Mobility Management in Hierarchical Multi-hop Mobile Wireless Networks

Guangyu Pei and Mario Gerla
Computer Science Department
University of California, Los Angeles
Los Angeles, CA 90095

Abstract

In this paper, we consider the mobility management in large, hierarchically organized multihop wireless networks. The examples of such networks range from battlefield networks, emergency disaster relief and law enforcement etc. We present a novel network addressing architecture to accommodate mobility using a “Home Agent” concept akin to mobile IP. The performance of the mobility management scheme is investigated through simulations.

I. INTRODUCTION

In an area where there is no or little communication infrastructure, or accessing to the existing infrastructure is ineffective or impossible, wireless mobile users may still be able to maintain the communications through the ad hoc network. One of the important issues in providing ubiquitous untethered communications is mobility management. In general, the mobility management (also referred to as mobility tracking or location management) is the set of mechanisms providing a time-varying mapping between the identifier of a mobile node (i.e., what the node is named) and its address (i.e., where the node is located with respect to the network structure). If network address is “flat”, mobility management is simply handled by the routing algorithm. The routing table contains all destinations and thus can deliver a packet to any destination. If the address and the network are hierarchical (as may be required when size grows large), several challenges emerge. There exist different mobility management schemes for different hierarchical networks. In the sequel, we overview the mobility management schemes used in two popular hierarchical networks, the cellular PCS networks and Internet. Then, we introduce a novel mobility management scheme for the hierarchical ad hoc network and contrast it to the former schemes.

A. Mobility management in PCS

Current cellular systems (such as IS-41 [16, 7] and GSM [8, 7]) partition the service areas into location areas (LA). Mobility management is implemented through a two-level hierarchy of databases called home location register (HLR) and the visitor location register (VLR). HLR permanently registers each subscriber while VLR registers any mobile user that has moved into its network temporarily. Once an mobile user enters a new LA, it transmits a location update message to the new Mobile Switching Centers (MSCs). The MSC then sends a registration query to its associated VLR. VLR updates its record on the location of the mobile user. If the VLR serves both the new LA and the old LA, no further action is required. Otherwise, the new VLR sends a location registration message to HLR. The HLR updates the location entry corresponding to the mobile user so that the entry points to new serving VLR. The HLR sends back to the VLR a response which may contain relevant parts of the user’s service profile and a registration cancelation message to the old VLR.

B. Mobility management in the Internet

The mobility management protocol for the Internet, Mobile IP [3, 5], is being standardized by the Internet Engineering Task Force (IETF). In Mobile IP, each mobile host is allocated a permanent IP address on a home network for identification. When a mobile moves to new location outside of the home network, it obtains a temporary forwarding address called the care-of address (CoA) in the foreign network. One way of obtaining the CoA is through the foreign agent (FA) in the visiting network. The mobile host registers with the foreign agent and uses the IP address of FA as the CoA. Alternative way of obtaining the CoA is through an address discovery protocol such as the Dynamic Host Configuration Protocol (DHCP) [13]. The mobile host can obtain the CoA as a DHCP client and reports its newly allocated CoA to its home agent. In this case, the mobile host will effectively operate as its own FA. Each mobile host must have a home agent (HA) on its home network to keep track of the mobile host’s current CoA (called mobility binding). Each time the mobile host establishes a new CoA, it must register with HA so that the HA always knows the current binding of every mobile host it serves. Hosts which communicate with a mobile host are known as correspondent hosts. A correspondent sends packets to the permanent address of the mobile host. The HA of the mobile host encapsulates and tunnels the packet to the mobile host. The home agents and foreign agents cooperate to provide the illusion for correspondent host that the mobile host is still in the home network. Route optimization is being developed by IETF to enable direct communication between correspondent host and the mobile host by maintaining a binding cache in which the correspondent host caches the CoA of the mobile hosts.

C. Mobility management in ad hoc wireless networks

In a “flat” ad hoc network, the mobility management is not necessary since all of the nodes are visible to each other via routing protocols. In global, proactive type flat
routing algorithms (e.g., DSDV [12], WRP [9], etc) based on a “flat” network, the routing tables have entries to all destinations. For on demand routing algorithms (e.g., DSR [6], AODV [14], TORA [10], ABR [17] etc), the routes are discovered when there is a need. All the nodes are discoverable by using the query-reply approaches and therefore, no mobility management is necessary. However, as discussed in [4], a flat network organization is not scalable, and hence some form of hierarchical clustering should be employed to hide detail and reduce overhead. MMWN (short for Multimedia over Mobile Wireless Networks) [2, 15] is a prime example of hierarchically organized network.

In MMWN, the network contains two kinds of nodes, endpoints and switches. Only endpoints can be sources and destinations for user data traffic, and only switches can perform routing functions. To form the lowest level partitions in the hierarchy, endpoints choose the most convenient switches to which they will associate by checking radio link quality. Autonomously, they group themselves into cells around those switches (cluster heads). This procedure is called “cell formation”. The switches, in turn, organize themselves hierarchically into clusters, each of which functions as a multihop packet-radio network. First level cluster heads organize to form higher-level clusters, and so on. This procedure is called “hierarchical clustering”. (A switch is a 0th level cluster).

To support data transfer between mobile nodes, it is necessary to keep track of their hierarchical addresses. To this end both paging and query/response are used in conjunction with a location manager to track the changes of hierarchical addresses due to the mobility of nodes. Each cluster has its location manager which keeps track of nodes within the cluster and assists in locating nodes outside the cluster. Each node has a roaming level which is specified with respect to the clustering hierarchy and which implicitly defines a roaming cluster at the corresponding level. Paging is used to locate a mobile node within its current roaming cluster. When a node moves outside of its current roaming cluster, it sends a location update to the location manager. This update propagates to the highest level from which inter-cluster movement is visible. By combining these hierarchical topology management and location management functions, hierarchical routing is extended to the mobile environment.

In this scheme there are several features which are potentially complex to implement and hinder performance. First, the mobility management agent (the location manager) is closely tied with the network hierarchical topology. This feature makes the location updating and location finding quite complex. A location updating/finding has to travel along the hierarchical tree of the location managers. Furthermore, the changing in hierarchical cluster membership of an location manager will cause the reconstructing of the hierarchical location management tree and complex consistency management. Second, MMWN system is not based on approved standard protocols such as IP. It provides its own protocols for addressing and routing. An encapsulation interface with translation capabilities is needed to enable MMWN internetwork with IP-centric networks. Finally, the paging and query/response approach used to locate mobile nodes may lead to control message overhead and introduce latency.

In this paper, we focus on the mobility management for Hierarchical State Routing (HSR) protocol recently proposed in [11], where the ad hoc mobile wireless network is hierarchically organized into multilevel clustering structure. The key feature is the notion of logical subnets (e.g., brigade in the battle field, colleagues in the same organization, or a group of students in the same class) in order to handle the mobility. The network addressing scheme is described in section II. In section III, we explain the mobility management scheme in detail. The performance evaluation of our proposed scheme is given in section IV. Section V concludes the paper.

II. NETWORK ADDRESSING ARCHITECTURE IN HSR

The most basic aspect of a network is the addressing scheme by which the nodes of the network are used to communicate with each other. HSR maintains a hierarchical topology, where elected cluster heads at the lowest level become members of the next higher level. These new members in turn organize themselves in clusters, and so on. The goals of clustering are multifaceted: efficient utilization of radio channel resources; reduction of network-layer routing overhead (i.e., routing table storage, processing and transmission overhead); transmission power control and battery saving etc. Here, routing overhead is the main concern. In addition to multilevel clustering, HSR also provides multilevel logical partitioning. While clustering is based on geographical proximity, physical relationship between nodes, (hence, it will be referred to as physical clustering), logical partitioning is based on logical, functional affinity between nodes (e.g., employees of the same company, members of the same family, etc). Logical partitions play a key role in mobility management.

A fundamental assumption in “ad hoc routing” networks is that all nodes are “equal” and therefore any node can be used as an independent router. In the following discussion, we define a node in the ad hoc network to be an entity capable of being moving independently from the other nodes in the network.

A. Addressing for physical clustering

The physical clustering hierarchy used in our architecture is illustrated in Figure 1. Different clustering algorithms can be used for the dynamic creation of clusters and the election of cluster heads. At lowest level (level 1), we have 4 physical clusters. Level 2 and level 3 clusters are generated recursively, new cluster heads are elected at each level and become members of the higher level cluster. Generally, there are three kinds of nodes in a cluster, namely, cluster-head node (e.g., Node 1, 2, 3, and 4), gateway node (e.g., Node 6, 7, 8, and 11), and internal node (e.g., 5, 9, 10, and 12). The cluster-head node acts as a local coordinator of transmissions within the cluster.
Each node has a unique identifier NodeID. NodeIDs are the physical hardwired addresses (i.e., MAC addresses). The NodeIDs shown in Figure 1 are MAC addresses. In HSR, we define the HID (Hierarchical ID) of a node as the sequence of the MAC addresses of the nodes on the path from the top hierarchy to the node itself. For example, in Figure 1 the hierarchical address of node 12, HID(12), is \(<3.2.12>\). We use the dotted notation for the hierarchical address. In this example, node 3 is a member of the top hierarchical cluster (level 3). It is also the cluster head of C2-3. Node 2 is a member of C2-3 and is the cluster head of C1-2. Node 12 is a member of C1-2 and can be reached directly from node 2. The HID address of node 12 is the concatenation of MAC addresses of the nodes along the path to reach node 12 from the top hierarchy. In other words, node 3 is the top hierarchy representative of node 12, node 2 is the next level hierarchy representative of node 12.

The hierarchical address is sufficient to deliver a packet to its destination from anywhere in the network using HSR tables. Referring to Figure 1, consider for example the delivery of a packet from node 5 to node 10. Note that HID(5)=\(<1.1.5>\) and HID(10)=\(<3.3.10>\). The packet is forward upwards to the top hierarchy by node 5 (i.e., to node 1). Node 1 delivers the packet to node 3, which is the top hierarchy node for destination 10. Node 1 has a "virtual link", i.e. a tunnel, to node 3, namely, the path \((1,6,2,8,3)\). It thus delivers the packet to node 3 along this path. Finally, node 3 delivers the packet to node 10 along the downwards hierarchical path, which in this case reduces to a single hop. The advantage of this hierarchical address scheme is that each node can dynamically and locally update its own HID upon receiving the routing updates from the nodes higher up in the hierarchy. No central control is required.

HSR reduces the routing table overhead by the aforementioned physical clustering. Let us assume that the average number of nodes in a cluster (at any level) is \(N\), and the number of hierarchical levels is \(M\). Then, the total number of nodes is \(N^M\). A flat link state routing requires \(O(N^M)\) entries. The proposed hierarchical routing requires only \(O(N \times M)\) entries in the hierarchical map. This maximum occurs in the top hierarchy nodes which belong to \(M\) levels (i.e., clusters) simultaneously and thus must store \(N\) entries per cluster. Thus, routing table storage, processing and updating at each node is greatly reduced by introducing the hierarchical topology. The drawback of HSR with respect to flat link state routing is the need of continuously updating the cluster hierarchy and the hierarchical address as nodes move. In principle, a continuously changing hierarchical address makes it difficult to locate and keep track of nodes. Fortunately, logical partitioning comes to help, as discussed in the next section.

**B. Addressing for logical partitions**

In addition to MAC addresses, nodes are assigned logical addresses of the type \(<\text{subnet}, \text{host}>\). These addresses have format similar to IP, and can in fact be viewed as private IP addresses for the wireless network. Each IP subnet corresponds to a particular user group with common characteristics (e.g., tank battalion in the battlefield, search team in a search and rescue operation, professionals on the move belonging to the same company, students within the same class, etc). The notion of a subnet is important because each subnet is associated with a home agent, as explained later. Also, a different mobility pattern can be defined independently for each subnet. This allows us to independently define the mobility models for different formations (e.g., members of a police patrol). The transport layer delivers to the network a packet with the private IP address. The network must resolve the IP address into a hierarchical (physical) address which is based on MAC addresses.

A node does not know which cluster a particular destination belongs to, except for those destinations within the same lowest level cluster. The distributed location server assists in finding the destination. The approach is similar to mobile IP, except that here the home agent may also move. Note that the IP subnetwork is a "logical" subnetwork which spans several physical clusters. Moreover, the subnet address is totally distinct from the MAC address. Each logical IP subnetwork has at least one home agent (which is also a member of the subnet) to manage membership. For simplicity, we assume that all home agents advertise their HIDs to the top hierarchy. The home agent HIDs are appended to the top level routing tables. Optionally, the home agent HIDs can be propagated downwards to all nodes together with such routing tables. As the number of logical subnets (and therefore Home Agents) grows, it may be desirable to define hierarchy also in the logical subnets. Only the top hierarchy Agent HIDs are then broadcast to all nodes via routing tables.

The advantages of this logical partition is the separation of mobility management from physical hierarchy. As mentioned earlier, it is very complex and expensive to couple location manager with physical clustering as in MMWN. The tree of location managers has to change upon the variations of physical clustering. The hand-off and consistency management of location database are formidable in the dynamic changing environment. HSR solves these problems by introducing the
home agent for each IP subnet to manage the hierarchical address changes of its subnet members. Furthermore, logical IP subnets can take advantage of the “movement locality” since the nodes from same subnets tend to have common tasks and move as a group. The location of a home agent (physical hierarchical address) roughly represents the routing direction for the nodes in the subnet it represents.

C. Integration with Internet Routing

IP type logical partition provides convenient framework for interconnection between wireless ad hoc networks and wired Internet. Since the IP address is used to deliver a packet in HSR, it provides the illusion to the outside world that the ad hoc network is simply a concatenation of IP subnets. Local delivery within the ad hoc “subnets” is accomplished using the HSR protocol while the standard IP routing mechanisms decide which packets should enter and leave the ad hoc wireless subnets.

![Ad Hoc network](image)

Figure 2: Interconnection between ad hoc network and Internet

Figure 2 depicts how an ad hoc network can be connected to the Internet. Node G is a gateway between the ad hoc network and the Internet. Routing on G’s interface internal to the ad hoc network is accomplished using HSR, while its interface connected to the Internet is configured to use normal IP routing mechanisms. G is configured as the default gateway for all the IP subnets outside the ad hoc network. In order for all the nodes in the ad hoc network to reach the gateway, node G advertises its hierarchical address with top hierarchy and its hierarchical address is piggybacked with the HSR routing table. Node S can send the packet directly to node G using the hierarchical address. After node G receives the packet, it will then forward the packet to destination node D using the standard IP routing.

On the other hand, any external node which would like to communicate with a node in the ad hoc network, sends the packet using standard IP mechanism to node G. Node G then forwards the packet to the destination via the ad hoc interface using HSR.

III. MOBILITY MANAGEMENT IN HSR

The mobility management in HSR involves the maintenance of a dynamic mapping database for the hierarchical address of each IP subnet member at the home agents. In a large mobile network, the problem of locating users (their physical hierarchical addresses) by their names is not a trivial task. In a wired Internet, the DNS provides a mapping between symbolic names and network addresses. The network address is then processed by the routing tables and leads directly to the destination. In wired networks, Mobile IP was developed to handle the last hop indirection, from Home Agent to a mobile user. In ad hoc networks there is no fixed home agent. This requires two new functional components: HID registration and HID finding, which are introduced in the following sections.

A. HID registration

HID registration is the process of updating the HID address database of a logical IP subnet member at its home agent. There are two addresses associated with each mobile node, one is the logical, IP address used for identification for packet delivery by transport layers and the other is the HID used for routing in HSR. Through the HID registration process, the home agent maintains up-to-date mobility binding for each subnet member.

Each subnet member registers with its home agent by sending a REGISTRATION message to the home agent. The REGISTRATION message contains two fields: its own host ID and current HID. As mentioned in section B, each member of a logical subnet network knows the HID of its home agent (it is listed in the routing table), therefore it uses that HID to send the REGISTRATION message to the home agent directly. Upon receiving an REGISTRATION message, the home agent creates a mobility binding for the mobile node if there is none, or modifies the mobility binding accordingly. Registration is both periodic and event driven (e.g., whenever the member moves to a new cluster). At the home agent, the registered address is timed out and erased if not refreshed. Since in most applications, the members of the same subnet move as a group (e.g., tanks in a battalion), they tend to reside in neighboring clusters. Thus, registration overhead is modest.

Instead of sending the small REGISTRATION messages individually by each node, the cluster head can in fact periodically send to the home agent of a subnet a single aggregated REGISTRATION message, which includes the registration information for all mobile nodes from same subnet in the cluster. The aggregation of small individual REGISTRATION messages into larger blocks improves the MAC (medium access control) layer efficiency considering the large per block overhead in most wireless MAC layers.

B. HID finding

HID finding is the process for sender to learn the HID of the destination node it wants to communicate with. Each node maintains a HID mapping cache. When a source wants to send a packet to a destination of which it knows the IP address, it first checks its local HID cache. It will send the packet directly to the destination if its cache has the valid HID of the destination. Otherwise, it extracts the subnet address from the IP address. From its internal list (or from the top hierarchy) it obtains the hierarchical address of the corresponding home agent (recall that all home agents advertise their HIDs to the top level hierarchy). It then sends the packet to the home
agent using such hierarchical address. The home agent finds the registered address from the host ID (in the IP address) and delivers the packet to destination. The home agent then sends the registered address to the sender through HID REPLY message. Once the source and destination have learned each other hierarchical addresses, packets can be delivered directly without involving the home agent.

HSR utilizes the address mapping cache to avoid the need to perform the HID finding for every packet. Aggressive caching helps minimizing the cost incurred by routing the packets through the home agent first. We assume that nodes can operate in a promiscuous receive mode on their wireless network interface hardware, causing the interface passing all heard packets to the network layer without filtering based on the destination address. Although our mobility management does not require this feature to work properly, this facility is common in current LAN hardware for broadcast media including wireless. Overheard control packets do help the optimization of the mobility management. Whenever a node overhears the REGISTRATION and HID REPLY messages, it will update its cache entries accordingly. Furthermore, when home agent sends the HID REPLY message to the sender, it could return multiple most recently refreshed HIDs beside the HID for the destination the sender is interested in. It could include as many HIDs as the payload is allowed for a MAC frame. The penalty for caching is primarily due to the wasted effort in delivering a packet using an incorrect address. We adopt soft state approach to maintain the cache, i.e., each entry of the HID address (mobility binding) is purged if it is not refreshed.

IV. PERFORMANCE Evaluation

The multihop, mobile wireless network simulator was developed on a simulation platform built upon the language PARSEC [1]. The simulator is very detailed. It models all the control message exchanges at the MAC layer (e.g., IEEE 802.11) and the network layer (e.g., HSR Protocol control messages). This is critical in order to monitor and compare the traffic overhead (O/H) of the protocol. In our experiments, the network consists of 400 mobile hosts. Two-level physical clustering is used. The number of logical partitions (i.e., IP subnets) is 8. Exponentially correlated random group mobility [15] is used for each logical partition. Each group moves independently from the others. The simulation area is in a 2000x2000 meter square. A reflecting boundary is assumed. Radio transmission range is 120 meters. A free space propagation channel model is assumed. Data rate is 2Mb/s.

The first experiment (Figure 3) reports the registration packet overhead (O/H) versus node mobility. The registration O/H increases as node speed increases. The registration O/H is greatly reduced if the cluster heads send the aggregated registration message to the home agents. Furthermore, when a cluster head of a cluster is changed, all the nodes in the cluster will change their clustering membership. But the number of distinct clustering membership after the change will always be less than or equal to the number of nodes in the previous cluster.

Figure 4 shows the average number of data transmissions per data packet delivery, \( \eta \), as the function of traffic rate. The lower value of \( \eta \), the more efficient is the data packet delivered. Since the packet will be routed via the home agent if the HID of the destination is not known at the sender, aggressive caching will decrease the value of \( \eta \). Depending on (1) whether the node will operate in promiscuous mode and cache HID addresses from the overheard mobility management control packets (promiscuous caching); or (2) whether the home agent will send multiple addresses in the HID REPLY message to the sender, we investigate how the various combinations of these two situations impacts on \( \eta \) for different traffic rate. The traffic is UDP sessions between random node pairs. As shown in Figure 4, promiscuous caching greatly improves the performance. Multiple address replies are also important for the networks with high random traffic rate.

The performance of the cache as the function of the cache purge time is shown in Figure 5. The hit ratio decreases as the purge time increases (Figure 5(a)) which is expected. The penalty of a miss will waste the network resources in the attempt to deliver a packet to the wrong physical address. Note the
η in Figure 5(b) is calculated by dividing the number of data transmissions with the number of successfully received packets, therefore the low hit ratio of cache will increase η. On the other hand, although the cache hit ratio is high when the purge time period is short, the value of η will also increase since more and more packets have to be routed via the home agents.

![Graphs](image)

Figure 5: Cache performance

V. CONCLUDING REMARKS

In this paper we have presented the mobility management in a hierarchically organized ad hoc wireless network. This scheme is the extension of the Mobile IP concept to the multihop mobile networks. Logical partitions of the ad hoc network to IP type subnets using the home agent concept provide an efficient and convenient framework for mobility management and inter connection with Internet. We also presented simulation results that validate the approach and show the drastic improvement of the scheme using address caches.

As the wireless and embedded computing technologies continue to advance, and increasing number of small size and high performance wireless devices will be used. This will create a future of large scale, dynamic and heterogeneous wireless networks. Our architecture and protocols are well-suited to provide efficient and robust routing and mobility management for applications in this context.

VI. REFERENCES


