

NSF CAREER Proposal: Project Summary

Adaptive Architectures for Multicast Service Support in Large-Scale Mobile Ad Hoc Networks: Design and Evaluation Framework

Ahmed Helmy, Electrical Engineering Department, University of Southern California

This proposal outlines a comprehensive career plan for advancement of research and education in the field of wireless mobile ad hoc networks. The research plan consists of two complementary parts. The first is to design and develop a novel architecture for scalable multicast routing in wide-area ad hoc networks, while the second part proposes new methodologies for evaluation of wireless ad hoc routing protocols. The educational plan includes developing wireless networking courses coupled with a laboratory, in addition to proposing novel methods to improve student participation in the learning process.

Ad hoc networks are evolving as very interesting architectures to support pervasive mobile wireless devices. At the same time, *group communication* represents a very important class of applications in future networks. Multicast is the enabling technology for efficient group communication. Providing a scalable architecture for multicast services in large-scale networks has proven to be a challenge for years in the networking community. The research challenges are even greater for ad hoc networks, mainly due to the *highly dynamic* nature of wireless nodes and their unexpected *mobility*. The developed multicast service should allow group participants to join and leave at will, and should impose no restrictions on node mobility. It should also provide automatic multicast address allocation and session advertisement. Most existing work on ad hoc networks ignores such requirement. We propose to design a new scalable architecture for multicast services support in large-scale ad hoc networks, the *first* of its kind.

Since ad hoc networks are *infrastructure-less*, the developed protocols must be *self-configuring*. Scalability and robustness should be addressed carefully with growth in the size of the network and the number of group participants. Existing approaches for multicast, usually employ flooding or cores for discovery of group participants. These approaches, however, only apply to small networks due to cost of flooding and maintenance of the core tree. Questions are often asked about which routing mechanisms to use, pro-active or re-active routing? Hierarchical or flat? It is usually the case, however, that each of these mechanisms has its strengths and weaknesses depending on network conditions. We introduce an approach to use *hybrid* schemes that *adapt* to network dynamics. But how to adapt in a meaningful way? Such adaptation should be a function of *mobility* and *power*.

The proposed research program intends to fully utilize *highly adaptive* mechanisms in the various components of the architecture. Specifically, we design an adaptive zone-based hierarchy augmented by the notion of *contact* nodes to increase network coverage. In addition, we provide novel schemes for multicast service support, based on *adaptive anycast* for resource discovery. We also introduce geographic multicast address allocation to map groups into *rendezvous regions*. All the proposed mechanisms are self-configurable, and promise desirable scaling and robustness properties.

Once the mechanisms are developed, they shall be thoroughly evaluated using a rich set of test suites. We intend to use systematic formal methods, network simulation and laboratory experiments to test, analyze and evaluate the developed mechanisms. In addition, we plan to develop rich scenario-based mobility models to deeply understand the effect of mobility on ad hoc routing performance.

The above research plan is complemented by a detailed *educational plan*, potentially leading to significant *improvements in learning*. The plan includes development of a series of courses, by the PI, on protocol design and ad hoc networking and the establishment of a *wireless networking laboratory*, to improve theoretical and practical aspects of the networking curriculum at USC. In addition, the PI introduces a *novel and unique* approach to increase student involvement in classroom activities using what we call *student-assisted* grading and course development. These methods will help create a *tighter feedback loop* between instructors and students. We also plan to incorporate architectures and results of our research into the curriculum for the networking program at USC, at the undergraduate and graduate levels.

1. Introduction

Emerging ad hoc networks are expected to have a significant impact on service paradigms in various vital fields; such as military, disaster relief, and bio-sensing. These infrastructure-less networks consist mostly of heterogeneous wireless mobile devices with various *power* constraints, capabilities and *mobility* characteristics. An essential capability in future ad hoc networks is the ability to provide scalable *group services* and communication. This proposal presents a *novel adaptive architecture* to support multicast services in large-scale ad hoc networks. Unlike existing work on multicast in ad hoc networks that mainly addresses local areas with tens to hundreds of nodes, our work targets *large-scale wide-area* ad hoc networks with, perhaps, tens of thousands of nodes. Issues of *scalability*, *robustness* and *efficiency* in such highly dynamic environment affect the essence of our architectural design and guide our choice of multicast model. In order to address these issues, we divide our research proposal into two main areas. The first is the architectural design and the second is the evaluation methodology. We introduce our architectural design by discussing the design requirements and presenting overview of the architectural components. Evaluation methodology is discussed later in the proposal.

1.1 Design Requirements

The main factors driving our design are scalability, the multicast service support, and robustness.

- (a) **Scalability:** Unlike most related work that considers tens to hundreds of mobile nodes, our architecture should be able to support two to three orders of magnitude more nodes. We believe that mobile nodes will be pervasive, replacing PCs and cellular phones, with tens of new classes of application supported by mobile wireless devices (e.g., navigation, location-based services). Flat architectures are known not to scale well, mainly due to the far-reaching effects of network dynamics; mobility, failures and topological changes. Such effects consume network resources (i.e., bandwidth and power), and lead to recovery delays and increased route oscillations. Hierarchical architectures, on the other hand, alleviate the above problems, as they tend to localize and dampen network dynamics, and scale routing tables using aggregation. Many existing hierarchical architectures are based on *clustering* mechanisms, in which a single node per cluster (called master, cluster-head or parent) is chosen to manage or organize the cluster. Such architectures suffer from single point of failures, in which the failure (or movement) of the master may have severe negative effects on the hierarchy. Furthermore, establishment and maintenance of clusters entails electing the master and mechanisms for joining/leaving clusters, which usually incur a lot of overhead and complexity. We design an architecture that leverages hierarchical advantages while alleviating effects of master and hierarchy maintenance. We adopt a two-level distributed hierarchical architecture. For the first level of the hierarchy we adopt a *zone-based* approach (a variant of the zone routing protocol ZRP[28]), in which each node has its own view of a zone. For the second level, we introduce a novel concept of *contacts* (based on the concept of small-world[29][55]) to enhance a node's view, and aid in route and resource discovery. Also affecting scalability, is the choice of routing protocol. In general, ad hoc routing protocols are either pro-active (i.e., table driven) or re-active (on-demand). Pro-active protocols exchange periodic messages to keep routes up-to-date. Pro-active route discovery has low delay, with significant overhead of periodic route exchange (many of which may become invalid due to mobility). Re-active protocols, by contrast, maintain routes on-demand. Re-active route discovery does not incur periodic overhead, but incurs more route discovery delay, which usually involves request broadcasts throughout the network. We believe that neither protocol perform well in all network conditions. We attempt to combine the strengths of both protocols using a hybrid approach.
- (b) **Multicast Service Support:** Multicast participants should be able to join, leave or send packets to the group at will. Participants are not known a priori and are allowed to move freely during a multicast session. We design a novel adaptive anycast architecture for resource discovery to facilitate the *rendezvous* of group participants. We also introduce a new multicast paradigm instead of the traditional approaches of broadcast-and-prune or rendezvous cores. Our paradigm is based on *sender push*, *server cache*, *receiver pull*, approach, that better fits large-scale ad hoc networks. Also, multi-

sender groups should be supported. Global multicast groups as well as locally-scoped groups should be supported. Mechanisms providing participants with active session information should be provided (This is part of our *bootstrap* architecture) and a multicast address allocation scheme should be introduced (We provide a new approach for geographic multicast address allocation). No unicast routing is required for multicast support. Also, minimum configuration of nodes is desirable (In our scheme, nodes only need to know a *well-known* session announcement group address and an algorithmic mapping function to map groups into regions).

- (c) **Robustness:** In a highly dynamic environment, such as ad hoc networks, where mobility and crashes are likely to occur, robustness is of prime concern. Being able to adapt to network dynamics to achieve correct behavior and reasonable performance plays a major role in our design. We incorporate mobility and stability models into our hierarchy formation to achieve adaptivity. In addition, our distributed adaptive resource discovery architecture avoids single point of failure scenarios and promises continued operation and graceful recovery during network partitions. We also incorporate path redundancy mechanisms in our multicast routing protocol. Multicast *trees* do not provide sufficient robustness against mobility and failures. We use *mesh* structures for robust delivery. Unlike existing proposals for mesh construction, however, our mechanisms will be designed to build meshes in anticipation of movement. Not only does that achieve better performance, but also provides path redundancy that may be used in case of failures. In addition, mesh branches are activated on-demand to reduce overhead without affecting robustness.

In addition, our mechanisms should be *energy-efficient* and *self-configuring*. Energy-efficiency is one side-effect of scalability. In addition, we attempt to limit communication (one main source of energy consumption) by using localized mechanisms for advertisement and query. In addition, in resource-discovery and anycast (where only a single resource is sought), localized broadcast techniques are adopted to reduce overhead. Self-configurability is an inherent feature of our hierarchy formation and resource discovery mechanisms.

1.2. Brief Architectural Overview

In order to address the above challenging requirements and support the proposed mechanisms, we provide an architectural framework based on the following components:

- i. Self-configuring hierarchy formation and adaptation
- ii. The multicast service architecture consisting of: (a) the multicast model, (b) multicast routing, and (c) adaptive resource discovery and multicast address allocation.

The rest of this proposal is outlined as follows. Section 2 introduces the hierarchy formation and adaptation mechanisms. Section 3 proposes a new multicast services architecture including a new multicast paradigm, multicast routing protocols, and novel architectures for adaptive anycast and geographical multicast address allocation. Section 4 presents the evaluation and analysis methodology based on a new scenario-based mobility model. Section 5 discusses related work. Section 6 provides the proposed research, and Section 7 provides the educational plan.

2. Self-Configuring Hierarchy Formation and Adaptation

We first provide the mechanisms for the self-configuring hierarchy formation, stemming from the concept of zone-based routing, augmented by our novel idea of *contacts*, along with description of contact selection and update mechanisms. Then we propose to integrate mobility adaptation mechanisms into our hierarchical architecture based on link availability and mobility estimation models.

2.1 Hierarchy Formation

As was discussed earlier, a flat routing architecture does not scale well for wide-area networks, and more so for ad hoc networks. Also, a hierarchical approach based on concept of *master* (or *cluster-head*), through which traffic from the cluster funnels, is undesirable.

One possible approach to consider is to use the master for infrequent coordination but not for forwarding packets. One such approach is the Landmark hierarchy (LMH)[32]. Although designed mainly for wired networks, LMH has the ability to self-configure dynamically, without relying on administrative domains and exhibits path lengths and routing table sizes comparable to conventional cluster-based hierarchies (e.g., Internet). The traffic from/to a cluster need not go through the landmark, which adds robustness. LMH, as presented in[32], however, was designed mainly for wired networks, without accounting for mobility dynamics or power constraints. After examination we identified several drawbacks of LMH. LMH employs complex promotion, demotion, and adoption operations for hierarchy maintenance. Furthermore, effects of mobility on the hierarchy were found to be drastic, sometimes leading to total re-configuration of the hierarchy. For example, movement of a high level landmark triggers re-election in its old region, then demotion and adoption for the moving node are triggered (potentially for as many times as the levels of the hierarchy). This may trigger a chain reaction that consumes many resources unnecessarily. In addition, sub-optimal paths may be common due to hierarchical routing. A variant of LMH (called LANMAR) was presented in[38]. LANMAR, however, was based on the premise that group mobility will be dominant in ad hoc networks. Hence, it does not provide mechanisms to solve LMH drawbacks¹.

Another approach that avoids complex coordination for architectural setup is the zone routing protocol (ZRP)[28] which defines a zone for every node as the number of nodes reachable within a radius of R hops away, shown in Figure 1 (a). Inside the zone proactive (intra-zone) routing is used, so nodes obtain routes to all nodes within their zone. To discover nodes outside of the zone reactive (inter-zone) routing is performed by flooding through periphery or *border* nodes of each zone (known as *bordercasting*). ZRP routing overhead depends heavily on the choice of one parameter; the *zone radius*. If the radius is too small the routing overhead is dominated by reactive overhead, and vice versa. Optimizing such a parameter is nontrivial. A hybrid min-search and traffic adaptive approach is used in[52] but requires relative stability of the network to approach optimality.

ZRP seems appealing, but experiences excessive delays and overheads in large-scale networks, where much of the traffic maybe destined out-of-zone. Therefore, we develop a novel approach that goes beyond the *zone* while maintaining similar simplicity and stability. Our approach is based on a concept we call *contacts*. Contacts, of a certain node x , are nodes that previously existed in x 's zone but are drifting out of zone, and hence have a network view beyond that of x or any of its border nodes². While drifting away gradually, x may maintain route to (some of) these drifting nodes using low overhead, by virtue of having been in x 's zone and are close to its border nodes. Figure 1 shows a simple illustrative example of zoning, contacts and effects of mobility.

The concept of contacts relates to the concept of *small world* graphs[29]. Watts observes that small world graphs can have low average path length with high clustering. In other words, introducing a small number of far away links (i.e., short cuts) leads to significant reduction in path length. He also found that picking random far away nodes may lead to small world graphs, depending on the probability of choosing the far link. We believe, however, that picking contacts at random is undesirable, since this may lead to unpredictable overhead for contact route discovery and maintenance. In addition, knowing mobility characteristics and stability of a node helps identify better (more useful) contacts. We *take advantage of node mobility* and pick contacts from those nodes drifting away from the zone. The eventual characteristics of the formed graphs, however, still need to be studied. The resulting graph is a function of time and depends on the initial choice of contacts and the node mobility. We plan to study these characteristics extensively in our research plan.

¹ More discussion of hierarchical architectures is given in the related work section.

² It seems this is one of the few concepts that actually takes advantage of mobility. Mobility is often viewed as a disadvantage, and for good reasons, but we think it may also be taken advantage of, when possible.

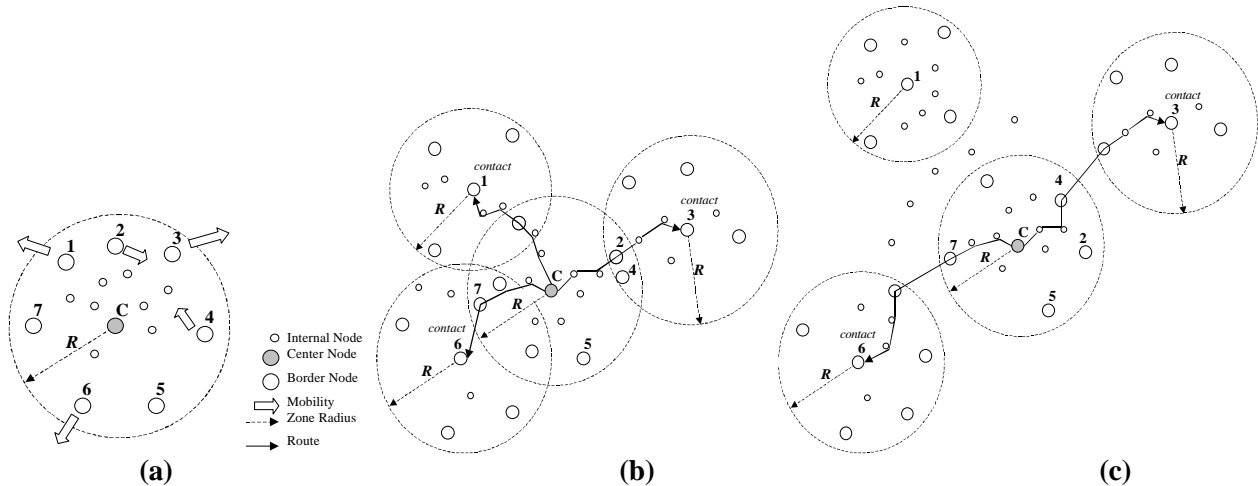


Figure 1. Example of zoning, contacts and effect of mobility: (a) Zone for center node C is shown (with radius R). Border nodes are numbered (1-7). Nodes 1,3 and 6 are moving/drifting out of zone. (b) Radii for the drifting nodes are shown. C stays in *contact* with the drifting nodes, which enables it to obtain better network coverage with low overhead. (c) After moving away, contact nodes drift up to a point where their zones no longer intersect with C 's zone. In this example, C maintains contact with those nodes not more than $(2R+1)$ hops away, i.e. nodes 3 and 6, and loses contact with node 1 as it drifts farther than the contact zone.

Contacts play an important role in route and resource discovery as will be explained later. A node should choose its contacts carefully to attempt to maintain a relatively long and useful contact as long as the contact route is kept. The *contact list*, maintained by a node, changes *adaptively* as the network conditions change. We further discuss how is the contact list chosen and changed in the next section.

2.2 Hierarchy Adaptation

The architecture presented thus far provides a framework for hierarchy formation. Such hierarchy needs to be highly adaptive to network dynamics and conditions. Unlike traditional protocols, our protocols must integrate concepts of *mobility* and *energy*. To achieve this, mechanisms that integrate measures of stability and power consumption must be introduced. We use such mechanisms for hierarchy formation and route selection. In hierarchy formation, the zone radius and contact selection may be adapted dynamically to establish certain stability measures for zones and contacts. Furthermore, each node should adapt its behavior to achieve desirable collective performance. For example, number of contacts chosen by a node should be a function of the contacts already established by other nodes in its zone. Increase in number of contacts increases bandwidth and must be done only as necessary.

In[13] a scheme was proposed to provide mobility-based adaptive behavior, where a model is used to measure link and path *stability*. The model used is (a,t) , where a is the probability that the link will be available for time t , and was utilized to build adaptive clustering schemes for ad hoc networks. The basic idea is to build stable clusters in which internal route availability is probabilistically bounded. The proposed model is a random walk-based model that determines the conditional probability that the nodes will be within range of each other at time t_0+t given that they are located within range at time t_0 . Although quite useful for analytical reasoning about mobility, the random walk mobility assumption may not hold for correlated mobility. Potential benefits of this approach warrant further study and investigation for different mobility models. Another work[57] defines a mobility metric derived from signal strengths, using a free space propagation model; $\frac{RxPwr}{TxPwr} a \frac{1}{d^2}$ where $RxPwr$ ($TxPwr$) is received (transmitted) signal

power, and d is the distance between transmitter and receiver. By measuring the signal strength between two consecutive packets received from the same transmitter a *relative mobility* metric is defined as

$10 \log \frac{RxPwr_{new}}{RxPwr_{old}}$, if negative high value then nodes are moving away quickly, and vice versa. This metric

was presented to aid in choosing cluster-heads. We plan to study such concepts, analyze their limitations, suggest improvements and extend them to use with other criteria such as power and QoS measures[13], and integrate it with our hierarchy formation architecture. Such integration helps in contact selection, where a node chooses its contacts with probability p , as a function of the border's mobility, the number of zone contacts, energy and activity. One simple model to be initially used is to have p proportional to the energy estimates of both nodes E_{est} , their relative stability S_{est} , the activity level of the node A_{est} measured as rate of discovery requests. Also, p is inversely proportional to the number of zone contacts Z_{est} . Hence, $p \propto \frac{E_{est} S_{est} A_{est}}{Z_{est}}$. These measures are based on either local or zone-scoped (e.g., piggybacked on intra-zone

pro-active messages). E_{est} includes energy estimates at the node choosing the contact, and the contact node drifting out of zone. To accommodate heterogeneous nodes, the estimates should include energy left E_{left} , and the drainage DE . Thus, a simple equation for the energy estimate is: $E_{est} = \left(\frac{E_{left}}{\Delta E} \right)_{node} \left(\frac{E_{left}}{\Delta E} \right)_{contact}$.

Stability estimates will be derived from modified models based on the (α, t) model and the signal power mobility metric described above, subject to further research. Furthermore, as was pointed earlier, the choice of contacts (along with network dynamics) affects the resulting connectivity graph. Thus, we consider such choice to be non-trivial, especially if we want to attain small world characteristics. We plan to pursue this issue further in our research. These adaptive mechanisms are also used in the multicast service model and resource discovery in section 3. In such case, contact's choice should also be affected by the capabilities of the contact node (e.g., GPS capability, or being a session or sender discovery server).

3. Multicast Service Architecture

Providing a scalable multicast service architecture is the focal point of our proposal. The hierarchical architecture proposed thus far provides the basis for efficient support of our multicast service model, presented in this section. As the mechanisms for our multicast model unfold, the essential role played by the adaptive hierarchical architecture will become very clear in support of multicast routing and resource discovery. In general, existing work on ad hoc multicast concentrates on multicast routing in small to medium size networks. Particularly focusing on mechanisms to establish distribution trees (or meshes) between senders and receivers, mainly using periodic broadcasts (either from the sender[36] or receiver[39] side) or relying on the existence of *cores*[26]. In this proposal, however, we address multicast in large-scale ad hoc networks, and propose new schemes for *bootstrapping* multicast services, by providing efficient *resource discovery* using *popularity-based adaptive anycast* and *geographic multicast address allocation*. We are not aware of existing or on-going work on multicast for wide-area ad hoc networks. Furthermore, we do not know of other work for *bootstrapping* such service, resource discovery or multicast address allocation in ad hoc networks. Our work does not assume any underlying unicast routing protocol.

Our multicast service architecture consists of three main components; (a) the multicast model (i.e., how senders and receivers meet), (b) multicast routing (i.e., establishing multicast distribution paths), and (c) *adaptive resource discovery* architecture using anycast and geographic multicast address allocation.

3.1. The Multicast Model

The basic premise for scalable multicast is that sources do not know who/where receivers are a priori. This model enables any node to join or leave a multicast group at any point in time. Hence, one of the main components of multicast is a mechanism for group participants to *meet* or *rendezvous*. Traditionally, this problem has been addressed, in wired and ad hoc networks, using two different approaches; *broadcast-and-prune*, and *rendezvous cores*. In the former[15][42][36], a participant (usually the sender) announces its presence by broadcasting data packets (or control messages) throughout the network. Network nodes not interested in the group send prune messages to stop the flow of packets (or

simply do not respond in case of control messages broadcast). These broadcasts are periodic to capture network and membership dynamics. It has been shown that such model is best suited for small to medium size networks with densely populated groups, but does not scale for wide-area networks[60]. The rendezvous cores approach, by contrast, avoids periodic broadcasts by providing *explicit join* mechanisms. Participants join (or send packets) towards a common core, which relays the packets from the senders to the receivers using a shared tree (or mesh)[2][26]. Problems of traffic concentration or single point of failure scenarios for core-based approaches may be alleviated by using multiple cores or dynamic core election mechanisms. We believe, however, that the major research problem associated with the core approach is the core bootstrap and consistency problem. How do participants know the core's address/location? Senders and receivers need to maintain a consistent view of the cores in order to *meet*. This problem was addressed for wired networks[14] within a single domain, and uses a flooding to disseminate core-to-group mapping. This scheme does not scale well for wide-area networks and its convergence performance degrades with the size of the network. The problems of rendezvous cores are exacerbated by ad hoc network dynamics and mobility.

Unlike the previous approaches for multicast, we instead propose a new multicast model. We refer to our model simply as *sender push, server cache, receiver pull* model. Unlike previous work on ad hoc multicast that requires periodic broadcasts throughout the entire network, our scheme incurs less overhead, and only when necessary as necessary and as localized as possible. We introduce the notion of *sender discovery servers (SDS)* to aid in sender location and information dissemination. As shown in Figure 2, a sender sends an Advertisement (*Adv*) using *localized broadcast*. *SDS*s receiving the *Adv* store this information. Receivers send joins toward the sender based on *backward learning*; every node forwarding the *Adv* adds its address to the message to construct a path back to the source. Other, farther, receivers not receiving the *Adv* message attempt to find a nearby *SDS*, first by checking in their own zone (*SDS*s advertise in their zone), then by checking with their contact list. If *SDS* is found it is queried for group information and responds with a join reply, with approximate source location or possible routes (if available at the *SDS*). Depending on the quality of the provided routes (if any), the querying receiver(s) may opt to use these routes or use zone/contact search for other routes (eventually, geographic routing may also be used for route discovery as in LAR[35], as described later). If *SDS* is not found, a receiver may send a localized broadcast to discover other nearby receivers of the group. In case this process fails to get information about the group then a fallback mechanism, described later in this section, is used. Once the information about the group/senders is available the route discovery/construction is initiated. An illustrative example is shown in Figure 2.

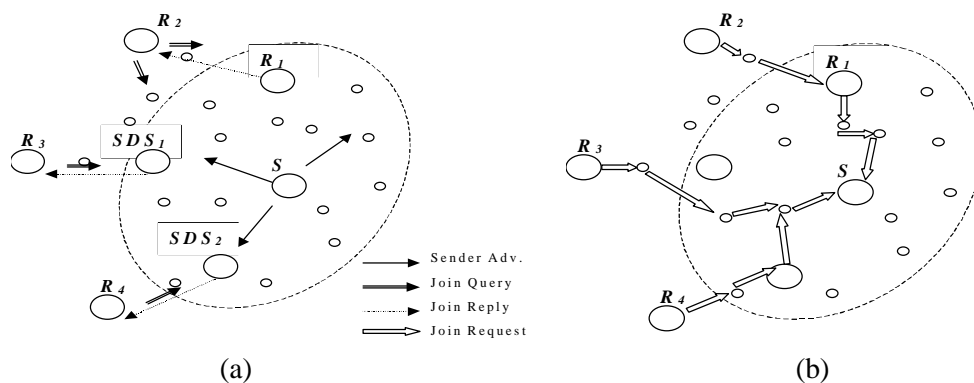


Figure 2. Multicast service scenario. (a) Sender S becomes active and broadcasts an advertisement (*Adv.*) locally (covered region is shaded). Sender discovery servers SDS_1 , SDS_2 , and joined receiver R_1 receive the *Adv*. When new receivers join the group, they try to find: (i) a sender discovery server, (ii) nearby members, and/or (iii) active senders for the group. Receiver R_2 finds R_1 using local broadcast, while R_3 and R_4 find SDS_1 and SDS_2 . (b) Once sender information is obtained, receivers send Join Requests to establish multicast distribution paths.

3.2 Multicast Routing

Establishing multicast distribution paths for ad hoc networks has been shown to be more robust using *mesh* structures, as opposed to conventional *tree* structures[27]. For single-source groups or when sources are sparsely distributed, however, even a mesh does not provide the desired path redundancy. Local recovery mechanisms[56] may be used to alleviate such problem. Conventional receiver-initiated multicast schemes setup reverse path forwarding (RPF) trees[2] due to dependence on unicast routing. For ad hoc networks, however, using RPF paths is not desirable due to possible path asymmetry due to wireless channels. Furthermore, we do not use RPF nor rely on existence of a unicast routing protocol.

We propose to use mesh structures with local recovery mechanisms. In addition, multiple paths may be selected by the receiver (during the discovery process) to increase robustness of the multicast distribution mesh. Depending on the number of senders in the group and their location (if available), in addition to the mobility of the receiver, a receiver may opt to choose several stable paths to join the group (when propagated, route information includes stability metrics). The receiver sets an ‘active’ flag in only one join to activate only one path at a time. Only active paths forward data packets. This reduces packet transmission overhead (a very significant factor in ad hoc networks) while maintaining robustness. If performance or stability of the active path degrades, or local recovery fails, the receiver may activate another path with high stability. Also, when the receiver moves, it may activate another path containing one of the new neighbors, thus achieving fast handoff using the concept of multicast-based mobility[1]. Our investigations show that, on average, a moving node traverses 2.5 hops to reach the nearest point of the multicast distribution structure in very large networks (with up to 5000 nodes)³. Moreover, mobility prediction[37] may be used to achieve advance joining, further reducing the effects of mobility handoff (i.e., mobility of the receiver in ad hoc networks). Rules for activating/de-activating should be carefully selected to avoid black holes. For example, if *any* member exists downstream a branch then the branch must be activated. In order for a branch to be inactive, all downstream branches must be inactive.

So far, we have assumed that the receivers’ efforts in searching for group information are successful, using the above localized search techniques. In case of sparse groups, where participants are far apart, or in case of multi-sender groups, where information obtained from other group members may not be complete, a *fallback* mechanism should be used. Such mechanism should be efficient, avoiding frequent global flooding and should be adaptive to membership dynamics. To achieve this, we introduce a novel *bootstrapping* anycast architecture for multicast service in ad hoc networks, discussed next.

3.3. Adaptive Resource Discovery and Multicast Address Allocation

The major research challenge for multicast resource discovery (i.e., discovery of group address, senders), is the lack of any (centralized or distributed) infrastructure to hold and distribute such information. Inter-domain multicast for wired networks[45][3] utilizes the AS hierarchy of the Internet and uses BGP extensions to distribute multicast routes that map group prefixes into root-domains (established by the multicast address allocation). In turn, receivers join towards the root-domain and senders send their packets towards it. Intra-domain multicast is used within the domains and is built on top of unicast routing. All these infrastructures (unicast routing, AS hierarchy) do not exist in ad hoc networks.

An architecture for *anycast* routing in the Internet was recently proposed in[31]. Utilizing hierarchical routing and aggregation. This work provides a scalable mechanism to discover members of anycast groups that are closer to the requester than other members. The architecture identifies two mechanisms, a low overhead mechanism, using default routes, for non-popular anycast groups (in which case the request is routed to the home domain, derived from the anycast address itself), and another mechanism for popular anycast groups that caches routes for nearby members.

³ These results were obtained for Internet topologies. We shall investigate these measures further in the context of ad hoc networks to study how they differ.

The only global infrastructure we can probably utilize in ad hoc networks is *geographic location*. Based on geographic multicast address allocation, we devise a new adaptive anycast architecture as follows. The multicast address space is broken into prefixes. Each multicast address prefix is assigned to a geographic region called the *rendezvous region (RR)*⁴. Nodes located in the *RR* have a collective responsibility of maintaining information about the groups belonging to the group prefix assigned to their current region. Since it is ‘collective’ and could be done by only a small subset of *SDS* nodes (say 3-7 uncorrelated nodes) we can use a probabilistic promotion scheme. Each node decides locally whether it will become a *SDS* based on its own configuration (some nodes maybe configured as servers), capabilities (e.g., GPS), power and stability estimates. If so, it obtains its (approximate) geographical location⁵ and determines the group prefix to which its current location maps, using algorithmic mapping in the general form of $f(x_1 \text{ to } x_2, y_1 \text{ to } y_2) = G_{prefix}$, or similar⁶. At that point, the node acts as a member of the anycast group of *SDS*s responsible for G_{prefix} , and advertises this information in its zone and to its contacts⁷. Other *SDS*s for the same prefix reply to update the new *SDS* (the reply is localized to reduce overhead). As *SDS* nodes move out of the *RR* for the corresponding G_{prefix} , they advertise their latest group information and leave notice to the *RR* (using geocast[43], for example), which increases the probability of other nodes promoting themselves to become *SDS*s. This ensures constant replenishing of the pool of *SDS*s serving as members of the anycast group for that *RR*.

The above scheme requires approximate knowledge of geographic location. We do not assume that all nodes are GPS capable. We do assume, however, that nodes are heterogeneous, and that *some* nodes will be more capable than others, specifically with localization capability⁸. We shall investigate GPS-less techniques[58][59] in our work.

Session Initiation A node initiating a multicast session is expected (without necessity) to use the above algorithmic mapping to obtain a multicast address that maps into a geographical vicinity as its *RR*. In any case, group/session initiation requests/updates are sent to the *RR* to avoid collisions in multicast address allocation. When a new sender of a group belonging to G_{prefix} becomes active it performs a localized broadcast (as described earlier) and, if far from its *RR*, issues an update to *RR*. The requests and updates are sent using *lollipop-LAR* (our modified location aided routing (LAR) [35]) to improve scalability. In *lollipop-LAR*, a far away sender chooses a contact that is closer to the *RR*. If the distance between the contact and the *RR* is less than a limit l , the contact sends the request/update to the *RR* directly using LAR, otherwise it chooses one of its contacts closer to the *RR*. Estimating l is part of our research.

When a node joins the group, it first attempts localized search for *SDS*. If this fails it sends/geocasts a join query to the *RR* using *lollipop-LAR*. A join reply is issued by a *SDS* in the *RR* and follows the backward route created by the join query (this path is only used for the reply so it need not be the shortest path). Once the receiver has group/sender information it sends a Join request as described earlier to establish multicast distribution path(s). Join latency is subject to further study.

Popularity-based Dynamic Adaptation If used for the common case, geocasting to the *RR* may incur a lot of overhead to maintain group information, especially for a very dynamic environment as is ad

⁴ This does not necessarily imply, however, that these groups are geographically limited to that region.

⁵ Geographic location update need not be done frequently, only when a node moves noticeable distances. Afterall, this information is approximate and is used for distributed resource discovery (not for forwarding packets).

⁶ We assume that nodes wishing to use multicast will express their desire to use multicast to their neighbors, zone or contacts (similar to a subscription), which will result in a reply with the well-known algorithmic mapping function. The scheme should allow for changing this mapping to another well-known mapping (very infrequently, though). The choice of rendezvous region and the algorithmic mapping function will be subject to our further investigation.

⁷ Alternatively, a node may advertise this information using localized broadcast, or geographically scoped broadcast (where only nodes, within a specific region, broadcast the packets), or Geocast[43]. We shall investigate these alternatives in our research.

⁸ This poses questions about accuracy of location estimation and its effect on performance of the proposed mechanisms. This is an important issue to study through investigation of space of possibilities of number of localization capable nodes, and their distribution (dense/sparse), under a wide array of mobility patterns. We shall address this issue in our proposed research.

hoc networks. Hence, our multicast service and resource discovery paradigm should adapt to dynamics of membership, to achieve better performance for popular groups. As explained above, sending or joining groups entails local advertisement (Adv) or query. This gives indications of the popularity of the group in the vicinity of the participants. Nodes receiving Advs and queries, those that are willing to become *SDS* for that group (based on their configuration, stability and capability), estimate the popularity of the group. The initial estimate is based on Advs/queries heard. If this initial estimate exceeds a threshold $pop_{query-th}$ a local group query is sent by the candidate-*SDS* to its own zone (using pro-active routing updates readily sent) and to its contacts. This group query gives a better estimate of group popularity Grp_{est} , in addition to information about existing *SDS*s nearby SDS_{est} . Popularity estimate pop_{est} is obtained as $pop_{est} = a \frac{Grp_{est}}{SDS_{est}}$. If

$pop_{est} > pop_{th}$ where pop_{th} is the popularity threshold (subject to further research), then a node advertises itself as *SDS* for the group to its zone and contacts, and contacts the *RR SDS* for updates (using lollipop-LAR). Future nearby join queries for the group reach the local *SDS* and are answered locally, reducing overhead and delay. Furthermore, this adaptive mechanism also achieves better robustness and continued operation during network partitions.

Discussion We note that the probability of success of the localized search is affected mainly by two factors. The first is the group address, obtained during session initiation, which decides the location of the *RR*, and in turn determines the location of the *RR-SDS*s. The second factor is the nearby popularity of the group nearby, which decides the promotion of nearby *SDS*s. Many of the offered services are expected to be *location-based* services, meaning it is targeted to a specific location. Hence, these groups will tend to be popular within certain locations (more than others). In addition, initiators are expected to choose group addresses that have *RR* in the geographical vicinity. Both these factors increase the probability of success for localized search, and lead us to believe that, in the common case, our architecture is capable of high performance, with low overhead and low delays. Our schemes work for global as well as locally-scoped groups.

Several questions arise for the above proposed architecture. Some of which we try to address here. For example, how do participants know about newly initiated sessions and their properties? This can be provided using the same scheme provided above, as follows. When multicast participants express their interest in multicast service they obtain the algorithmic mapping function (described above) as well as a *well-known session advertisement* group address. This address also has *RR*. As groups are initiated, they are updated at the *RR-SDS* and the local *SDS*s (if any), then information about new sessions is obtained as above. This provides an *bootstrapping* mechanism essential for multicast service.

Other questions include the frequency of update to the *SDS*s, and refresh of multicast state using soft-state. Also, what about multicast state kept in the nodes? How does such state scale with the number of active groups and sources? What if for many applications statistical information is sent periodically, but infrequently, to the whole group (as in RTCP). Should these infrequent senders be treated as regular senders or should they use a shared tree? What are the effects of node failures, network partitions and different mobility patterns on the performance? These issues are subject to further research.

4. Modeling, Analysis and Evaluation Methodology

Unlike traditional networks, ad hoc networks tend to be highly dynamic, mainly due to mobility. Mobility affects protocol behavior and performance directly and drastically. Some even suggest that many results obtained for ad hoc network studies merely reflect the mobility model used[61]. We believe that it is imperative to study and understand the effects of mobility on the performance and correctness of our proposed mechanisms. Specifically, we follow an approach that allows us to systematically relate certain mobility characteristics (e.g., movement correlation) on our mechanistic building blocks.

In general, previous studies on ad hoc networks used a single mobility model, usually the random way-point model. Such model fails to capture mobility correlation that may exist in real world scenarios. Instead, we propose a novel, yet simple, approach based on *scenarios* or state machines. In a specific state (e.g., pedestrian or vehicular) the node has a certain mobility model. Upon change of state (e.g., bus) the characteristics of the model change. Sequence of states is derived from real-life scenarios. Such model

includes richer mobility semantics; such as group mobility (e.g., bus), correlated mobility (e.g., vehicles on a congested highway), or random movement (e.g., pedestrian).

Often, network simulation is used to evaluate ad hoc routing protocols. Choosing a meaningful set of simulation and testing scenarios, however, is often hard. The *STRESS* methodology[5] attempts to alleviate such problem via test scenario synthesis.

We plan to use three different approaches for evaluation; systematic algorithms for robustness validation using *STRESS*, packet level simulation using the network simulator (NS-2)[4], and implementation in a laboratory setting. The tools shall be adapted adequately for our studies. We focus on analysis and understanding of mobility effects, rather than attempting to develop an accurate mobility model based on actual traces.

STRESS The *STRESS* methodology[5] (designed and developed by the PI) provides algorithms for *systematic and automatic* evaluation of networking protocols. Such methodology targets *multiparty protocols*, in that it is *different than traditional validation* techniques. To capture the nature of multicast protocols, *STRESS* provides a model for the *network topology* and the states of the different group members using a global finite state machine (GFSM) model. In addition, network *fault* models (such as packet losses, or machine crashes) are incorporated into the algorithms to facilitate robustness checking. In addition, this method introduced the concept of *topology synthesis* that enables backward search; found in our case studies to be much more efficient than traditional forward search. Hence, we believe that *STRESS* includes rich models and semantics needed for systematic evaluation and robustness verification of our proposed protocols. We are not aware of other work with these characteristics.

STRESS provides two main algorithms to search the protocol state space, *fault-independent test generation (FITG)* based on forward search[46], and *fault-oriented test generation (FOTG)* based on topology synthesis and backward search[6]. Case studies have been conducted using this methodology for analysis of multicast routing protocols and reliable multicast protocols, among others. However, *STRESS* has only been applied to LANs and end-to-end models (using a *virtual LAN* model). It has *not* been applied to multi-hop models. It also does not include rich mobility models. This entails *changing the connectivity* of a node (or several nodes to account for correlation). These mobility models need to be embedded into the methodology. The scenario-based approach we take for mobility modeling lends itself readily to a finite state machine (FSM) representation. Whereas correlation between node movements may be captured using the global FSM. We shall also investigate other aspects of mobility (such as velocity and direction) and location, in the context of *STRESS*. [We do not elaborate for brevity.] Scenarios resulting from the stress studies will carry on to the simulation stage.

Simulation The network simulator (NS-2)[4] (developed partially by the PI) supports a variety of multicast and ad hoc routing protocols, and will be used for simulations. Several evaluation metrics shall be used to evaluate the performance of the proposed architecture, including packet delivery ratio, join latency, route discovery delay, and control overhead, among others. The salient feature of our work, however, is the integration of physical constraints into the topology map. This facilitates modeling of mobility scenarios. For example, overlaying of highways, routes and bus movement enables us to capture different correlation models. We plan to develop a centralized *mobility coordinator* object to correlate node movements.

Implementation We plan to implement the proposed architecture in a laboratory test bed. (The USC EE networking laboratory is founded and directed by the PI). Using carefully designed experiments, the abilities of each component and their integration shall be demonstrated. The test bed currently consists of around 20 nodes (to be extended later) of laptops and iPAQs running Linux and wireless Motes running TinyOS. Also, the simulations may be run in emulation mode[17] (supported by NS-2) enabling simulations to interface to an operational network. Earlier developed tests will be leveraged. Details omitted for brevity.

5. Related Work

Related work lies in the areas of ad hoc routing (unicast and multicast), hierarchy and cluster formation, anycast architectures, inter-domain multicast, geocasting, and mobility modeling. Note that only closely related work is discussed, given the research interest given to ad hoc networks recently.

In the area of unicast ad-hoc routing, protocols are generally classified as either pro-active (or table-driven) or re-active (or on-demand) protocols. Pro-active protocols include DSDV[18], CGSR[19], and WRP[20], and rely upon routing updates to maintain consistency of route information. Re-active protocols include AODV[21], DSR[22], TORA[23], ABR[24], and SSA[25], and create routes only when required by the source node. One feature of SSA is that it selects routes based on the signal strength between nodes. Fisheye state routing (FSR)[65] is used to reduce routing update overhead of link state routing. For a node, the route update frequency to a certain destination is inversely proportional to the distance (in hops) of the destination. That is, routes to nodes within a small distance are sent to neighbors with higher frequency than routes to far away nodes. This reduces route overhead (not table size) and reduces accuracy of routing with distance. Routing efficiency decreases and delays increase, however, with dynamics of mobility, and routing table size grows linearly with network size[38]. CEDAR[40] uses a simple approximate algorithm was proposed for building a core graph consisting of the minimum dominating set of nodes. The mechanism was designed for a network of small (10s of nodes) to medium (100s) size network. A variant of CEDAR[41] was used for multicast by joining to the core graph. The effects of mobility and concentration on the core graph, were not clear in the study. Other works on multicast ad-hoc routing in[26][27] generally extend existing multicast routing for the Internet, such as PIM-SM[2]. Other recently proposed multicast ad hoc routing protocols include tree-based and mesh-based protocols. Tree-based protocols include AMRoute[33] and AMRIS[34]. AMRoute creates a bi-directional shared core-based tree using unicast tunnels, it uses virtual mesh links for tree creation and needs unicast, but incurs temporary loops and chooses sub-optimal routes with mobility. AMRIS uses a shared tree and an ID number per node, does not need unicast, broadcasts new-session messages and uses beacons to detect disconnection and re-joins to potential parents. However, it uses the expanding ring search mechanism for branch re-construction due to node failure, which does not scale well. Mesh-based protocols include ODMRP[36] and CAMP[47]. CAMP uses a shared mesh, and all nodes keep membership, routing and packet information. New members use expanding ring search to find other member neighbors. However, CAMP needs a special unicast protocol for its proper operation. ODMRP floods packets within mesh, but follows an on-demand policy for establishment and update of the mesh. It uses request and reply phases, broadcasts source announcements, and does not require unicast routing. The mesh is created when join requests from multiple receivers are sent to multiple-sources. Hence, for sparse groups or single-sender groups ODMRP may not be as robust. A local route recovery scheme[56] may be used to address this problem. In our routing protocol, we utilize the concept of mesh construction and local recovery, but we attempt to avoid floods for resource discovery, using a contact-based query approach. Moreover, we allow on-demand-activated multiple paths to be constructed to a single source, to increase robustness and achieve better handoff performance during mobility. Multi-path routing was proposed[49] for parallel data distribution. We will leverage mechanisms provided for multi-path discovery, but we do use multiple-paths for multicast differently, by activating on path at a time on-demand.

Recently, a scalable anycast architecture was proposed for the Internet[31], and was discussed earlier. Work on location-based routing was presented in location-aided routing LAR[35], Geocast[43][44]. We use concepts of geocasting for route and resource discovery. We modify it for scalability using lollipop shaped regions with the aid of contacts. We shall also leverage work on GPS-less positioning[58][59] to determine relative approximate positions of the nodes, assuming GPS capability in *some* nodes.

One of the earliest works on self-configuring hierarchical architectures includes work on the landmark hierarchy Landmark hierarchy uses the concept of area hierarchy. An area is constructed by grouping nodes into logical areas and subareas. Each node has a level in the hierarchy and a radius r associated with that level. Each node advertises information about itself to nodes within r hops. So, a node receives advertisements from nearby nodes that are the lowest level of the hierarchy, and faraway nodes that are at higher levels of the hierarchy, and so on. In effect, each node knows less information about farther nodes.

In[50] landmark hierarchy is used to form an object location architecture for sensor networks. Hierarchy levels are self-configuring and may be adapted using a promotion/demotion scheme. Drawbacks of landmark hierarchy were discussed earlier. LANMAR[37][38] uses the landmark hierarchy concepts to establish hierarchy in ad hoc networks. However, landmarks are used for sets of nodes moving together as a group to reduce routing information exchange. Other hierarchical ad hoc routing include the zone-based hierarchical link state (ZHLS)[51]. ZHLS is a global positioning system (GPS)-based routing protocol for ad hoc networks, where a network is divided into non-overlapping zones. A node only knows node connectivity within its zone and the zone connectivity for the network. This architecture does not use cluster head to mitigate traffic concentration, reduce routing protocol control/message exchange overhead and avoid single point of failure. It uses zone ID and node ID for routing. However, the zone map is defined by design for interzone routing, and hence does not adapt to network changes and dynamics. Another protocol called the zone routing protocol (ZRP)[28][52] discussed earlier. In[30] the ZRP approach is coupled with geographic (geodesic) routing for remote routing. In[13][53][54] the link availability model discussed earlier is proposed. The authors suggest to use it with a cluster based approach, in which a parent is selected based on the availability model to increase the lifetime of the cluster. Parent selection and cluster dynamics may complicate our architecture. Instead, we propose to incorporate the availability model with our modified ZRP approach and to use it for determining zone size, and in choosing contacts.

In the Internet, Hierarchical PIM[45] was proposed is an inter-domain architecture based on the PIM-SM protocol. It suggests a hierarchy of rendezvous points (equivalent to cluster heads) to communicate between multicast domains. The BGMP architecture[3] was proposed for hierarchical inter-domain multicast. It uses a bi-directional shared tree and the notion of a root domain. The problem of multicast address allocation is coupled with BGMP for the choice of the root domain. The same study proposes the MASC scheme for multicast address allocation. Such problem is still active in research.

In the area of mobility and movement modeling, several models have been developed for cellular and ad hoc networks. For cellular networks, the random-walk model was used in[7][8][9], where the next position is picked randomly then the speed and direction are picked uniformly from a given range. [7] studied the cell residence time and found the generalized gamma distribution to be suitable, while the negative exponential distribution was suitable for channel hold time modeling. Rappaport also studied the cell residence time (also called the dwell time) and used models based on multi-dimensional birth-death process. [10] used a model based on semi-Markov process to represent mobile user behavior. [11] noted that mobility patterns tend to increase during rush hours, while [12] used a group mobility model to evaluate ad hoc routing protocols. [13] used a random-walk and pause-time random models to derive probability of path availability as function of time, for clustering mechanisms in ad hoc networks. Other models include pursuit and tracking, nomadic and brownian motion.

These models affect the performance evaluation of our architecture. For example, group mobility adds correlation that is not exhibited by random moves, which may create traffic or state concentration in the network. We shall study these models, investigate their effect on performance, and possibly suggest improvements or modifications as necessary. Our protocols should be evaluated using a variety of mobility models.

6. Proposed Research

We propose to conduct in-depth studies to understand the architectural design trade-offs of the proposed protocols. We propose to pursue research in two major directions (1) *architectural and protocol design* and (2) *systematic evaluation methodology*. Particularly, we intend to investigate and attempt to answer the following questions (in addition to questions discussed earlier):

- How should the list of *contacts* be chosen? and what is the effect of introducing the *contact* list on the overall performance of the architecture? Some performance metrics are given in the *evaluation* section. What are the properties of the resulting network, as function of the probability of choosing contacts p , and how does it change with mobility? We believe that such network may approach a desirable small-world graph for careful choice of p and its parameters, resulting in reduction of route

and resource discovery overhead at the expense of increased latency than pure ZRP in some cases. We plan to investigate that trade-off in depth. Another (longer-term) issue includes several levels of contacts. We have only discussed nearby contacts (let us call them *level-1* contacts). Would it be beneficial to include farther contacts in the architecture (call them *level-2* or higher)? There is an interesting trade-off to investigate between increased contacts and increased maintenance overhead.

- What is the effect of using different algorithmic mapping functions, group initiation procedures, and *SDS* promotion scheme? We think that with the proper choice of the promotion parameters these effects will be reduced. How do we choose pop_{th} to achieve desirable effects? The effect of membership distribution will also have an effect, we expect. By thorough investigation, using various membership distribution patterns and densities, we hope to develop enough insight to suggest useful algorithms, or heuristics, to choose the parameter settings adaptively.
- How do characteristics of the *mobility model* affect behavior and performance of our protocols? We plan to investigate such effects thoroughly using a semantically rich scenario-based mobility model explained in the *evaluation* section. Our current intuition is that short and long term mobility characteristics matter. Specifically, we think correlated mobility, affecting node density, will have a strong effect. But to what extent? By using appropriate mobility and stability metrics we hope to reduce undesirable effects, where the mechanisms adapt intelligently to network dynamics. We plan to test our adaptive schemes by embedding mobility models into *STRESS* and NS-2. We plan to use a two-level FSM model. A *transportation* state (e.g., pedestrian, vehicular, train) captures the long-term characteristics and correlation of movement. While *mobility* states within a transportation state represents short-term characteristics (e.g., low or high mobility). Transportation literature is helpful in this process[63][64]. We also plan to develop a multi-hop model for *STRESS* using the notion of node *neighborhood*, which allows independent network view for each node.
- How are geographic information detected using partial capability of GPS in conjunction with GPS-free techniques? Furthermore, how is this information updated to the *SDSs* by the senders? We plan to use only approximate estimates of location, to reduce update frequency (as function of the sender mobility), and use location recovery and location-aided routing techniques to locate senders. Estimating these trade-offs and finding a reasonable mix of location techniques needs more research.
- What is the effect of *failures* on the performance and robustness? Power depletion, crashes, partitions and partition healing are but a few of these failures to study. We plan to fine-tune our adaptive anycast scheme and multicast routing to increase robustness. The effects of multi-path and channel fading, however, can only be studied in an implementation environment, and we plan to investigate them in our laboratory during the later stages of this research.
- How does the multicast state scale in the network nodes? This is an important aspect to analyze as function of the number of groups and sources. We leverage our previous work in that area[62].
- What is the quality of the routes discovered? We expect sub-optimal routes to result due to the local recovery mechanisms. We plan to study and analyze such mechanism in detail.

Other problems to investigate include modified or new mobility estimation and prediction schemes that attempt to avoid limitations of existing methods (such as depending on random models or using simple free-space model). Also, integrating QoS into these models, (as hinted in[13]) is to be investigated.

Longer-term research issues to pursue include the use of mobility prediction schemes for fast handoff using Multicast-based mobility concepts[1], the study of asymmetry and uni-directional links on route quality, and possible development of localized shared trees for low-rate, densely populated sources.

7. Networking Educational Plan

This educational plan includes two elements. (1) A series of courses (accompanied by a networking laboratory) to be developed and instructed by the PI to improve theoretical and practical aspects of the networking curriculum at USC. (2) A new approach to involve students in various aspects of the classroom, including course preparation, evaluation and examination.

(1) Lab Courses Plan (Practical Networking)

This plan includes a series of networking courses (integrated with the networking lab) both at the undergraduate and graduate levels, over the next few years.

- EE-555 Broadband Networking Architecture: LANs, WANs, congestion control, queuing disciplines, traffic characterization, admission control, switching, optical networks, multicast routing, wireless networking and IP mobility.
- EE-599 Internet Protocol Design, Simulation and Testing: DiffServ and AQM, unicast, multicast, geocast and hierarchies in ad hoc networks, TCP and reliable multicast over wireless, Stress testing, multicast congestion control. (Developed and instructed by the PI. For graduate students)
- EE-499 Introductory Computer and Wireless Networking Laboratory: packet, route and topology trace and measurement tools, route configuration and test bed setup, simulation vs. implementation, congestion control, mobile IP, wireless ad hoc networks. (Developed and instructed by the PI. For undergraduate and entry-level graduate students)
- TBD Advanced Wireless Networking Lab: IP mobility with fast handoff, ad hoc networks, sensor networks.
- TBD Protocol testing, verification and simulation: (with VLSI chip testing)

The PI has founded and is directing the networking lab at the EE department at USC.

(2) Student-Assisted Education and Evaluation

The PI proposes the concept of *student-assisted grading* and *course-development*, in which students grade assignments (as double-blind review), based on *adjustment* questions and supervision of the instructor. Students also participate by suggesting experiments, assignments and exam questions. A novel *instant feedback* system is proposed for *in/during/out-of-class* evaluation. This new approach for student participation gives the students the opportunity to interact, innovate, and review technical material in a new way, which would enhance their level of understanding.

Student-assisted grading system:

This is a new approach for assignment grading and evaluation, especially in a lab setting, to be included as one of the experiments in the classes. In order to evaluate experiment assignments, each student is assigned a number of assignments to grade for classmates (anonymously, using double-blind review). Before the students start the grading, they need to answer several *adjustment* questions⁹ carefully chosen by the instructor. The answers provided by a student give a *weight* for the student, denoting the level of understanding. After the students grade the assignments, the grades provided are adjusted according to the grader's weight. The instructor samples these graded assignments to evaluate the process and its fairness. This will be coupled with an incentive system that rewards good grading and punishes dishonesty and inconsistency (measured as high variation between adjustment question performance and grading performance).

This grading experiment provides a chance for the students to view colleagues' work in detail, to learn and be able to improve and compare their own performance with that of their colleagues. Since students grade essay questions or experiment reports, they have to read the answers in depth and understand them. This should increase their level of understanding and knowledge, and reflect on their answers in the final exams. The assignment will be phased in gradually into the courses (i.e., initially they will not carry a lot of weight).

To evaluate the initial effects of this method, the grading assignment will be optional and will be given after the midterm exams. Improvement in grades between midterms and the final exams for students participating in the assignment will be compared to that of the other students¹⁰.

⁹ That is, to *adjust* the graded score according to the score of these questions that measure the level of understanding of the material at hand.

¹⁰ Normal correlation between midterm exam grades and final exam grades (i.e., without this assignment) should also be factored in.

Student-assisted course development

To stimulate creativity in creating lab experiments, the students are asked to design new experiments for current and future classes. Each student shall design one experiment individually and two in a team. The students are also asked to provide sets of questions out of which a subset is chosen by the instructor to be included in upcoming assignments or exams. Guidelines are given to students to help develop high quality and relevant questions. These questions and their answers are graded for quality.

On the first day of class a *diagnostic exam* is given that covers background and fundamentals as well as elementary material to be taught in class. This gives indication of students' weaknesses and strengths. In addition, feedback forms are collected twice a semester (aside from conventional evaluation at end of semester). Furthermore, short 5-minute exams, automatically computer-graded, are given at the beginning and end of each lecture (called, *in-class/out-of-class* exams). Also, a novel *instant-feedback* system will be used during each lecture. In such system, several (e.g., 4) multiple-choice-questions are given by the instructor. Each student in class has a screen to view the questions and answers, as well as a small device to enter the answer (a, b, c or d). The order of the questions and answers on each screen is randomized to minimize cheating. The answers are automatically graded¹¹. Statistics of these answers may be automatically correlated with overall student performance (e.g., GPA, diagnostic and midterm exam scores) to help identify troubled students and provide better academic advice.

These procedures will be designed to measure the knowledge gained in class, to give students motivation to pay attention in class and to raise the students' level of understanding and education.

Broader Impact

The potential impacts of this research are significant. If successful, this research will provide the first *architecture for multicast service* support in *large-scale ad hoc* networks, enabling a very important class of applications in future networks. Our mechanisms are *highly adaptive* to network dynamics and *mobility*, which renders our architecture more *robust, efficient* and *scalable*. *Self-configurability* of our architecture facilitates ease-of-deployment. We also provide the first architecture for *adaptive anycast* in ad hoc networks. Such mechanism can potentially provide the *resource discovery* component for a wide-array of future applications and middle-ware. In addition, our *systematic evaluation* framework will provide in-depth understanding of the *effects of mobility* on ad hoc network protocols (an effect that is significant but not well-understood). One output of this research is a set of test-suites that can, potentially, provide a meaningful and richer *benchmarks* for evaluating ad hoc networks, based on semantically rich *mobility models*. Developed tools, simulation, implementation code, lab experiments, test scenarios, and publications will be publicly available through the project's web site. It will also be integrated into the networking curriculum and courses at USC at both the undergraduate and graduate levels.

The research plan is complemented by a strong *educational plan*, potentially leading to significant *improvements in the learning* process. The plan includes a wide-array of essential *courses* on Internet and ad hoc networking protocols, and the establishment of a joint research/instruction wireless networking *laboratory*. In addition, the PI provides a *unique and effective* approach to increase student involvement in classroom activities and new methods for better *evaluation of student abilities*, using *student-assisted* grading and course development. These methods will provide a *model for improving education*.

¹¹ A similar concept will also be used to provide instant feedback on instructor performance (e.g., on volume of voice, clarity or repetition, etc.).

References:

- [1] A. Helmy, "A Multicast-based Protocol for IP Mobility Support", ACM Second International Workshop on Networked Group Communication (NGC 2000), ACM SIGCOMM-sponsored, Palo Alto, November 2000.
- [2] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, V. Jacobson, M. Handley, C. Liu, P. Sharma, "Protocol Independent Multicast – Sparse Mode (PIM-SM): Protocol Specification", *RFC 2362/2117 of the Internet Engineering Task Force (IETF), Inter-Domain Multicast Routing (IDMR)*, March '97/'98.
- [3] S. Kumar, P. Radoslavov, D. Thaler, C. Alaettinoglu, D. Estrin, M. Handley, "The MASC/BGMP Architecture for Inter-domain Multicast Routing", *Proceedings of ACM SIGCOMM*, August 1998.
- [4] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, H. Yu, "Advances in Network Simulation", *IEEE Computer*, vol. 33, No. 5, p. 59-67, May 2000.
- [5] A. Helmy, STRESS Systematic Test Synthesis for Multipoint Protocol Design, Ph.D. Dissertation, USC-CS-TR-99-716, Oct. 1999.
- [6] A. Helmy, D. Estrin, S. Gupta, Fault-oriented Test Generation for Multicast Routing Protocol Design, Proc. FORTE/PSTV '98, IFIP, Nov. 1998.
- [7] M. Zonoozi, P. Dassanayake, "User mobility modeling and characterization of mobility patterns", *IEEE Journal on Selected Areas in Communications*, 1239-1252, September 1997.
- [8] A. Bar-Noy, I. Kessler, M. Sidi, "Mobile users: to update or not to update?", *INFOCOM'94*, pp 570-576, 1994.
- [9] I. Rubin, C. Choi, "Impact of the location area structure on the performance of signaling channels in wireless cellular networks", *IEEE Communications Magazine*, pp 108-115, February 1997.
- [10] H. Kobayashi, S. Yu, B. Mark, "An Integrated Mobility and Traffic Model for Resource Allocation in Wireless Networks", *WOWMOM*, Boston, MA, 2000.
- [11] M. Chung, T. Kim, H. Cho, D. Sung, "Modeling of Terminal to Evaluate the Number of Location Updates", *ICC'97*, Montreal, 1997.
- [12] X. Hong, M. Gerla, G. Pei, C. Chiang, "A group mobility model for ad hoc wireless networks", *2nd ACM Int'l workshop on modeling, analysis and simulation of wireless and mobile systems*, pp 53-60, 1999
- [13] A. McDonald, T. Znati, "A mobility-based framework for adaptive clustering in wireless ad hoc networks", *IEEE Journal on Selected Areas in Communications*, V. 17 (8), pp 1466-1487, Aug. 1999.
- [14] D. Estrin, M. Handley, A. Helmy, P. Huang, D. Thaler, "A Dynamic Bootstrap Mechanism for Rendezvous-based Multicast Routing", *Proceedings of IEEE INFOCOM '99*, New York, March 1999.
- [15] S. Deering, D. Cheriton, "Multicast routing in data internetworks and extended lans", *ACM Transactions on Computer Systems*, pp 85-111, May 1990.
- [16] A. Helmy, S. Gupta, D. Estrin, A. Cerpa, Y. Yu, "Systematic Performance Evaluation of Multipoint Protocols", *Proceedings of Formal Description Techniques & Protocol Specification, Testing, and Verification (FORTE/PSTV)*, IFIP, Kluwer Academic Publication, Pisa, Italy, p. 187-202, October 2000.
- [17] Kevin Fall. Network Emulation in the Vint/NS Simulator, *ISCC '99*.
- [18] C. E. Perkins, P. Bhagwat, Highly Dynamic Destination-Sequenced Distance Vector Routing (DSDV) for Mobile Computers, *Comp. Commun. Rev.*, Oct. 1994, pp. 234-44.
- [19] C.-C. Chiang, Routing in Clustered Multihop, *Mobile Wireless Networks with Fading Channel*, Proc. IEEE SICON '97, pp. 197-211.
- [20] S. Murthy, J. J. Garcia-Luna-Aceves, An Efficient Routing Protocol for Wireless Networks, *ACM Mobile Networks and App. J.*, Special Issue on Routing in Mobile Communication Networks, Oct. 1996, pp. 183-197.
- [21] C. E. Perkins, E. M. Royer, Ad-hoc On-Demand Distance Vector Routing, Proc. 2nd IEEE Wksp. Mobile Comp. Sys. And Apps., Feb. 1999, pp. 90-100.

- [22] D. B. Johnson, D. A. Maltz, Dynamic Source Routing in Ad-Hoc Wireless Networks, Mobile Computing, 1996, pp.153-181.
- [23] V.D. Park, M.S. Corson, A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks, Proc. INFOCOM '97, Apr. 1997.
- [24] C-K. Toh, A Novel Distributed Routing Protocol To Support Ad-Hoc Mobile Computing, Proc. 1996 IEEE 15th Annual Int'l. Phoenix Conf. Comp. And Commun., Mr. 1996, pp. 480-486.
- [25] R. Dube et al., Signal Stability based Adaptive Routing (SSA) for Ad-Hoc Mobile Networks, IEEE Pers. Commun., Feb. 1997, pp. 36-45.
- [26] J. J. Garcia-Luna-Aceves, E. L. Madruga, A Multicast Routing Protocol for Ad-Hoc Networks, Proc. IEEE INFOCOM '99, pp. 784-792, 1999.
- [27] M. Gerla, C.-C. Chiang, L. Zhang, Tree Multicast Strategies in Mobile, Multihop Wireless Networks, ACM Mobile Networks and Apps. J., 1998.
- [28] Z. Haas, "A new routing protocol for the reconfigurable wireless networks", *IEEE Conference on Universal Personal Communications*, Vol. 2, pp. 562-566, 1997.
- [29] D. J. Watts. In *Small Worlds, The dynamics of networks between order and randomness*. Princeton University Press, 1999.
- [30] L. Blazevic, S. Giordano, J. Boudec, "self-organizing wide-area routing", SCI/ISAS July 2000,
- [31] D. Katabi, J. Wroclawski, "A framework for scalable global IP anycast (GIA), Proceedings of ACM Sigcomm 2000.
- [32] P. F. Tsuchiya, "The Landmark Hierarchy: A new hierarchy for routing in very large networks", *Computer Communication Review*, Vol. 18, no. 4, pp. 35-42, Aug. 1988.
- [33] E. Bommaiah, M. Liu, A. McAuley, R. Talpade, "Ad-hoc Multicast Routing Protocol", *Internet-draft*, Aug. 1998.
- [34] C. Wu, Y. Tay, C. Toh, "Ad hoc Multicast Routing Protocol utilizing Increasing id-numbers (AMRIS): functional specification", *Internet-draft*, November 1998.
- [35] Y. Ko, N. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks", *Wireless Networks* 6, 4, p. 307-321, July 2000.
- [36] S. Lee, M. Gerla, C. Chiang, "On-demand multicast routing protocol", IEEE WCNC, p. 1298-1302, vol. 3, 1999.
- [37] S. Lee, W. Su, M. Gerla, "Ad hoc wireless multicast with mobility prediction", IEEE ICCN, p. 4-9, 1999.
- [38] P. Guangyu, M. Gerla, X. Hong, "LANMAR: landmark routing for large scale wireless ad hoc networks with group mobility", *MobiHOC '00*, p. 11-18, 2000.
- [39] C. Chiang, M. Gerla, L. Zhang, "Forwarding Group Multicast Protocol (FGMP) for Multihop, Mobile Wireless Networks", *ACM/Kluwer Journal of Cluter Computing*, vol. 1, no. 2, 1998.
- [40] R. Sivakumar, P. Sinha, V. Bharghavan, "CEDAR: a core-extraction distributed ad hoc routing algorithm", IEEE JSAC vol. 17, 8, p. 1454-1465, Aug. 1999.
- [41] P. Sinha, R. Sivakumar, V. Bharghavan, "MCEDAR: multicast core-extraction distributed ad hoc routing", IEEE WCNC, p. 1313-1317, 1999.
- [42] D. Estrin, D. Farinacci, A. Helmy, V. Jacobson, L. Wei, "Protocol Independent Multicast – Dense Mode (PIM-DM): Protocol Specification", Proposed RFC of the *IETF/IDMR*, September 1996.
- [43] Y. Ko, N. Vaidya, "Geocasting in mobile ad hoc networks: location-based multicast algorithms", IEEE WMCSA, p. 101-110, 1999.
- [44] Y. Ko, N. Vaidya, "GeoTORA: a protocol for geocasting in mobile ad hoc networks", IEEE ICNP, p. 240-250, 2000.
- [45] S. Deering, W. Fenner, D. Estrin, A. Helmy, D. Farinacci, L. Wei, M. Handley, V. Jacobson, D. Thaler, "Hierarchical PIM for Inter-Domain Multicast Routing", *Internet-Draft (I-D) of the IETF/IDMR*, December 1995.
- [46] A. Helmy, D. Estrin, S. Gupta, "Systematic Testing of Multicast Routing Protocols: Analysis of Forward and Backward Search Techniques", IEEE ICCCN, p. 590-597, October 2000.

- [47] J. J. Aceves, E. Madruga, "The Core-Assisted Mesh Protocol", *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1380-1394, August 1999.
- [48] A. Nasipuri, S. Das, "On-Demand Multipath Routing for Mobile Ad Hoc Networks", ICCN, p. 64-70, 1999.
- [49] S. Das, A. Mukherjee, S. Bandyopadhyay, K. Paul, D. Saha, "Improving quality-of-service in ad hoc wireless networks with adaptive multi-path routing", *GlobeCom*, p. 261-265, vol. 1, 2000.
- [50] S. Kumar, C. Alaettinoglu, D. Estrin, "Scalable Object-tracking through Unattended Techniques (SCOUT)", *IEEE ICNP*, Nov. 2000.
- [51] M. Ng, I Lu, "A peer-to-peer zone-based two-level link state routing for mobile ad hoc networks", *IEEE Journal on Selected Areas in Communications*, Vol. 17, no. 8, pp. 1415-1425, Aug. 1999.
- [52] M. Pearlman, Z. Haas, "Determining the optimal configuration for the zone routing protocol", *IEEE JSAC*, p. 1395-1414, vol. 17, 8, Aug 1999.
- [53] A. McDonald, T. Znati, "A path availability model for wireless ad-hoc networks", *IEEE WCNC*, p. 35-40, vol. 1, 1999.
- [54] A. McDonald, T. Znati, "Predicting node proximity in ad-hoc networks: a least overhead adaptive model for selecting stable routes", *MobiHOC*, p. 29-33, 2000.
- [55] D. Watts, S. Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, Vol. 393, June 4, 1998.
- [56] M. Ize, Y. Kim, "PatchODMRP: an ad-hoc multicast routing protocol", *ICIN*, p. 537-543, 2001.
- [57] P. Basu, N. Khan, T. Little, "A mobility based metric for clustering in mobile ad hoc networks", *ICDCS Workshop*, p. 413-418, 2001.
- [58] N. Bulusu, J. Heidemann, D. Estrin, "GPS-less Low Cost Outdoor Localization For Very Small Devices", *IEEE Personal Communications*, Vol. 7, 5, p. 28-34, October 2000.
- [59] S. Capkun, M. Hamdi, J. Hubaux, "GPS-free positioning in mobile ad-hoc networks", *ICSC*, p. 3481-3490, 2001.
- [60] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, V. Jacobson, M. Handley, C. Liu, P. Sharma, L. Wei, "Protocol Independent Multicast (PIM): Motivation and Architecture", Proposed RFC of the IETF/IDMR, October 1996.
- [61] A. Battacharya, S. Das, "LeZi-Update: An Information-Theoretic Approach to Track Mobile Users in PCS Networks", *MobiCom '99*.
- [62] A. Helmy, "State Analysis and Aggregation Study for Multicast-based Micro Mobility", *ACM-LANL-NCSTRL cs.NI/0105031*, May 2001.
- [63] H. Adeli, A. Karim, "Fuzzy-Wavelet RBFNN Model for Freeway Incident Detection", *Journal of Transportation Engineering*, Dec. 2000.
- [64] H. Zhang, "Recursive Prediction of Traffic Conditions with Neural Networks Models", *Journal of Transportation Engineering*, Dec. 2000.
- [65] G. Pei, M. Gerla, T. Chen, "Fisheye State Routing: A Routing Scheme for Ad Hoc Wireless Networks", *ICC* 2000.