Networking on the Battlefield: Challenges in Highly Dynamic Multi-hop Wireless Networks

Ricardo Sánchez, Joseph Evans and Gary Minden

Information and Telecommunications Technology Center
Department of Electrical Engineering and Computer Science
The University of Kansas, Lawrence, Kansas 66044-7541
Email: {rsanchez, evans, gminden}@ittc.ukans.edu

(To Appear In Proceedings of IEEE MILCOM '99, Atlantic City, New Jersey, October 1999)

Keywords: Tactical Mobile Networking.

ABSTRACT

A crucial aspect of effective networking on the battlefield is choosing the correct networking architecture. Multi-hop wireless networks provide the best model for tactical networking because of their ability to self-organize and rapidly adapt to change. We focus on a multi-hop wireless network model that is *highly dynamic* and that consists of mobile base stations and mobile hosts. In this model, there are two key requirements for enabling an effective networking infrastructure for the battlefield: the support of highly mobile nodes and the scalability to a large number of nodes. In this paper, we present some of the system-level challenges encountered in highly dynamic multi-hop wireless networks. In particular, we address the topology model, the location model, and the routing model in light of the aforementioned challenges.

INTRODUCTION

Wireless networking technology will play a key role in future battlefield communications. The choice of the network architecture model strongly impacts the effectiveness of the tactical applications proposed for the mobile military networks of the future. Broadly speaking, there are two major models for wireless networking: single-hop and multi-hop. The single-hop model [1], based on the cellular network model, provides one-hop wireless connectivity between mobile hosts and static nodes known as base stations. This type of network relies on a fixed backbone infrastructure that interconnects all base stations by high-speed wired links. Supporting a fixed infrastructure greatly simplifies network management operations. However, future mobile military networks cannot rely on static infrastructures for three reasons. First, fixed nodes are more vulnerable to enemy attacks. Second, highly mobile military forces need networks that are equally mobile. Third, the need for military networks to continue "in operation" even when some nodes are destroyed and/or some links are jammed. On the other hand, the multi-hop model [2] requires neither fixed, wired infrastructure nor predetermined interconnectivity. Ad hoc networking [3] is the most popular type of multi-hop wireless network because of its simplicity: a homogeneous system of mobile hosts connected by wireless links. In ad hoc networks, all nodes are presumed to have similar capabilities with respect to energy supply, processing power, memory capacity, etc. As a result, each node must be prepared to provide general support services to the entire network (such as buffering messages and routing). Although the ad hoc model has been suggested for future mobile military networks [4], we believe it is more suitable for military networks where the cost of communication equipment is small compared to the value of the platform. For instance, the ad hoc model seems appropriate for a Navy network in which every node is a ship. In contrast, it seems inadequate for battalion forces in which some of the nodes are foot soldiers (think of the burden in putting a router on the back of each soldier!). We distinguish yet another type of multi-hop wireless network in which some nodes are more powerful than others, yet they are all free to move around arbitrarily. We call this network a *highly dynamic multi-hop wireless network* (HDNet). This network is characterized by two types of nodes: mobile base stations and mobile hosts. An HDNet example is shown in Figure 1.

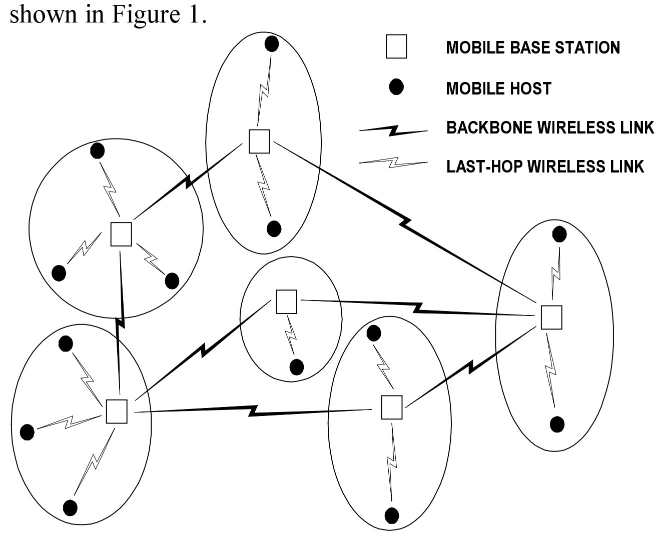

Figure 1- A Typical HDNet System

In HDNets, mobile base stations (or switches) have more resources than mobile hosts (or end-users) in terms of processing power, memory capacity, energy supply, etc. A mobile base station provides connectivity to mobile hosts in close proximity through last-hop bi-directional wireless links. In addition, neighboring mobile base stations interconnect with each other through backbone bi-directional wireless links to conform a mobile wireless backbone. This organization imposes a physical hierarchical structure on the network closely matching that of the military. Future mobile military networks are likely to follow a more heterogeneous trend in which the vast majority of the nodes are relatively simpler (like mobile hosts) while fewer nodes are more complex (like mobile base stations). Because of this, we believe that the HDNet approach offers a more practical and cost-effective solution to the battlefield communication problem than the aforementioned ad hoc model.

There are two important requirements to effectively enable networking on the battlefield:

- Support of highly mobile nodes (i.e., nodes that move so fast that the wireless links established from them to peers – end-user-to-switch and/or switch-to-switch – have extremely low duration thus giving the network little time to react to these rapid state changes).
- Ability to scale to a large number of nodes (i.e., hundreds to thousands of nodes) ranging from stationary to highly mobile nodes.

Dealing with highly mobile nodes implies coping with short-lived connections over highly variable wireless links that are easily prone to failure. This problem affects the design of a self-organizing architecture that faces fast and frequent changes not only at the backbone level but also at the end-user level. On the other hand, dealing with a large number of nodes implies the maintenance of a consistent global view of the network. But maintaining global consistency in a large network, while keeping track of highly mobile nodes, is not an easy task since the network has little time to react to the rapid and frequent state changes. Consequently, in a large, highly mobile HDNet system, the topology must be operated in a continuous reconfiguration mode to keep pace with the rapid and continuous state changes in the network.

Previous research on HDNet architectures has been reported in the context of the MMWN project [5] and the RDRN-I project [6]. Both of the projects have provided a system-level approach to self-configuring HDNet architectures in different ways. MMWN addresses the problem of quality of service provision in large HDNet system. On the other hand, RDRN-I focuses on techniques for rapid deployment of a medium-size wireless-ATM network. Our knowledge indicates, however, that no work addressing the design of a HDNet-like architecture to simultaneously support *scalability* and *highly mobile nodes* has ever been published.

In this paper, we focus on three challenging areas in the design of HDNets: topology management, location management, and routing management. We provide a system-level analysis of each area from the perspective of a large, highly mobile HDNet system. Table 1 shows some of the technical challenges in the listed areas. Subsequent sections discuss in more detail these three areas.

Challenge area	Technical Challenges
Topology management	- Providing self-configuring mechanisms for highly dynamic mobile wireless backbone control - Developing a network control organization that efficiently accommodates a large hierarchical network including highly mobile nodes
Location management	- Locating and tracking highly mobile nodes in a large network
Routing management	- Managing routes to/from/through highly mobile nodes in a large network

Table 1. Technical Challenges in Large, Highly Mobile HDNets

TOPOLOGY MANAGEMENT CHALLENGES

Topology management refers to the control mechanisms required to autonomously organize a variable number of nodes into a connected network. In a tactical environment, it is important that the network be *rapidly deployable* to ensure connectivity among nodes in the shortest amount of time. It is also important that the network be *rapidly reconfigurable* to provide timely reactions to changes in the topology caused by node destruction and/or jamming of links. In a large, highly mobile HDNet system the problem of topology management is aggravated by the fact that a highly mobile node can be intermittently connected to the network for a very short amount of time establishing extremely short-lived wireless links. Further, as the network is allowed to scale, the presence of highly mobile nodes is threatening to a network organization that strives to maintain the connectivity among the majority of its nodes. Moreover, since an HDNet system is inherently hierarchical (i.e., two-tier architecture), there is a need for some degree of coordination for topology management operations between the lower-tier (i.e., end-user level) and the higher-tier (i.e., backbone level). But considering the fact that the impact of high mobility at the lower-tier differs from that occurring at the higher-tier, coordination among tiers can be a complex task.

When all these issues are assimilated, one question arises: *What does it mean for a node to be highly mobile?* We believe that high mobility is a relative concept. For example, while one node may be highly mobile with respect to another node, it may appear quasi-static to a third node. We thus consider it necessary to introduce the notion of *relative mobility* to characterize the degree of mobility a node exhibits with respect to its peers[1] for a given period of time. This property states that, at a certain point in time, a node with high relative mobility is more prone to erratic (uncertain/unstable) behavior than a node with less relative mobility. The relative mobility of a node can be interpreted as a first approximation of future node behavior. A node computes its relative mobility by exchanging its mobility profile (i.e., position, velocity, direction, acceleration, etc.) with potential and current peer nodes[2]. But as a node moves, its relative mobility with peer nodes, which may also move changes. Therefore, the relative mobility of a node must be periodically re-evaluated to allow for adaptation to future states of the network. A network design that takes into account the relative mobility of its nodes is better positioned to yield higher efficiency (responsiveness) and stability (adaptability) to highly dynamic topological changes.

The relative mobility of a node can be used to characterize the *capabilities* of the node in question with respect to its peer nodes over the resulting peer links. For example, a highly mobile end-user may be limited to use datagram traffic, while a less mobile end-user can use either datagram or virtual circuit traffic. Similarly, a highly mobile switch may not offer critical backbone services (e.g., address database, intra-backbone traffic switching, etc.) and be limited to offer basic backbone access, with restricted datagram traffic, for mobile end-users.

At the system level, topology management involves network monitoring and control in a de-centralized fashion to decide the capabilities and interconnectivity of each node. But topology management in large dynamic networks is a burdensome task. Flat mechanisms are not scalable. If a flat scheme were used for topology representation, each node would have to maintain the entire topology of the network by including information for every link on the network and reachability information for every node in the network. Although feasible for small networks, this would create enormous overhead for larger networks. To address this problem, some form of hierarchical control is required. The goal of a hierarchical control strategy is to achieve topology aggregation by reducing nodal as well as link information so that the scaling in a larger network is more manageable. In a hierarchical network, the entire network is organized into groups of nodes known as *clusters* [7][8]. A hierarchy is formed by virtue of *nesting* so a cluster at a higher level is formed by cluster(s) at lower levels. In HDNets, the lowest cluster is represented as a single switch with zero or more associated end-users. Clusters act as information aggregator units by designating a node within each cluster to perform the function of a *clusterhead.* The clusterhead is not only a repository for the knowledge of the cluster but also a coordinator of the cluster operations. But in a highly dynamic environment, the use of a cluster-based strategy for topology management introduces new challenges.

One of the challenges in topology management for highly dynamic settings is the election of clusterheads. Since clusterheads can easily become a single point of failure within a cluster, the distributed election process should only consider those nodes with a higher degree of stability (i.e., low relative mobility). It is useless to elect a highly mobile node as a clusterhead because it can jeopardize the integrity of the cluster at any time. On the other hand, a node should continue to function as a clusterhead for as long as its degree of stability is maintained within a safe boundary. If the node running the clusterhead service is on the verge of experiencing some degree of unstability (e.g., sharp increase in relative mobility), then it should step down and make arrangements for *service migration* so a more stable node within the cluster can take over as the clusterhead. Note that this concept of service migration can be extended to any network services (e.g., address servers, routing caches, etc.) running on nodes that could possibly become hotspots in a highly dynamic environment.

Another topology management challenge is that of cluster creation and maintenance. Clusters are continuously rebuilt as switch nodes join or leave. But due to the high dynamics of the environment, spending too much effort in maintaining cluster consistency can be wasteful. Especially, if the integrity of the cluster is compromised by the presence of highly mobile nodes that join/leave the cluster very quickly. The topology should be constructed by placing more stable nodes at the top and less stable nodes at the bottom of the hierarchy. This scheme minimizes the impact of reconfiguration actions by localizing the effects to the lowest levels of the hierarchy and reducing the amount of update traffic required. Relative mobility can be used to rank nodes in the hierarchy as well as determine the depth (i.e., number of levels) needed to operate the hierarchy more efficiently. And since any node in the hierarchy may change its mobility profile at any time, the relative mobility always provides an indication of future node's stability, causing changes on the node's capabilities, and triggering reconfigurations in the hierarchy as needed. There is also the related issue of nodal degree, which affects intra-cluster and inter-cluster connectivity. Networks with a high degree of connectivity suffer from increased overhead and complexity. Therefore, relative mobility can also be used to control nodal degree such that more stable nodes have a richer connectivity than less stable nodes.

LOCATION MANAGEMENT CHALLENGES

Location management refers to the set of mechanisms required for tracking the location of mobile nodes within a

[1] Two peer nodes are neighboring nodes that communicate directly over a bi-directional point-to-point wireless link.

[2] In HDNet systems, we assume the existence of a separate broadcast-based channel for the exchange of control information. Therefore, transmissions from a given node reach only all neighboring nodes within line-of-sight.

network. This involves the creation of dynamic location directories that contain mappings between node static identities and dynamic addresses (i.e., location) in the network. When a node moves, its location has to be updated at the appropriate directory server. A source node that wants to communicate with a target node first needs to know the target's location. Locations are obtained by querying a location directory service. In a cluster-based system, location directories are normally provided by the clusterheads. Because of the hierarchical clustering model, location directories are configured as nested databases establishing parent-children relationships between nested clusters. Hierarchical addresses are then generated and maintained for all nodes in the network. Since end-users are associated to a switch (i.e., level-zero cluster), an address query always starts at the bottom of the hierarchy. If target is not there, then the query is propagated up the hierarchy to the next parent location directory that contains the target. When found, the query is propagated down the hierarchy until it finds the zero-level cluster, or switch, containing the target end-user. Once the hierarchical target location is known, a source node is able to deliver data by some means of hierarchical routing to the destination node.

In a large, highly mobile HDNet system the problem of location management revolves around providing a resilient multi-level directory structure that operates in a hierarchical network. Since the entire network (i.e., backbone and end-users) is mobile, changes on the backbone can cause not only disruption of location directory service but also location updates for, even static, end-users. One such scheme designed for tracking highly mobile end-users in an HDNet-like system has been reported in the literature [9]. However, we still note several challenges that need to be resolved.

One such issue is the design of a hierarchical addressing scheme with the ability of revealing the identity of highly mobile end-users. Because of global roaming, the dynamic binding between a user's address and its identity changes frequently. For example, highly mobile end-users tend to change their point-of-attachment much more quickly than more stable end-users. Therefore, it is important for an end-user to generate address updates in a controlled manner. This can be accomplished by triggering updates only when an end-user leaves a pre-defined cluster region. The region, which is known as the roaming cluster [9], is represented by the lowest-level cluster that contains a node for the purpose of triggering updates. For highly dynamic environments, however, roaming levels must be dynamically configured by each end-user. On the other hand, it is also important for a source node to know the relative stability of a target node upon querying its location. The reason for this is that the capability of the target provides hints as to what types of traffic can be used from a source. The relative mobility of a node can be used during the address update process to dynamically configure the most adequate roaming level for the node in question.

Another issue is the effect of highly mobile switches on their associated end-users. In such situations, the likelihood of hierarchical address changes at the end-user level is higher. As indicated earlier, highly mobile switches should be restricted to provide backbone access to neighboring end-users that cannot obtain backbone connectivity through more stable switches. Strategies for obtaining short-lived hierarchical addresses are thus needed for end-points affected by this scenario.

ROUTING MANAGEMENT CHALLENGES

Routing management refers to the set of mechanisms required to route a packet from a source to a destination mobile node through the mobile wireless backbone. In a hierarchical cluster-based system, routing can take place within the cluster itself (i.e., intra-cluster routing) or across different clusters (i.e., inter-cluster routing). The main advantage of this model is scalability by means of hierarchical routing [10]. In such a scheme, nodes can use compact (not full) route specifications to forward packets to distant regions, which incur in less traffic overhead. When a node decides to communicate with a target, it first obtains the target's hierarchical address from the location server. The hierarchical address, which may only provide aggregated information about the direction to the remote destination, is then used in establishing an intra/inter-cluster route. As the packet gets closer towards the destination, the route is refined with more detailed information until it finally reaches the target node.

In an HDNet system, a route includes a source node (i.e., end-user or switch), a destination node (i.e., end-user or switch) and possibly one or more intermediate nodes (i.e., switches). Movement of any of these nodes affects the validity of the route. We note several challenges in providing a resilient routing strategy for large, highly mobile HDNets.

One of the challenges is that of route discovery. As mentioned earlier, routes can be made inconsistent by movements of end-users and/or switches. Adapting multi-hop routes to high rates of mobility can be very difficult. A way to minimize the effects of highly mobility in this situation is by differentiating the role that highly mobile nodes should play during routing. For example, highly mobile end-users should be limited to use datagram service. Highly mobile switches, on the other hand, should be pushed towards the periphery (i.e., lowest-level of the hierarchy) of the mobile wireless backbone to provide entry access (i.e., inbound or outbound access) to the end-users offering only datagram services. We identify the need for three types of routing services in HDNets systems: virtual circuits, datagrams, and selective flooding. Virtual circuits can be established between nodes that are connected by paths that are relatively stable (i.e., do not include any highly mobile nodes). Datagram service can be offered for nodes in which at least one of the members of the route (i.e., source, intermediate, and destination) is highly mobile. But because of mobility, there may be consistency problems in routing datagrams across highly mobile outbound switches and/or to highly mobile end-user destinations. For this case, a selective-flood mechanism that includes the potential outbound clusters may help to alleviate this problem. Although routing techniques for choosing the most stable paths in multi-hop wireless networks has been explored [11], they are still not suitable for highly dynamic environments. Therefore, new techniques for scalable routing specifications that consider the future availability of paths, perhaps utilizing the concept of relative mobility, are needed.

In highly dynamic networks, it is futile to provide support for routing protocols that continuously evaluate routes within the network before they are actually requested (i.e., pro-active protocols). Pro-active techniques may cause tremendous traffic overheads on evaluating unnecessary routes. But routing procedures that work on an as-needed basis (i.e., reactive protocols) are more suitable to large, highly dynamic HDNets. One of the challenges in developing a reactive strategy is route maintenance. For example, virtual circuits should be able to adapt to the movements of any node in the routing path (i.e, switches and/or end-users) for as long as the path can guarantee the minimum level of stability required by the virtual circuit service. Repairing techniques that attempt to ensure such a guarantee may find useful the relative mobility information provided by the nodes around the affected area.

One final challenge is the provision of more direct routes for nodes within close physical proximity. Because of hierarchical routing, logical routes are setup to traverse paths within the hierarchy tree. Routes setup this way can be sometimes sub-optimal. Several mechanisms can be devised so those more direct routes take over the hierarchical routes. A feasible solution can be to allow associations of end-users with more than one switch at a time. Multiple associations to neighboring switches may provide not only faster entry points to the backbone but also more resilience for the routing problem over highly mobile switches.

CONCLUSIONS

In this paper, we have presented some of the system-level challenges encountered in highly dynamic multi-hop wireless networks (HDNet) that include mobile base stations and mobile hosts. Our analysis is presented from the perspective of an HDNet system that features highly mobile nodes and that scales to a large number of nodes. In particular, we have addressed the three following challenging areas: topology management, the location model, and the routing model.

We believe that the HDNet networking model offers a more practical and cost-effective solution to future mobile military networks than the classical ad hoc multi-hop wireless model. We hope that as research continues to explore options for highly dynamic multi-hop wireless architectures, including mobile base stations and mobile hosts, the boundaries that prohibit global freedom for wireless communications will continue to disappear. This simple fact will be key in enabling effective networking technology on the battlefield.

REFERENCES

[1] D. Raychaudhuri and N.D. Wilson, “ATM-based Transport Architecture for Multiservices Wireless Personal Communication Networks,” *in IEEE Journal of Selected Areas in Communications*, Vol. 12, No. 8, pp. 1401-1414, October 1994.

[2] G.S. Lauer, “Packet-Radio Routing,” in *Routing in Communications Networks* (ed. M. Streenstrup), Chapter 11, Prentice-Hall, 1995.

[3] M.S. Corson, S. Batsell, and J. Macker, “Architectural Considerations for Mobile Mesh Networking,” *in Proceedings of IEEE MILCOM ’96*, McLean, Virginia, October 1996.

[4] Z. Haas and S. Tabrizi, “On Some Challenges and Design Choices in Ad-Hoc Communications,” *in Proceedings of IEEE MILCOM ’98,* Bedford, Massachusetts, October 1998.

[5] R. Ramanathan and M. Steenstrup, “Hierarchically-Organized, Multihop Mobile Wireless Networks for Quality of Service Support,” in *ACM/Baltzer Mobile Networks Journal*, Vol. 3, No. 1, 1997.

[6] R. Sanchez, J. Evans, G. Minden, V. Frost, and K. S. Shanmugan, “RDRN: A Prototype for a Rapidly Deployable Radio Network,” in *ACM Mobile Computing and Communications Review*, Vol. 2, No. 2, 1998.

[7] C.V. Ramamoorthy, A. Bhide, and J. Srivastava, “Reliable Clustering Techniques for Large, Mobile Packet Radio Networks,” in *Proceedings of IEEE INFOCOM’87*, San Francisco, CA, March 1987.

[8] M. Gerla and J. T-C. Tsai, “Multicluster, Mobile, Multimedia Radio Networks,” in *ACM/Baltzer Wireless Networks Journal*, Vol. 1, No. 3, 1995.

[9] K. Kasera and R. Ramanathan, “A Location Management Protocol, for Hierarchically Organized Multihop Mobile Wireless Networks,” in *Proceedings of IEEE ICUPC ’97*, San Diego, CA, October 1997.

[10] J. Wescott and G. Lauer, “Hierarchical Routing for Very Large Networks,” in *Proceedings of IEEE MILCOM ’84*, Los Angeles, CA, October 1984.

[11] C-K Toh, “Associativity-based Routing for Ad-hoc Networks,” in *Wireless Personal Communications*, Vol. 4, No. 2, March 1997.