*

1. Introduction

A Reconfigurable Wireless Network (RWN) is an *ad-hoc network* architecture that can be rapidly deployed without relying on preexisting fixed network infrastructure. The nodes in a RWN can dynamically join and leave the network, frequently, often without warning, and without disruption to other nodes' communication. Finally, the nodes in the network can be highly mobile, thus rapidly changing the node constellation and the presence or absence of links. Examples of the use of the RWNs are:

- tactical operation - for fast establishment of military communication during the deployment of forces in unknown and hostile terrain;
- rescue missions - for communication in areas without adequate wireless coverage;
- national security - for communication in times of national crisis, where the existing communication infrastructure is non-operational due to a natural disaster or a global war;
- law enforcement - for fast establishment of communication infrastructure during law enforcement operations;
- commercial use - for setting up communication in exhibitions, conferences, or sale presentations;
- education - for operation of wall-free (virtual) classrooms; and
- sensor networks - for communication between intelligent sensors (e.g., MEMS[2]) mounted on mobile platforms.

Nodes in the RWN exhibit nomadic behavior by freely migrating within some area, dynamically creating and tearing down associations with other nodes. Groups of nodes that have a common goal can create formations (clusters) and migrate together, similarly to military units on missions or

♥ This work is supported by the US Air Force/Rome Labs, under the contract number C-7-2544 and a grant from Motorola Corporation, the Applied Research Laboratory.

[2] Micro-Electro-Mechanical-Systems

similarly to guided tours on excursions. Nodes can communicate with each other at anytime and without restrictions, except for connectivity limitations and subject to security provisions. Examples of network nodes are pedestrians, soldiers, or unmanned robots. Examples of mobile platforms on which the network nodes might reside are cars, trucks, buses, tanks, trains, planes, helicopters, ships, UAV-s[3], or UFO-s.

In this paper, we concentrate on the issue of designing a routing protocol for the RWN. In particular, we address routing in a *flat* ad-hoc networks, as opposed to *hierarchical* ad-hoc networks that have been investigated in the past (e.g., [Lauer86,Westcott84]). The proposed protocol, the *Zone Routing Protocol,* allows efficient and fast route discovery in the RWN communication environment (i.e., large geographical network size, large number of nodes, fast nodal movement, and frequent topological changes). In what follows, we explain the elements of the proposed scheme.

2. Previous and Related Work

In this work, we address routing in a *flat* ad-hoc networks, as opposed to *hierarchical* ad-hoc networks that have been investigated in the past (e.g., [Lauer86,Westcott84]). Although routing in hierarchical ad-hoc networks involves simpler procedure, some salient features of the flat architectures, as mentioned above, make them that much more attractive for communication in the RWN environment. Comparison of the two architectures is outside the scope of the paper and the reader is referred to [Haas98] for further discussion on this topic.

The wired Internet uses routing protocols based on topological broadcast, such as the OSPF [Moy97]. These protocols are not suitable for the RWN due to the relatively large bandwidth required for update messages.

Routing in multi-hop packet radio networks was based in the past on shortest-path routing algorithms [Leiner87], such as Distributed Bellman-Ford (DBF) algorithm [Bertsekas92]. These algorithms suffer from very slow convergence (the “counting to infinity” problem). Besides, DBF-like algorithms incur large update message penalties. Protocols that attempted to cure some of the shortcomings of DBF, such as Destination-Sequenced Distance-Vector Routing (DSDV) [Perkins94], were proposed. However, synchronization 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., [Cheng89] and [Garcia-Luna-Aceves93]). However, the processing requirements of these protocols may be quite high, because of the way they process the update messages.

Routing protocols that are based on a source initiated query-reply process have also been introduced. Such techniques typically rely on the *flooding* of queries to discover a destinatio. In [Corson97] the route replies generated are also flooded, in a controlled manner, to distribute routing information in the form of directed acyclic graphs (DAGs) rooted at each destination. In contrast, other schemes *unicast* the route reply back to the querying source, typically by means of reversed routing information gathered during the query phase. In the case of [Perkins97], this routing information is in the form of next-hop routes to the querying node, while in [Johnson96], a route accumulation procedure is employed during the route query, allowing the route reply to be returned via source routing. The on-demand discovery of routes can result in much less traffic than standard distance vector or link state schemes, especially when innovative route maintenance schemes are employed. However, the reliance on flooding may still lead to considerable control traffic in the highly versatile RWN environment.

[Murthy95] and [Murthy] present a new distance-vector routing protocol for packet radio networks (WRP). Upon a 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, as compared with

[3] Unmanned Aerial Vehicles

other known routing algorithms, such as, for example, DBF. Compared with our routing protocol, the main disadvantage of WRP 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. Our protocol, in contrast, rapidly finds routes, only when transmission is necessary. Moreover, multiple routes are maintained, so that when some of these routes become obsolete, other routes can be immediately utilized. This is especially important when the network contains large number of very fast moving nodes, as is the case in the RWN architecture.

3. The Notion of a Routing Zone and Intrazone Routing

A *routing zone* is defined for each node and includes the nodes whose minimum distance in hops from the node in question is at most some predefined number, which is referred to here as the *zone radius*. An example of a routing zone (for node *S)* of radius 2 is shown in Figure 1.

Note that in this example nodes A through K are within the routing zone of S. Node L is outside *S'*s routing zone. *Peripheral nodes* are nodes whose minimum distance to the node in question is equal exactly to the zone radius. Thus, in Figure 1, nodes G-K are peripheral nodes. Zones of different nodes overlap heavily.

Related to the definition of a zone is the coverage of a node's transmitter, which is the set of nodes that are in direct communication with the node in question. These nodes are referred to as *neighbors*. The transmitter's coverage depends on the propagation conditions, on the transmitter power, and on the receiver sensitivity. In our simulation, we define conceptually a radius, d_{xmit}, which is the maximal distance that a node's transmission will be received without errors. Of course, it is important that each node be connected to at least one other node. However, more is not, necessarily, better. As the transmitter's coverage includes all the nodes with distance 1 hop from the node in question, the larger the d_{xmit} is, the larger is the content of its routing zone. A large routing zone requires large amount of update traffic.

For the purpose of simplification, we will depict zones as circles around the node in question. However, one should keep in mind that the zone is not a description of distance, but rather nodal connectivity (measured in hops).

Each node is assumed to maintain the routing information to all nodes that are within its routing zone and those nodes only. Consequently, in spite of the fact that a network can be quite large, the updates are only locally propagated. We assume that the protocol through which a node learns its zone is some sort of a proactive scheme, which we refer to here as the *IntrAzone Routing Protocol (IARP)*. In this paper, we use a modification of the Distance Vector algorithm. However, any other proactive scheme would do. Of course, in principle, the performance of the ZRP depends on the choice of IARP. However, our experience suggests that the tradeoffs are not strongly affected by the particular choice of the proactive scheme used.

3.1 Interzone Routing and the *Zone Routing Protocol*

IARP finds routes within a zone. The *IntErzone Routing Protocol (IERP),* on the other hand, is responsible for finding routes between nodes located at distances larger than the zone radius. IERP relies on what we call *bordercasting*. Bordercasting is a process by which a node sends a packet to all its peripheral nodes. A node knows the identity of its peripheral nodes by the virtue of the IARP. Bordercasting can (and should) be implemented by multicasting, if multicasting is supported within the subnet.[4] Alternatively, unicasting the packet to all the peripheral nodes achieves the same goal, albeit at much higher cost in resources.

[4] It is not clear whether multicasting is, indeed, feasible in a highly dynamic network topology. Examination of applicability of multicasting in ad-hoc networks is outside the scope of this paper. Here, we assume that bordercasting is performed using unicasting.

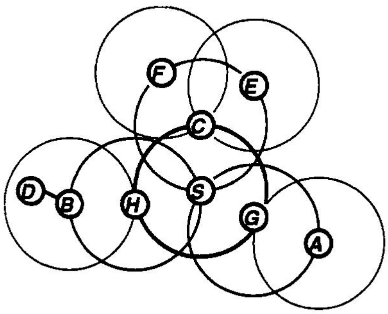

Figure 2: An example of IERP operation

The IERP operates as follows: The source node first checks whether the destination is within its zone.[5] If so, the path to the destination is known and no further route discovery processing is required. If the destination is not within the source Routing Zone, the source bordercasts a *route request* (which we call simply a *request)* to all its peripheral nodes.[6] Now, in turn, all the peripheral nodes execute the same algorithm: check whether the destination is within their zone. If so, a *route reply* (which we call simply a *reply)* is sent back to the source indicating the route to the destination (more about this in a moment). If not, the peripheral node forwards the query to its peripheral nodes, which, in turn, execute the same procedure.

An example of this *Route Discovery* procedure is demonstrated in Figure 2. The source node *S* sends a packet to the destination *D.* To find a route within the network, *S* first checks whether *D* is within its routing zone. If so, *S* knows the route to node *D.* Otherwise, *S* sends a query to all the nodes on the periphery of its zone; that is, to nodes *C, G,* and *H.* Now, in turn, each one of these nodes, after verifying that *D* is not in its routing zone forwards the query to its "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.*

A nice feature of this distributed route discovery process is that a single route query can return multiple route replies. The quality of these returned routes can be determined based on hop count (or any other path metric[7] accumulated during the propagation of the query. The best route can be selected based on the relative quality of the route (e.g., choose the route with the smallest hop count, or shortest accumulated delay).

Two main issues need to be addressed: When sending a reply to the source node, how does the "last peripheral node" know the whole path, to be included in the reply to the source? (A related question is, how does the responding node know how to send the reply to the source?) The second question is, how does the IERP process terminate?

Let us start with the first question. The process by which the node receiving a query knows the path back to the source of the query is the *Route Accumulation* procedure. In the Route Accumulation procedure, each node that forwards a query writes into the query packet its identification. The sequence of these identifications represents a path from the source node to the current node, and, by reversing the order, a path from the current node to the source node. Thus, the routes within the network are specified as a sequence of nodes, separated by approximately the zone radius. A node, which identifies that the destination is in its zone, simply adds its own identification to the query and returns the accumulated route to the source.

The second issue, that of termination of the IERP process, is a more difficult one. Of course, similarly to the standard flooding algorithm, a node that previously received the query will discard it. This, however, does not solve the whole problem, since as the zones heavily overlap, the query will be forwarded to many network nodes. In fact, it is very possible that the query will be forwarded to all the network nodes, effectively flooding the network. But a more disappointing result is that, due to fact that bordercasting involves sending the query over a path of length equal to the zone radius, the IERP

[5] Remember that a node knows the identity, distance to, and a route to all the nodes in its zone.

[6] Again, the identity of its zone peripheral nodes are known to the node in question.

[7] Typical path metrics include hop count, delay, capacity, etc.

will result in much more traffic than the flooding itself! What is needed is a more efficient termination criterion.

Let us look at this problem more closely. The idea behind IERP is that the search for a node advances in the "quantum" of zone radius, instead of flooding the network by forwarding the query among neighbors. The gain that we expect is due to the fact that only some network nodes will be involved in such a "flood." The challenge is to "steer" the search in the direction outwards of the original Routing Zone (see Figure 3), rather than going back into the areas that were already covered by other threads of the search. There are a number of ways that such a redirection of the search could be accomplished. We discuss here two possibilities. The first improvement, termed the *Backwards Search Prevention (BSP),* makes sure that peripheral nodes of the current node that lie within the routing zone of the previous node are not included in the next bordercast. To prevent the backward propagation of queries, a bordercasting node must send its queried peripheral nodes a list of its routing zone nodes (perhaps appended to the IERP query packet). Thus, in the example in Figure 4, after S bordercasts to the nodes F and C, the nodes A, C, and S will not be included in the consecutive bordercast by node F, as the nodes A, C, and S are all within the routing zone of the previous bordercasting node S. While the BSP may be impractical for large routing zones (due to the long list of routing zone nodes), it could be quite effective when used by nodes which maintain smaller routing zones.

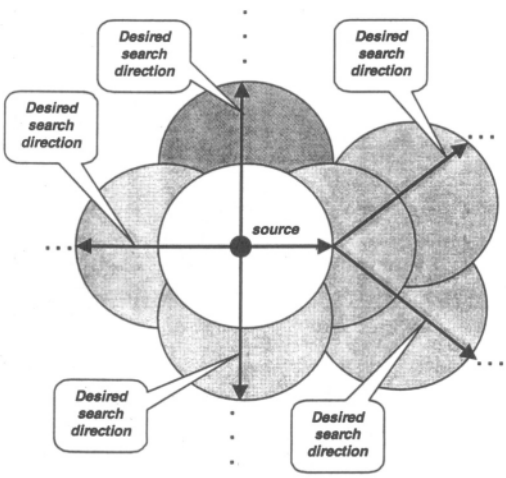

Figure 3: Guiding the search in desirable directions

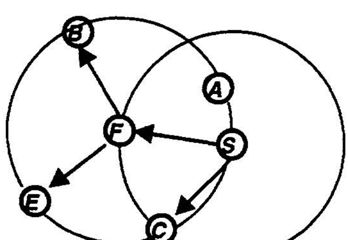

Figure 4: Preventing backwards IERP search

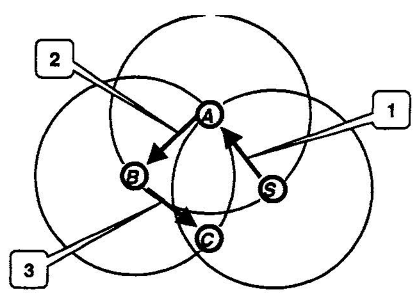

Figure 5: Preventing loopback IERP search

The second improvement, which we call the *Loopback Search Prevention (LSP),* involves pruning any search that goes into areas previously searched. This is accomplished by terminating bordercast at nodes that either have received the query before or that have overheard the query transmitted by their neighbors.[8] Note that this includes all threads of the query. Additional modification include termination of a thread at nodes whose routing zone include any of the nodes in the currently accumulated path. An example is shown in Figure 5. In this example, S bordercasts to A, which bordercasts to B, which in turn bordercasts to C. C will terminate the search of this thread (i.e., will stop bordercasting), as S, who is in the thread's accumulated route, is within its routing zone.

Both of these improvements reduce the amount of control traffic of the IERP protocol. From our simulation runs, we have learned that the contribution of the two schemes in reducing the control traffic is approximately equal. Note that the two techniques are not overlapping. The Backwards Search Prevention avoids sending a route request to nodes that should not forward it. On the other hand, the Loopback Search Prevention would send the route request to such a node, but will subsequently terminate the search thread at this node.

The main advantage of the *Zone Routing* Protocol is in the fact that the number of "flood" messages to discover a route is significantly smaller, as compared with other reactive-type protocols. This decrease is due to the directed propagation of queries to specified peripheral nodes. Since for radius greater than one the routing zones heavily overlap, the routing tends to be extremely robust.

Zone Routing, as described earlier, discovers multiple routes to a destination. However, the Route Discovery process can be made much more efficient in resources, at the expense of longer latency. This could be done by sequentially, rather than simultaneously, querying the peripheral zone nodes, either one-by-one or in groups. Thus, there is a tradeoff between the cost and the latency of the Route Discovery procedure.

We omit here correctness proof of the ZRP. Am interested reader is referred to [Haas98-2].

3.2 The *Route Maintenance* Procedure

In the Route Discovery procedure, each node, proactively and continuously learns the topology within its zone radius and, reactively, on-demand, discovers routes by hopping in steps of the routing radius. Because the number of nodes within a zone is much smaller than the number of network nodes, the penalty for dissemination of routing information within a zone is limited. So is the cost of the route discovery, when the zone radius is sufficiently large. For a small radius (zone radius =1), the ZRP behaves as a reactive scheme (flooding). On the other extreme, for a large radius **(zone radius = ∞)**, the scheme exhibits proactive behavior. In general, the size of the zone radius determines the ratio between the proactive and reactive behavior of the protocol. The *Route Maintenance Procedure* adaptively adjusts the zones' radii, as to reduce the "cost" of the Route Discovery Procedure.

The adjustment may be performed, based on the value of Call-to-Mobility-Ratio (CMR) measured independently at each node. CMR is a ratio of the rate at which queries are initiated to the rate at which connections with the neighbors are broken. Large CMR indicates that the network mobiles are very active in connection initiation and, thus, larger zone radius would decrease its frequent route discovery costs. Small CMR suggests that mobiles rarely place outgoing connections and, to reduce the overall cost of learning the routing within the nodes' routing zones, a smaller zone radius is preferable. Similarly, for fast moving mobiles (small CMR) the local zone routing information becomes obsolete quickly. Thus, a smaller zone radius carries smaller penalty. The routing zone radii may be configured prior to network deployment, based on *a priori* knowledge of network call activity and mobility patterns. Alternatively, and more typically, the routing zones may be resized dynamically, allowing the ZRP to adapt to local changes in call activity or node mobility.[9]

The Route Maintenance Procedure also significantly reduces the routing costs by employing the Route Discovery procedure only when there is a substantial change in the network topology. More specifically, active routes are cached by nodes: the communicating end nodes and intermediate nodes.

[8] This is done by having each node eavesdropping on all its neighbor communications and requires that the IERP communicate with the MAC layer.

[9] Dynamic adjustment of the routing zone radius requires minor modifications to the basic Zone Routing Protocol. Discussion of these enhancements is outside the scope of this paper.

Inactive paths are purged from the caches after some timeout period.[10] Upon a change in the network topology, such that a link within an active path is broken, a local path repair procedure is initiated. The path repair procedure substitutes a broken link by a mini-path between the ends of the broken link. A path update is then generated and sent to the end points of the path. Path repair procedures tend to reduce the path optimality (e.g., increase the length for shortest path routing). Thus, after some number of repairs, the path end points will initiate a new Route Discovery procedure to replace the path with a new optimal one.

4.0 Evaluation of the ZRP

We use the OPNET™ Network Simulator from MIL3, an event driven simulation package, to evaluate the performance of the ZRP over a range of routing zone radii, from reactive routing $(\rho=1)$ to proactive routing $(\rho\rightarrow\infty)$. Performance is gauged by measuring the control traffic generated by the ZRP and its effects on the average session delay. Our results can be used to determine the optimum ZRP routing zone radius for a given nodal velocity and for a given route query rate.

The ZRP control traffic consists of the intrazone (IARP) route update packets and the interzone (IERP) route request/reply/failure packets. While the neighbor discovery beacons could be considered control overhead, this additional traffic is independent of both mobile velocity and routing zone radius. Furthermore, the neighbor discovery process is not an exclusive component of the ZRP; various MAC protocols are also based on neighbor discovery. As such, the beacons do not contribute to the relative performance of the ZRP and are not accounted for in our analysis. Because the IERP packets are variable length (due to the route accumulation procedure), we measure control traffic in terms of node ID fields, rather than packets.

A meaningful measure of ZRP delay is the *average route query response time* $(\bar{T}_{rqr})$, which is defined as the average duration from the time a route is initially requested by the *Network* layer until the route is discovered.[11] If the destination appears in the routing tables (which will occur with probability (1–Prob[route discovery])), the query is immediately answered and the route query response time is assumed to be zero.[12] Otherwise, a route discovery is required (which will occur with probability (Prob[route discovery])) and the route query response time is measured as the time elapsed between the generation of the route request and the reception of the first route reply, $T_{route-reply}$.

$$\bar{T}_{rqr} = (1-\text{Prob}[\text{route discovery}])\cdot 0 + \text{Prob}[\text{route discovery}]\cdot T_{route-reply}$$

$$= \text{Prob}[\text{route discovery}]\cdot T_{route-reply}$$

For a fixed network size and fixed nodal density, the probability of a route discovery for an initial query is only dependent on the routing zone radius. The behavior of the route reply time is far more complicated. Not only is it dependent on the arrival rate of control packets, it is also affected by such factors as the network traffic load and the average length of IERP control packets. Our study provides some insight into the effect of these factors on the ZRP delay.

Our simulated RWN consists of 52 mobile nodes, whose initial positions are chosen from a uniform random distribution over an area of 600 [m] by 600 [m]. Each node j moves at a constant

---

[10] The determination of what constitutes an "active" or “inactive” path depends on the CMR of the network nodes. A cache management algorithm that determines the path activity is outside the scope of this paper.

[11] This delay metric does not reflect the delays associated with subsequent route repairs. We assume here that routes can be adequately repaired through the local route repair procedure described earlier. These *limited depth* queries produce much less control traffic and much lower delays compared with the initial full depth query.

We assume that the local processing time (e.g., table lookup) is negligible, compared with transmission delays.

speed, v, and is independently assigned an initial direction[13], which is uniformly distributed between 0 and 2π. When a node reaches the edge of the simulation region, it is reflected back into the coverage area.

Each simulation runs for duration of 125 seconds. No data is collected for the first 5 seconds of the simulation to avoid measurements during the transient period and to ensure that the initial intrazone route discovery process stabilizes.

In order to measure the delay resulting only from the ZRP overhead, the network load is assumed to be low. Route failures are detected and acted upon. The route queries are generated according to a Poisson arrival process, with the arrival intensity being a simulation parameter. The route queries represent both the initial query performed at the beginning of a session and subsequent queries due to reported route failures. Each route query is for a destination selected from a uniform random distribution of all other nodes in the network. Since the average time between a node's query for the *same* destination is longer than the expected interzone route lifetime, discovered interzone routes are effectively used only once and then discarded.

For the purposes of our simulation, we have made a number of simplifying assumptions regarding the behavior of the lower network layers and channel. This simplified model helps to improve understanding of our routing protocol behavior by providing our performance measures with some immunity from lower layer effects.

From the media access control (MAC) perspective, we assume that there is no channel contention. This assumption is necessary to separate the delays associated with a particular MAC scheme (e.g., collision avoidance algorithms) from the delays related to the routing protocol. These MAC independent results could be used as a benchmark for future analysis of the interaction between the Routing and the MAC layers

Our assumption of a collision-free media access protocol means that the average SIR of a received packet is limited by the ambient background noise and receiver noise. For fixed transmitter and noise powers, we assume that the BER is reasonably low within a distance, which we call d_{xmit}. Beyond $d_{xmit,}$ the BER increases rapidly. This behavior results from a rapid decrease in received power as the separation distance is increased. We approximate this rapid increase in BER by the following simplified path loss model:

$$PL(d) = \begin{cases} 0[dB] & \text{for} \quad d \le d_{xmit} \\ \infty[dB] & \text{for} \quad d > d_{xmit} \end{cases}$$

We interpret this behavior as follows: any packet can be received, error-free, within a radius of d_{xmit} from the transmitter, but is lost beyond d_{xmit}. Since packet delivery is guaranteed to any destination in range of the source, we are able to further reduce the complexity of our model by eliminating packet retransmission at the data link level.

Fixed Simulation Parameters

Parameter	Symbol	Value
network coverage area	A	600 m x 600 m
transmission radius	d_{xmit}	100 m
beacon period	T_{beacon}	0.2 sec
transmission rate	R_{xmit}	1.0 Mbps

[13] Direction is measured as an angle relative to the positive x-axis.

Variable Simulation Parameters

Parameter	Symbol	Values
routing zone radius	ρ	1-5
node speed	v	50,100,150 km/h
mean route query rate	R_{query}	0.25, 0.50, 1.00 query/sec/node

5.0 Performance Results

Results of our simulation are presented in the following figures. Figure 11 shows the dependence of intrazone control traffic on the routing zone radius, ρ, for various rates of network reconfiguration. All else being equal, the rate at which the networks reconfigure increases linearly with the speed of the nodes. For unbounded networks with a uniform distribution of nodes, we expect the increase in intrazone control traffic to be proportional to ρ^2. However, because our network is of finite size and the nodes are distributed randomly, we find that increase is actually somewhere between ρ and ρ^2. It should be noted that there is no intrazone control overhead for $\rho = 1$. All nodes within a routing zone of $\rho = 1$ are, by definition, neighbors. Consequently, the Neighbor Discovery Protocol provides all of the information needed to maintain connectivity within the routing zone.

The performance of the reactive portion of the ZRP is exhibited in Figure 12. As we increase the routing zone radius, we find that the rate of interzone control traffic decreases. This decrease can be attributed to three factors. First, as the size of the routing zone increases, more destinations can be found within a routing zone, requiring fewer IERP route requests. Second, as routing zones become larger, the redundant route query traffic is reduced through the increasingly directed propagation of queries to peripheral nodes. Lastly, the average number of peripheral nodes between a querying source and destination is inversely proportional to the routing zone radius. Thus, as the routing zone radius increases, the IERP accumulated routes are specified, on average, by fewer node IDs.

For $\rho = 1$, all nodes are peripheral nodes and bordercasting is equivalent to flooding. For $\rho > 1$, we observe a significant reduction in the interzone control traffic, indicating a potential benefit from a hybrid routing scheme compared to purely reactive routing.

The total control traffic (i.e., the sum of the control packets from the intrazone and interzone protocols), depicted in Figures 13, gives an indication of the performance of our hybrid routing scheme. For low route query rates, we find that relatively reactive routing (i.e., small routing zone radii) produces the least amount of control traffic. As the route query rate increases, the control over-

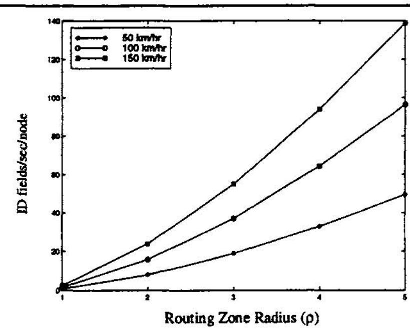

Figure 11: Intrazone Routing Control Traffic

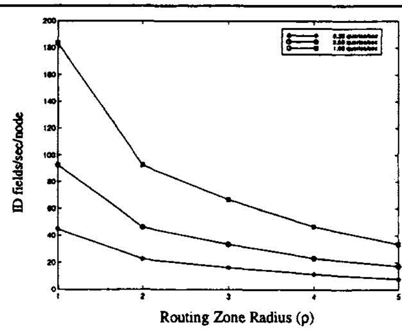

Figure 12: Interzone Routing Control Traffic

overhead can be minimized by increasing the routing zone radius. For the network configurations and operational conditions that we assume in our simulation, configuring the ZRP for a routing zone of $\rho = 2$ is shown to reduce the rate of ZRP control traffic by approximately 45% of the purely reactive schemes. For larger ρ, the overhead required to maintain larger routing zones outweighs the benefits gained from bordercasting

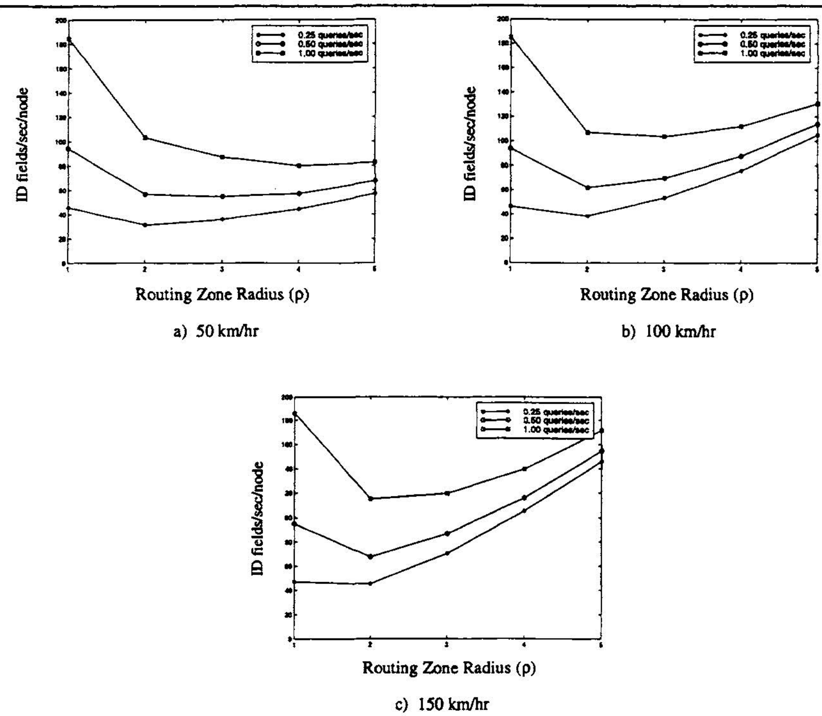

Figure 13: Total ZRP Control Traffic

Figures 14 show the performance of the ZRP as measured by the average route query response time. The delay characteristics appear to be heavily influenced by the behavior of the interzone route discovery protocol.

Under the conditions that the average amount of control traffic is small to moderate (small routing zones, small query rates and moderate nodal velocities), most of the instantaneous network load is due to a single route discovery. When the routing zones become relatively large and the network topography more volatile, the overall ZRP control traffic becomes large and begins to have a noticeable impact on the instantaneous network load. This behavior is exhibited at v=150[km/h] (Figure 14c). We note that for the load of 0.5 and of 1.0 [queries/second] (representing short route lifetimes), a minimum in average route query response time appears at $\rho = 4$. Although we've simulated the ZRP for a medium sized networks with routing zones of $\rho \leq 5$ hops, it is reasonable to assume that for wider networks and larger routing zones, minimum will also be present for relatively low velocities and low route query rates as well. Neglecting the effects of additional data traffic, we find that the ZRP can provide as much as a 50% reduction in the average route query response time compared with purely reactive routing. This improvement is somewhat smaller for networks with highly dynamic topologies, but even our most volatile networks exhibit significant improvements of 38% compared to purely reactive routing.

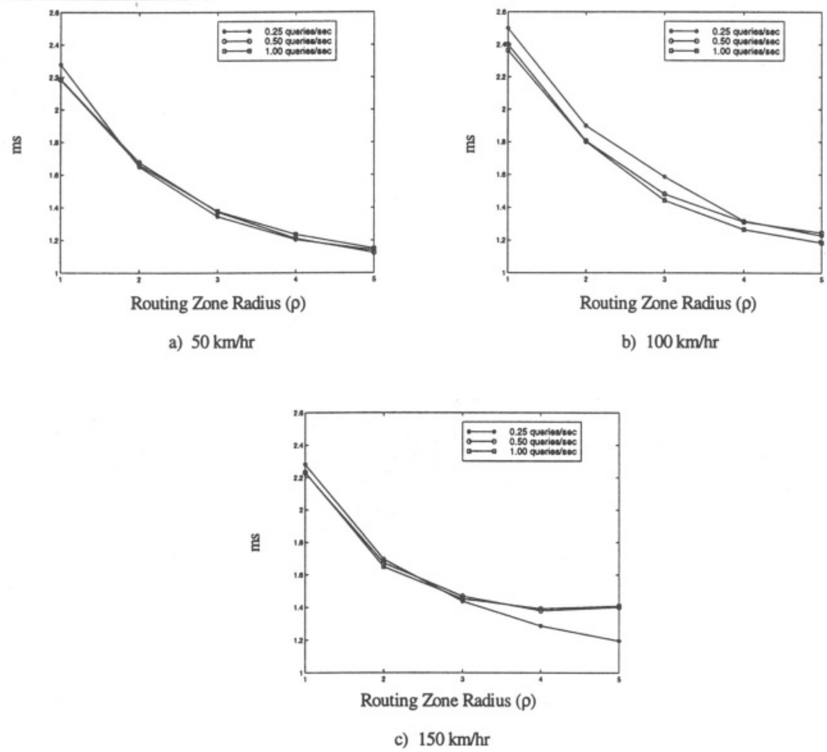

Figure 14: Average Route Query Response Time

6.0 Summary and Concluding Remarks

The Zone Routing Protocol (ZRP) provides a flexible solution to the challenge of discovering and maintaining routes in the Reconfigurable Wireless Network communication environment. The ZRP combines two radically different methods of routing into one protocol. Interzone route discovery is based on a reactive route request/route reply scheme. By contrast, intrazone routing uses a proactive protocol to maintain up-to-date routing information to all nodes within its routing zone.

The amount of intrazone control traffic required to maintain a routing zone increases with the size of the routing zone. However, through a mechanism, which we refer to as bordercasting, we are able to exploit the knowledge of the routing zone topography to significantly reduce the amount of interzone control traffic. For networks characterized by highly mobile nodes and very unstable routes, the hybrid proactive-reactive routing scheme $(\rho>1)$ produces less average total ZRP control traffic than purely reactive routing $(\rho=1)$. Purely reactive schemes appear to be more suitable for networks with greater route stability. Furthermore, for highly active networks (frequent route requests), more proactive networks produce less overhead (i.e., larger routing zones are preferred).

We note that for networks with low activity, the instantaneous network load is generally dominated by the control traffic from a single route discovery. Consequently, the ZRP exhibits minimum delay for relatively large routing zone radii (for the networks we simulated, $\rho \geq 4$), even for cases where relatively reactive routing minimizes the average ZRP control traffic. For highly volatile networks, the ZRP has been shown to provide 38% less delay than reactive routing. For slower, more stable networks, the optimal-delay ZRP configuration produced a nearly 50% reduction in delay compared to reactive routing. Based on the performance of the ZRP under heavy *control* traffic, we expect that additional data traffic will further reduce the optimal size of the routing zone.

7.0 References

[Bertsekas92] D. Bertsekas and R. Gallager, *Data Networks,* Second Edition, Prentice Hall, Inc., 1992.

[Cheng89] 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.

[Corson97] M.S. Corson and V. Park, "Temporally - Ordered Routing Algorithm (TORA) Version 1 Functional Specification," IETF MANET Internet Draft, Dec. 1997

[Ephremides87], A. Ephremides, J.E. Eieselthier, and D.J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proceedings of the IEEE,* vol.75, pp.56-73, January 1987.

[Garcia-Luna-Aceves93] J.J. Garcia-Luna-Aceves, "Loop-Free Routing Using Diffusing Computations," *IEEE/ACM Transactions on Networking,* vol.1, no.l, February 1993, pp.130-141.

[Gerla95] M. Gerla and J.T-C. Tsai, "Multicluster, Mobile, Multimedia Radio Network," *ACM/Baltzer Wireless Networks Journal,* vol.1, no.3, pp.255-265 (1995).

[Haas97] Z.J. Haas, "A Routing Protocol for the Reconfigurable Wireless Networks," *IEEE ICUPC'97,* San Diego, CA, October 12-16, 1997.

[Haas98] Z.J. Haas and S. Tabrizi, "On Some Challenges and Design Choices in Ad-Hoc Communications," submitted for publication.

[Haas98-2] Z.J. Haas and M.R. Pearlman, "Providing Ad-Hoc Connectivity with the Reconfigurable Wireless Networks," submitted for journal publication.

[Johnson96] D.B. Johnson and D.A. Maltz, "Dynamic Source Routing in Ad-Hoc Wireless Networking," in *Mobile Computing,* T. Imielinski and H. Korth, editors, Kluwer Academic Publishing, 1996.

[Lauer88] G. Lauer, "Address Servers in Hierarchical Networks," IEEE International Conference on Communications '88, Philadelphia, PA, 12-15 June 1988.

[Leiner87] B.M. Leiner, D.L. Nielson, and FA. Tobagi, "Issues in Packet Radio Network Design," *Proceedings of the IEEE,* vol.75, pp.6-20, January 1987.

[Moy97] J. Moy, "OSPF Version 2", *RFC 2178,* March 1997.

[Murthy] S. Murthy and J.J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks," *MONET,* vol.1, no.2, pp.183-197, October 1996.

[Murthy95] 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.

[Perkins94] 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, pp.234-244.

[Perkins97] C.E. Perkins "Ad Hoc On-Demand Distance Vector (AODV) Routing,", IETF MANET Internet Draft, Dec. 1997.

[Westcott84] J. Westcott and G. Lauer, "Hierarchical routing for large networks," *IEEE MILCOM'84,* Los Angeles, CA, October 21-24, 1984.