1 INTRODUCTION

There are increasing interests and importance in supporting group communications over Mobile Ad Hoc Networks (MANETs). Example applications include the exchange of messages among a group of soldiers in a battlefield, communications among the firemen in a disaster area, and the support of multimedia games and teleconferences. With a one-to-many or many-to-many transmission pattern, multicast is an efficient method to realize group communications. However, there is a big challenge in enabling efficient multicasting over a MANET whose topology may change constantly.

Conventional MANET multicast protocols [3]–[8], [28] can be ascribed into two main categories, tree-based and mesh-based. However, due to the constant movement as well as frequent network joining and leaving from individual nodes, it is very difficult to maintain the tree structure using these conventional tree-based protocols (e.g., MAODV [3], AMRIS [4], MZR [5], MZR [28]). The mesh-based protocols (e.g., FGMP [6], Core-Assisted Mesh protocol [7], ODMRP [8]) are proposed to enhance the robustness with the use of redundant paths between the source and the destination pairs. Conventional multicast protocols generally do not have good scalability due to the overhead incurred for route searching, group membership management, and creation and maintenance of the tree/mesh structure over the dynamic MANET.

For MANET unicast routing, geographic routing protocols [11]–[14] have been proposed in recent years for more scalable and robust packet transmissions. The existing geographic routing protocols generally assume mobile nodes are aware of their own positions through certain positioning system (e.g., GPS), and a source can obtain the destination position through some type of location service [15] [16]. In [13], an intermediate node makes its forwarding decisions based on the destination position inserted in the packet header by the source and the positions of its one-hop neighbors learned from the periodic beaconing of the neighbors. By default, the packets are greedily forwarded to the neighbor that allows for the greatest geographic progress to the destination. When no such a neighbor exists, perimeter forwarding is used to recover from the local void, where a packet traverses the face of the planarized local topology subgraph by applying the right-hand rule until the greedy forwarding can be resumed.

Similarly, to reduce the topology maintenance overhead and support more reliable multicasting, an option is to make use of the position information to guide multicast routing. However, there are many challenges in implementing an efficient and scalable geographic multicast scheme in MANET. For example, in unicast geographic routing, the destination position is carried in the packet header to guide the packet forwarding, while in multicast routing, the destination is a group of members. A straightforward way to extend the geography-based transmission from unicast to multicast is to put the addresses and positions of all the members into the packet header, however, the header overhead will increase.
significantly as the group size increases, which constrains the application of geographic multicasting only to a small group [17]–[19]. Besides requiring efficient packet forwarding, a scalable geographic multicast protocol also needs to efficiently manage the membership of a possibly large group, obtain the positions of the members and build routing paths to reach the members distributed in a possibly large network terrain. The existing small-group-based geographic multicast protocols [17]–[19] normally address only part of these problems.

In this work, we propose an efficient geographic multicast protocol, EGMP, which can scale to a large group size and large network size. The protocol is designed to be comprehensive and self-contained, yet simple and efficient for more reliable operation. Instead of addressing only a specific part of the problem, it includes a zone-based scheme to efficiently handle the group membership management, and takes advantage of the membership management structure to efficiently track the locations of all the group members without resorting to an external location server. The zone structure is formed virtually and the zone where a node is located can be calculated based on the position of the node and a reference origin. In topology-based cluster construction, a cluster is normally formed around a cluster leader with nodes one hop or k-hop away, and the cluster will constantly change as network topology changes. In contrast, there is no need to involve a big overhead to create and maintain the geographic zones proposed in this work, which is critical to support more efficient and reliable communications over a dynamic MANET. By making use of the location information, EGMP could quickly and efficiently build packet distribution paths, and reliably maintain the forwarding paths in the presence of network dynamics due to unstable wireless channels or frequent node movements.

In summary, our contributions in this work include:

1) Making use of the position information to design a scalable virtual-zone-based scheme for efficient membership management, which allows a node to join and leave a group quickly. Geographic unicast is enhanced to handle the routing failure due to the use of estimated destination position with reference to a zone and applied for sending control and data packets between two entities so that transmissions are more robust in the dynamic environment.

2) Supporting efficient location search of the multicast group members, by combining the location service with the membership management to avoid the need and overhead of using a separate location server.

3) Introducing an important concept zone depth, which is efficient in guiding the tree branch building and tree structure maintenance, especially in the presence of node mobility. With nodes self-organizing into zones, zone-based bi-directional-tree-based distribution paths can be built quickly for efficient multicast packet forwarding.

4) Addressing the empty zone problem, which is critical in a zone-based protocol, through the adaption of tree structure.

5) Evaluating the performance of the protocol through quantitative analysis and extensive simulations. Our analysis results indicate that the cost of the protocol defined as the per-node control overhead remains constant regardless of the network size and the group size. Our simulation studies confirm the scalability and efficiency of the proposed protocol.

We organize the rest of this paper as follows. In Section 2, we discuss some related work. We present a detailed design of the EGMP protocol in Section 3, and quantitatively analyze the per-node cost of EGMP in Section 4. Finally, we give our simulation results in Section 5 and conclude the paper in Section 6.

2 Related Work

In this section, we first summarize the basic procedures assumed in conventional multicast protocols, and then introduce a few geographic multicast algorithms proposed in the literature.

Conventional topology-based multicast protocols include tree-based protocols (e.g., [3]–[5], [28]) and mesh-based protocols (e.g., [6], [8]). Tree-based protocols construct a tree structure for more efficient forwarding of packets to all the group members. Mesh-based protocols expand a multicast tree with additional paths which can be used to forward packets when some of the links break. Although efforts were made to develop more scalable topology-aware protocols [7], the topology-based multicast protocols are generally difficult to scale to a large network size, as the construction and maintenance of the conventional tree or mesh structure involve high control overhead over a dynamic network. The work in [26], [27] attempts to improve the stateless multicast protocol [2], which allows it a better scalability to group size. In contrast, EGMP uses a location-aware approach for more reliable membership management and packet transmissions, and supports scalability for both group size and network size. As the focus of our paper is to improve the scalability of location-based multicast, a comparison with topology-based protocols is out of the scope of this work. However, we note that at the similar mobility and system set-up, the delivery ratio of [26] is much lower than that of EGMP, and the delivery ratio in [27] varies significantly as the group size changes. In addition, topology-based routing by nature is more vulnerable to mobility and long path transmission, which prevents topology-based protocols from scaling to a large network size.

Besides the need of managing group membership as well as constructing and maintaining a multicast structure, a geographic multicast protocol also requires a location service [15] [16] to obtain the positions of the members. The geographic multicast protocols presented in [17], [18] and [19] need to put the information of the entire tree or all the destinations into packet headers, which would create a big header overhead when the group size is large and constrain these protocols to be used only for small groups. In DSM [17], each node floods its location in the network. A source constructs a Steiner tree and encodes the multicast tree into each packet, and delivers the packet by using source routing. LGT [18] requires each group member to know the locations of all other group members, and
proposes two overlay multicast trees: a bandwidth-minimizing LGS tree and a delay-minimizing LGK tree. In PBM [19], a multicast source node finds a set of neighboring, next-hop nodes and assigns each packet destination to one next-hop node. The next-hop nodes, in turn, repeat the process. Thus, no global distribution structure is necessary. GMP in [29] attempts to build a more efficient multicast tree through a centralized calculation for tree construction, and is also more applicable for a smaller group. The focus of EGMP, however, is to improve the scalability and efficiency of geometric multicast.

The HRPM [30] and SPBM [20] are more related to our work, as they also support hierarchical group management. HRPM consists of two key design ideas: 1) hierarchical decomposition of a large group into a hierarchy of recursively organized manageable-sized subgroups and 2) the use of distributed geographic hashing to construct and maintain such a hierarchy. Although it is interesting to apply hashing to find the rendezvous point (RP) for the network to store and retrieve state information, the hashed location is obtained with the assumption of the network size, which is difficult for a dynamic network. Also, as the hashed location is virtual, it is possible that the nodes could not find the (consistent) RP. This can happen when a message (e.g., Join) reaches a node whose transmission range covers the virtual point, but the node is neither the one closest to the RP, nor aware of the node (which may be out of its transmission range) closest to the RP. The mobility of nodes will introduce additional challenge to the protocol, which may not only result in frequent RP handoff, but also increase the chance of RP search inconsistency and failure. Additionally, requiring a node to contact RP first for a Join will increase joining delay. In contrast, EGMP does not make any assumption of the network size in advance, and the change of the membership of a zone does not need to be sent to a far-away RP but only needs to be updated locally. Instead of using one RP as a core for group membership management, which may lead to a point of failure, EGMP introduces the root zone which is much more stable than a single point, and manages group membership more efficiently within the local range. Instead of using the overlay-based multiple unicast transmissions, EGMP takes advantage of the promiscuous mode transmission to forward packets along more efficient transmission paths. We did not directly compare our work with HRPM, as we do not know the hashing algorithm used and a different RP distribution scheme would lead to different performance. However, we evaluated the performance of EGMP using a much larger default network size, which is known to have much more challenge to ensure reliable multicast transmissions than a smaller network.

In SPBM [20], the network terrain is divided into a quadtree with L levels. The top level is the whole network and the bottom level is constructed by basic squares. Each higher level is constructed by larger squares with each square covering four smaller squares at the next lower level. All the nodes in a basic square are within each other’s transmission range. At each level, every square needs to periodically flood its membership into its upper level square. Such periodic flooding is repeated for every two neighboring levels and the top level is the whole network region. Significant control overhead will be generated when the network size increases as a result of membership flooding. With this proactive and periodic membership updating scheme, the membership change of a node may need to go through L levels to make it known to the whole network, which leads to a long multicast group joining time. Instead of using multiple levels of flooding for group membership management, EGMP uses more efficient zone-based tree structure to allow nodes to quickly join and leave the group. EGMP introduces root zone and zone depth to facilitate simple and more reliable group membership management. EGMP does not use any periodic network-wide flooding, thus it can be scalable to both the group size and network size.

Finally, a lot of work have been done on geocasting [31], [32], [34]. Different from general multicasting, in which the destinations are a group of receivers, the destination of geocasting is one or multiple geographic regions (squares are normally defined). When packets reach the destined region, they will be sent to the nodes in the region through flooding or other methods. There is no need of forming multicast infrastructure to deliver packets to group members that may distribute widely in the whole network domain and change their positions as nodes move.

In [33], we proposed an efficient and robust geographic multicast protocol for MANET. In this paper, we further introduce zone-supported geographic forwarding to reduce the routing failure, and provide mechanism to handle zone partitioning. In addition, we introduce a path optimization process to handle multiple paths, and provide a detailed cost analysis to demonstrate the scalability of the proposed routing scheme.

3 Efficient Geographic Multicast Protocol

In this section, we will describe the EGMP protocol in details. We first give an overview of the protocol and introduce the notations to be used in the rest of the paper in Section 3.1. In Sections 3.2 and 3.3, we present our designs for the construction of zone structure and the zone-based geographic forwarding. Finally, in Sections 3.4, 3.5 and 3.6, we introduce our mechanisms for multicast tree creation, maintenance and multicast packet delivery.

3.1 Protocol Overview

EGMP supports scalable and reliable membership management and multicast forwarding through a two-tier virtual-zone-based structure. At the lower layer, in reference to
a pre-determined virtual origin, the nodes in the network self-organize themselves into a set of zones as shown in Fig. 1, and a leader is elected in a zone to manage the local group membership. At the upper layer, the leader serves as a representative for its zone to join or leave a multicast group as required. As a result, a network-wide zone-based multicast tree is built. For efficient and reliable management and transmissions, location information will be integrated with the design and used to guide the zone construction, group membership management, multicast tree construction and maintenance, and packet forwarding. The zone-based tree is shared for all the multicast sources of a group. To further reduce the forwarding overhead and delay, EGMP supports bi-directional packet forwarding along the tree structure. That is, instead of sending the packets to the root of the tree first, a source forwards the multicast packets directly along the tree. At the upper layer, the multicast packets will flow along the multicast tree both upstream to the root zone and downstream to the leaf zones of the tree. At the lower layer, when an on-tree zone leader receives the packets, it will send them to the group members in its local zone.

Many issues need to be addressed to make the protocol fully functional and scalable. The issues related to zone management include: the schemes for more efficient and robust zone construction and maintenance, the strategies for election and maintenance of a zone leader with minimum overhead, zone partitioning as a result of severe wireless channels or signal blocking, potential packet loss when multicast members move across zones. The issues related to packet forwarding include: the efficient building of multicast paths with the zone structure, the handling of empty zone problem, the efficient tree structure maintenance during node movements, the reliable transmissions of control and multicast data packets, and obtaining location information to facilitate our geometric design without resorting to an external location server.

For the convenience of presentation, we first introduce the terminologies used in the paper. In EGMP, we assume every node is aware of its own position through some positioning system (e.g., GPS [10]) or other localization schemes. The forwarding of data packets and most control messages is based on the geographic unicast routing protocol GPSR [13] described in Section 1. EGMP, however, does not depend on a specific geographic unicast protocol.

Some of the notations to be used are:

**zone**: The network terrain is divided into square zones as shown in Fig. 1.

\[ r \]: Zone size, the length of a side of the zone square. The zone size is set to \( r \leq r_t/\sqrt{2} \), where \( r_t \) is the transmission range of the mobile nodes. To reduce intra-zone management overhead, the intra-zone nodes can communicate directly with each other without the need of any intermediate relays.

**zone ID**: The identification of a zone. A node can calculate its zone ID \((a,b)\) from its position coordinates \((x,y)\) as:

\[
\begin{align*}
 a &= \left\lfloor \frac{x - x_0}{r} \right\rfloor, \\
 b &= \left\lfloor \frac{y - y_0}{r} \right\rfloor,
\end{align*}
\]

where \((x_0, y_0)\) is the position of the virtual origin, which can be a known reference location or determined at network setup time. A zone is **virtual** and formulated in reference to the virtual origin. For simplicity, we assume all the zone IDs are positive.

**zone center**: For a zone with ID \((a,b)\), the position of its center \((x_c, y_c)\) can be calculated as:

\[
\begin{align*}
 x_c &= x_0 + (a + 0.5) \times r, \\
 y_c &= y_0 + (b + 0.5) \times r.
\end{align*}
\]

A packet destined to a zone will be forwarded towards the center of the zone.

**zLdr**: Zone leader. A zLdr is elected in each zone for managing the local zone group membership and taking part in the upper tier multicast routing.

**tree zone**: The zones on the multicast tree. The tree zones are responsible for the multicast packet forwarding. A tree zone may have group members or just help forward the multicast packets for zones with members.

**root zone**: The zone where the root of the multicast tree is located.

**zone depth**: The depth of a zone is used to reflect its distance to the root zone. For a zone with ID \((a,b)\), its depth is:

\[
\text{depth} = \max(|a_0 - a|, |b_0 - b|),
\]

where \((a_0, b_0)\) is the root-zone ID. For example, in Fig. 1, the root zone has depth zero, the eight zones immediately surrounding the root zone have depth one, and the outer seven zones have depth two.

In EGMP, the zone-structure is **virtual** and calculated based on a reference point. Therefore, the construction of zone structure does not depend on the shape of the network region, and it is very simple to locate and maintain a zone. The zone is used in EGMP to provide location reference and support lower-level group membership management. A multicast group can cross multiple zones. With the introduction of virtual zone, EGMP does not need to track individual node movement but only needs to track the membership change of zones, which significantly reduces the management overhead and increases the robustness of the proposed multicast protocol. We choose to design the zone without considering node density so it can provide more reliable location reference and membership management in a network with constant topology changes.

### 3.2 Neighbor Table Generation and Zone Leader Election

For efficient management of states in a zone, a leader is elected with minimum overhead. As a node employs periodic BEACON broadcast to distribute its position in the underneath geographic unicast routing [13], to facilitate leader election and reduce overhead, EGMP simply inserts in the BEACON message a flag indicating whether the sender is a zone leader. With zone size \( r \leq r_t/\sqrt{2} \), a broadcast message will be received by all the nodes in the zone. To reduce the beaconing overhead, instead of using fixed-interval beaconing, the beaconing interval for the underneath unicast protocol will be adaptive. A non-leader node will send a beacon every period of \( Intval_{max} \) or when it moves to a new zone. A zone leader has to send out a beacon every period of \( Intval_{min} \) to announce its leadership role.

A node constructs its neighbor table without extra signaling. When receiving a beacon from a neighbor, a node records the node ID, position and flag contained in the message in its neighbor table. Table 1 shows the neighbor table of node 18 in Fig. 1. The zone ID of the sending node can be calculated from its position, as discussed earlier. To avoid routing failure due to
outdated topology information, an entry will be removed if not refreshed within a period $\text{Timeout}_{NT}$ or the corresponding neighbor is detected unreachable by the MAC layer protocol.

TABLE 1: The neighbor table of node 18 in Fig. 1.

<table>
<thead>
<tr>
<th>nodeID</th>
<th>position</th>
<th>flag</th>
<th>zone ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>$(x_{16}, y_{16})$</td>
<td>1</td>
<td>$(1, 1)$</td>
</tr>
<tr>
<td>1</td>
<td>$(x_{1}, y_{1})$</td>
<td>0</td>
<td>$(1, 1)$</td>
</tr>
<tr>
<td>7</td>
<td>$(x_{7}, y_{7})$</td>
<td>1</td>
<td>$(0, 1)$</td>
</tr>
<tr>
<td>13</td>
<td>$(x_{13}, y_{13})$</td>
<td>1</td>
<td>$(1, 2)$</td>
</tr>
</tbody>
</table>

A zone leader is elected through the cooperation of nodes and maintained consistently in a zone. When a node appears in the network, it sends out a beacon announcing its existence. Then it waits for an $\text{Interval}_{max}$ period for the beacons from other nodes. Every $\text{Interval}_{min}$ a node will check its neighbor table and determine its zone leader under different cases: 1) The neighbor table contains no other nodes in the same zone, it will announce itself as the leader. 2) The flags of all the nodes in the same zone are unset, which means that no node in the zone has announced the leadership role. If the node is closer to the zone center than other nodes, it will announce its leadership role through a beacon message with the leader flag set. 3) More than one node in the same zone have their leader flags set, the one with the highest node ID is elected. 4) Only one of the nodes in the zone has its flag set, then the node with the flag set is the leader.

3.3 Zone-supported Geographic Forwarding

With a zone structure, the communication process includes an intra-zone transmission and an inter-zone transmission. In our zone-structure, as nodes from the same zone are within each other’s transmission range and are aware of each other’s location, only one transmission is required for intra-zone communications. Transmissions between nodes in different zones may be needed for the network-tier forwarding of control messages and data packets. As the source and the destination may be multiple hops away, to ensure reliable transmissions, geographic unicasting is used with the packet forwarding guided by the destination position. However, in normal geographic unicast routing, location service is required for the source to obtain the destination position. In EGMP, to avoid the overhead in tracking the exact locations of a potentially large number of group members, location service is integrated with zone-based membership management without the need of an external location server. At the network tier, only the ID of the destination zone is needed. A packet is forwarded towards the center of the destination zone first. After arriving at the destination zone, the packet will be forwarded to a specific receiving node or broadcast depending on the message type. Generally, the messages related to multicast group membership management and multicast data will be forwarded to the zone leader to process.

In the above design, for scalability and reliability, the center of the destination zone is used as the landmark for sending a packet to the group members in the zone although there may be no node located at the center position. This, however, may result in the failure of geographic forwarding. For example, in Fig. 1, node 7 is the only node in zone $(0, 1)$, while node 18 in zone $(1, 1)$ is closest to the center of zone $(0, 1)$. When node 16 sends a packet to zone $(0, 1)$ with its center as the destination, the underlying geographic unicast protocol (for example, GPSR [13] described in Section 1) will forward the packet to node 18 greedily as it is closer to the destination. As node 18 cannot find a neighbor closer to the center of zone $(0, 1)$ than itself, the perimeter mode [13] may be used to continue the forwarding. This still cannot guarantee the packet to arrive at node 7, as the destination is a virtual reference point. Such a problem is neglected by the previous geographic protocols that use a region as destination (e.g., [20]).

To avoid this problem, we introduce a zone forwarding mode in EGMP when the underlying geographic forwarding fails. Only when the zone mode also fails, the packet will be dropped. In zone mode, a sender node searches for the next hop to the destination based on its neighbor table, which can more accurately track the local network topology. The node selects as its next hop the neighboring node whose zone is the closest to the destination zone and closer to the destination zone than its own zone. If multiple candidates are available, the neighbor closest to the destination is selected as the next hop. To compare the distances of different zones to the destination zone, the node can calculate the distance value $\text{dis}_{(a,b)}$ of a zone $(a,b)$ to the destination zone $(a_{dst}, b_{dst})$ as:

$$\text{dis}_{(a,b)} = (a-a_{dst})^2 + (b-b_{dst})^2.$$  \hspace{1cm} (1)

A zone with a smaller $\text{dis}$ value is closer to the destination zone. In the above example, if the underlying geographic unicast forwarding fails at node 18, it will try to continue the forwarding using zone mode. It checks its neighbor table (Table 1). Since the $\text{dis}$ value of zone $(0, 1)$ has zero value to the destination zone $(0, 1)$, node 18 selects its neighbor node 7 in zone $(0, 1)$ as the next hop and forwards the packet to node 7. To avoid possible routing loop, an intermediate node only forwards a packet that is received for the first time.

To validate our zone-supported geographic forwarding strategy, we compare our scheme (represented as EGMP_forward) with two other schemes through simulations: SPBM [20] (represented as SPBM_forward), in which a sender uses the point in the square which is closest to itself as the destination position; a unicast strategy which also uses the zone center to estimate the destination position (represented as EGMP_non_zonemode) but does not use the zone forwarding mode. The settings and protocol parameters are the same as in Section 5. We simulated 30 CBR traffic flows with each
flow sent at 8 Kbps between a randomly chosen source and a non-empty zone. A packet is considered to be successfully delivered if it is received by any node in the destination zone. In Fig. 2, when the destination is a zone, the zone center is a better estimation of the destination position than the closest point in the destined zone. As the estimated closest point in the destined zone could be very close to the zone border, compared to the zone center, it is more likely for an out-of-zone node to be closer to the estimated point and become the forwarder than an intra-zone node. Hence, the forwarder may have a higher chance of dropping the packet when not able to find a next-hop node closer to the destination for forwarding the packet. The simulation results confirm that using zone forwarding mode can help reduce the number of undelivered packets.

3.4 Multicast Tree Construction

In this subsection, we present the multicast tree creation and maintenance schemes. In EGMP, instead of connecting each group member directly to the tree, the tree is formed in the granularity of zone with the guidance of location information, which significantly reduces the tree management overhead. With a destination location, a control message can be transmitted immediately without incurring a high overhead and delay to find the path first, which enables quick group joining and leaving. In the following description, except when explicitly indicated, we use G, S and M respectively to represent a multicast group, a source of G and a member of G.

3.4.1 Multicast session initiation and termination

When a multicast session G is initiated, the first source node S (or a separate group initiator) announces the existence of G by flooding a message NEW_SESSION(G, zoneID_S) into the whole network. The message carries G and the ID of the zone where S is located, which is used as the initial root-zone ID of group G. When a node M receives this message and is interested in G, it will join G using the process described in the next subsection. A multicast group member will keep a membership table with an entry (G, root_zoneID, isAcked), where G is a group of which the node is a member, root_zoneID is the root-zone ID and isAcked is a flag indicating whether the node is on the corresponding multicast tree. A zone leader (zLdr) maintains a multicast table. When a zLdr receives the NEW_SESSION message, it will record the group ID and the root-zone ID in its multicast table. Table 2 is an example of one entry in the multicast table of node 16 in Fig. 1. The table contains the group ID, root-zone ID, upstream zone ID, downstream zone list and downstream node list. To end a session G, S floods a message END_SESSION(G). When receiving this message, the nodes will remove all the information about G from their membership tables and multicast tables.

**TABLE 2:** The entry of group G in multicast table of node 16

<table>
<thead>
<tr>
<th>group ID</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>root-zone ID</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>upstream zone ID</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>downstream zone list</td>
<td>(0, 1), (0, 0)</td>
</tr>
<tr>
<td>downstream node list</td>
<td>1</td>
</tr>
</tbody>
</table>

3.4.2 Multicast group join

When a node M wants to join the multicast group G, if it is not a leader node, it sends a JOIN_REQ(M, Pos_M, G, {M_old}) message to its zLdr, carrying its address, position, and group to join. The address of the old group leader M_old is an option used when there is a leader handoff and a new leader sends an updated JOIN_REQ message to its upstream zone. If M did not receive the NEW_SESSION message or it just joined the network, it can search for the available groups by querying its neighbors. If a zLdr receives a JOIN_REQ message or wants to join G itself, it begins the leader joining procedure as shown in Fig. 3. If the JOIN_REQ message is received from a member M of the same zone, the zLdr adds M to the downstream node list of its multicast table. If the message is from another zone, it will compare the depth of the requesting zone and that of its own zone. If its zone depth is smaller, i.e., its zone is closer to the root zone than the requesting zone, it will add the requesting zone to its downstream zone list; otherwise, it simply continues forwarding the JOIN_REQ message towards the root zone.

If new nodes or zones are added to the downstream list, the leader will check the root-zone ID and the upstream zone ID. If it does not know the root zone, it starts an expanded ring search. As the zone leaders in the network cache the root-zone ID, a result can be quickly obtained. With the knowledge of the root zone, if its upstream zone ID is unset, the leader will represent its zone to send a JOIN_REQ message towards the root zone; otherwise, the leader will send back a JOIN_REPLY message to the source of the JOIN_REQ message (which may be multiple hops away and the geographic unicasting described in Section 3.3 is used for this transmission). When the source

---

**Procedure** LeaderJoin(me, pkt)

- me: the leader itself
- pkt: the JOIN_REQ message the leader received

**BEGIN**

if (pkt.srcZone == me.zoneID) then
  /* the join request is from a node in the local zone */
  if (isAcked) then
    pkt.isAcked = true;
    SendReply(pkt);
  else
    SendJoinRequest(pkt);
  end if
else
  /* the join request is from another zone */
  if (depth_me < depth_pkt) then
    SendJoinRequest(pkt);
  else
    SendReply(pkt);
  end if
end if

**Fig. 3:** The pseudocode of the leader joining procedure.
of the JOIN_REQ message receives the JOIN_REPLY, if it is a node, it sets the isAcked flag in its membership table and the joining procedure is completed. If the leader of a requesting zone receives the JOIN_REPLY message, it will set its upstream zone ID as the ID of the zone where the JOIN_REPLY message is sent, and then send JOIN_REPLY messages to unacknowledged downstream nodes and zones.

An example is given in Fig. 1, in which the root zone of G is (2, 2), and the double circled nodes are zone leaders. Suppose currently zone (0, 0) and (1, 1) are not on the multicast tree, and their leader nodes 15 and 16 already know the root zone ID from the NEW_SESSION message. Now node 15 plans to join G with the leader joining procedure. As it finds its upstream zone ID is unset, node 15 sends a JOIN_REQ message towards root zone (2, 2). The message reaches zone (1, 1) and is intercepted by leader node 16, which then starts its leader joining procedure. Node 16 compares the depths of zone (0, 0) and its own zone. Since $depth(0, 0) = 2$ and $depth(1, 1) = 1$, $depth(0, 0) > depth(1, 1)$, node 16 adds zone ID (0, 0) to its downstream zone list. As node 16 finds its upstream zone ID is unset, it sends a JOIN_REQ message towards root zone. This message is received by the leader of the root-zone, node 3, and triggers the joining procedure of node 3. Node 3 adds the zone ID (1, 1) to its downstream zone list after comparing the depth. As the root zone does not have an upstream zone, node 3 sends back a JOIN_REPLY message to the zone (1, 1). On receiving this message, node 16 sets the upstream zone ID as (2, 2) and sends a JOIN_REPLY message to its unacknowledged downstream zone (0, 0). Node 15 sets its upstream zone as (1, 1) on receiving the JOIN_REPLY message and the joining process is completed, with two multicast branches built between zones (2, 2) and (1, 1), and between zones (1, 1) and (0, 0) respectively.

Through the joining process, the group membership management is implemented in a distributed manner. An upstream zone only needs to manage its downstream zones, and the group membership of a local zone is only managed by its leader. The zone depth is used to guide efficient tree construction and packet forwarding.

### 3.4.3 Multicast group leave

When a member M wants to leave G, it sends a LEAVE(M, G) message to its zone leader. On receiving a LEAVE message, the leader removes the source of the LEAVE message from its downstream node list or zone list depending on whether the message is sent from an intra-zone node or a downstream zone. Besides removing a branch through explicit LEAVE, a leader will remove a node from its downstream list if it does not receive the beacon from the node exceeding $2 \times Interval_{max}$. If its downstream zone list and node list of G are both empty and it is not a member of G either, the leader sends a LEAVE(zoneID, G) message to its upstream zone. Through the leave process, the unused branches are removed from the multicast tree.

### 3.5 Multicast Packet Delivery

In this subsection, we explain how the multicast packets are forwarded to the members.

#### 3.5.1 Packet sending from the source

After the multicast tree is constructed, all the sources of the group could send packets to the tree and the packets will be forwarded along the tree. In most tree-based multicast protocols, a data source needs to send the packets initially to the root of the tree. If this scheme is used and node 5 in Fig. 1 is a source, node 5 needs to unicast the packets initially to root zone (2, 2). The sending of packets to the root would introduce extra delay especially when a source is far away from the root. Instead, EGMP assumes a bi-directional-tree-based forwarding strategy [23], with which the multicast packets can flow not only from an upstream node/zone down to its downstream nodes/zones, but also from a downstream node/zone up to its upstream node/zone.

A source node is also a member of the multicast group and will join the multicast tree. When a source S has data to send and it is not a leader, it checks the isAcked flag in its membership table to find out if it is on the tree. If it is, i.e., its zone has joined the multicast tree, it sends the multicast packets to its leader. When the leader of an on-tree zone receives multicast packets, it forwards the packets to its upstream zone and all its downstream nodes and zones except the incoming one. For example, in Fig. 1, source node 1 sends the packets to its leader node 16, which will send the packets to its upstream zone (2, 2) and its downstream zones (0, 1) and (0, 0), but not to the downstream node 1 which is the incoming node. When the packets are received by leader node 3 of the root zone, it continues forwarding the packets to its downstream zones (1, 3), (3, 3), (2, 1) except the incoming zone (1, 1). The arrows in the figure indicate the directions of the packet flows.

When a source node S is not on the multicast tree, for example, when it moves to a new zone, the isAcked flag will remain unset until it finishes the rejoining to G through the leader of the new zone. To reduce the impact of the joining delay, S will send packets directly to the root zone until it finishes the joining process.

#### 3.5.2 Multicast data forwarding

In our protocol, only zLdrs maintain the multicast table, and the member zones normally cannot be reached within one hop from the source. When a node N has a multicast packet to forward to a list of destinations ($D_1, D_2, D_3, \ldots$), it decides the next hop node towards each destination (for a zone, its center is used) using the geographic forwarding strategy described in Section 3.3. After deciding the next hop nodes, N inserts the list of next hop nodes and the destinations associated with each next hop node in the packet header. An example list is $(N_1 : D_1, D_3; N_2 : D_2; \ldots)$, where $N_1$ is the next hop node for the destinations $D_1$ and $D_3$, and $N_2$ is the next hop node for $D_2$. Then N broadcasts the packet promiscuously (for reliability and efficiency). Upon receiving the packet, a neighbor node will keep the packet if it is one of the next hop nodes or destinations, and drop the packet otherwise. When the node is associated with some downstream destinations, it will continue forwarding packets similarly as done by node N.
For example, in Fig. 1, after node 3 receives the multicast packet from zone (1, 1), it will forward the packet to downstream zones (2, 1), (1, 3) and (3, 3). It determines the next hop node for each destination and inserts the list (12: (1,3),(3,3); 14: (2,1)) in the packet header. After broadcasting the packet promiscuously, its one-hop neighbors node 12, node 14 and node 8 will receive the packet. Node 8 will drop this packet, while node 12 and node 14 will continue the forwarding. Node 12 replaces the list carried in the packet header with (17: (1,3); 2: (3,3)) and broadcasts this packet. Node 14 finds group information from its multicast table, and broadcast the packet with a header (9: (1,0); 5:(3,0)).

3.6 Multicast Route Maintenance and Optimization

In a dynamic network, it is critical to maintain the connection of the multicast tree, and adjust the tree structure upon the topology changes to optimize the multicast routing. In the zone structure, due to the movement of nodes between different zones, some zones may become empty. It is critical to handle the empty zone problem in a zone-based protocol. Compared to managing the connections of individual nodes, however, there is a much lower rate of zone membership change and hence a much lower overhead in maintaining the zone-based tree. As the tree construction is guided by location information, a disconnected zone can quickly re-establish its connection to the tree. In addition, a zone may be partitioned into multiple clusters due to fading and signal blocking. In this subsection, we discuss our maintenance schemes.

3.6.1 Moving between different zones

When a member node moves to a new zone, it must rejoin the multicast tree through the new leader. When a leader is moving away from its current zone, it must handover its multicast table to the new leader in the zone, so that all the downstream zones and nodes will remain connected to the multicast tree.

Whenever a node M moves into a new zone, it will rejoin a multicast group G by sending a JOIN_REQ message to its new leader. During this joining process, to reduce the packet loss, whenever the node broadcasts a BEACON message to update its information to the nodes in the new zone, it also unicasts a copy of the message to the leader of its previous zone to update its position. Since it has not sent the LEAVE message to the old leader, the old leader will forward the multicast packets to M. This forwarding process helps reduce the packet loss and facilitates seamless packet transmissions during zone crossing. When the rejoining process finishes, M will send a LEAVE message to its old leader.

To handle leader mobility problem, if a leader finds its distance to the zone border is less than a threshold or it is already in a new zone, it assumes it is moving away from the zone where it was the leader, and it starts the handover process. To look for the new leader, it compares the positions of the nodes in the zone it is leaving from and selects the one closest to the zone center as the new leader. It then sends its multicast table to the new leader, which will announce its leadership role immediately through a BEACON message. It will also send a JOIN_REQ message to its upstream zone.

During the transition, the old leader may still receive multicast packets. It will forward all these packets to the new leader when the handover process is completed. If there is no other node in the zone and the zone will become empty, it will use the method introduced in the next subsection to deliver its multicast table. In the case that the leader dies suddenly before handing over its multicast table, the downstream zones and nodes will reconnect to the multicast tree through the maintenance process described in Section 3.6.4.

3.6.2 Dealing with empty zones

A zone may become empty when all the nodes move away. The probability that a zone is empty is approximately \( P = e^{-\rho r^2} \) when the node density is \( \rho \) and the zone size is \( r \). Assume \( r = 150m \), which is the zone size that allows all the nodes in the same zone to be within the transmission range, the probability of the zone being empty is: \( P = 0.207 \) if \( d = 70\text{nodes}/\text{km}^2 \), and \( P = 0.509 \) if \( d = 30\text{nodes}/\text{km}^2 \).

We can see that the probability of a zone becoming empty is not negligible and it is critical to address the empty zone problem.

In EGMP, if a tree zone becomes empty, the multicast tree will be adjusted correspondingly to keep the multicast tree connected. Because of the importance of the root zone, we will treat it differently. When a leader is moving away from a non-root tree-zone and the zone is becoming empty, it will send its multicast table to its upstream zone. The upstream zone leader will then take over all its downstream zones, and delete this requesting zone from its downstream zone list. The new upstream zone needs to send JOIN_REPLY messages to all the newly added downstream zones to notify them the change. When receiving the JOIN_REPLY messages, these downstream zones will change their upstream zone ID accordingly.

If the to-be empty zone is the root zone, since the root zone has no upstream zone, the leader will check its neighboring zones and choose the one closest to the root zone as the new root zone. The leader then forwards its multicast table to the new root zone, and floods a NEW_ROOT message to announce the change.

3.6.3 Handling Multiple Clusters per Zone

When there is severe shadowing/fading or a hill/building that prevents the radio communication between nodes in a zone, the nodes in the same zone may form multiple clusters as shown in Fig. 4, where the two clusters are not connected in the zone although they are connected through some nodes outside the zone. In this case, two nodes in different clusters can communicate with each other by using unicasting because...
they are connected on the network topology graph, but an intra-zone flooding message initiated in one cluster may not reach other clusters. This problem is also a key problem for zone-based protocols.

EGMP handles the zone partitioning problem as follows. If there are multiple clusters in a zone, because these clusters are not aware of the existence of each other, each cluster will elect its leader. When an upstream zone leader receives JOIN_REQ messages from multiple leaders of the same zone and the new message is not sent as a result of leader handover (in which case the old leader’s address needs to be carried), it detects that the downstream zone has partitioned into multiple clusters. It identifies a cluster by its zID and the leader’s address. When sending a packet to the cluster, it uses the leader’s position instead of the zone center (in which case the zone ID is carried as the destination) as the transmission reference. Even though the leader may move, its position carried in JOIN_REQ message can still be used as a reference to forward packets to its cluster. When receiving a packet with the position of the leader as the reference, a cluster leader can learn that multiple clusters exist within its zone. In case that not all the clusters of a partitioned zone send JOIN_REQ messages, the upstream zone leader may not be aware of the partitioning of the downstream zone. When a cluster leader receives a packet destined to its zone but does not match its status, it will send an update message to its upstream zone. For example, when a cluster leader receives a JOIN_REPLY message or a multicast packet but did not send JOIN_REQ message, it will send a LEAVE message to the upstream zone. When receiving messages from multiple leaders of the same zone, the upstream leader can detect zone partitioning. It will resend the previous message to the target cluster with the position of the zone leader as the destination.

When the leader of a cluster changes, if the cluster is on-tree, the new leader sends a JOIN_REQ message to its upstream zone immediately which also carries the old leader’s address. With multiple clusters in its upstream zone, the JOIN_REQ message from a zone leader will generally be intercepted by one of the clusters, which will be responsible for forwarding the packets to the zone. Some clusters may merge later into a larger cluster, and through the leader election procedure, only one of the leaders will win as the new cluster’s leader. The new leader will send a JOIN_REQ message to the upstream zone to refresh the cluster’s information. The leaders of the other merged-in clusters will also send LEAVE messages to the upstream zone, which will remove their information from the multicast table.

3.6.4 Tree branch maintenance

To detect the disconnection of tree-branches in time, if there are no multicast packets or messages to deliver for a period of Intval_active, the leader of a tree-zone will send an ACTIVE message to its downstream nodes and zones to announce the activity of the multicast branches. The message is sent through multicast to multiple downstream entities. When a member node or a tree-zone fails to receive any packets or messages from its leader or upstream zone up to a period of 2 * Intval_active, it assumes that it loses the connection to the multicast tree and restarts a joining process.

3.6.5 Route Optimization

Sometimes a zone leader may receive duplicate multicast packets from different upstream zones. For example, as described in the above subsection, when failing to receive any data packets or ACTIVE messages from the upstream zone for a period of time, a tree-zone will start a rejoining process. However, it is possible that the packet and message were lost due to collisions, so the old upstream zone is still active after the rejoining process, and duplicate packets will be forwarded by two upstream zones to the tree-zone. In this case, the one closer to the root zone will be kept as the upstream zone. If the two upstream zones have the same distances to the root zone, one of them is randomly selected.

4 Cost Analysis

In this section, we will quantitatively analyze the per node cost of the protocol, which is defined as the average number of control messages transmitted by each node per second. The notations to be used in this section are listed in Table 3. The cost of the overall protocol consists of the following three components: zone building and geographic routing, tree construction, and tree maintenance.

<table>
<thead>
<tr>
<th>Table 3: Notations used in the cost analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
</tr>
<tr>
<td>( r )</td>
</tr>
<tr>
<td>( R )</td>
</tr>
<tr>
<td>( v )</td>
</tr>
<tr>
<td>( T )</td>
</tr>
<tr>
<td>( M_n )</td>
</tr>
<tr>
<td>( M_z )</td>
</tr>
</tbody>
</table>

4.1 Cost for zone building and geographic routing

The zone is virtual and determined by each node based on its position and the reference origin, without the need of extra signaling messages. The leader information is distributed with a flag inserted in the beacon messages of the underlying geographic unicast routing protocol. Therefore, the per node cost of the zone building and geographic routing is impacted by the beaconing frequency \( 1/Intval_{min} \) introduced in Section 3.2, and the cost is as follows:

\[
Cost_{unicast} \leq \frac{1}{Intval_{min}} = O(1).
\]

4.2 Cost for tree construction

The tree construction process is associated with the multicast session initiation and termination, and the member joining and leaving the multicast tree.

**Lemma 1:** The per node cost of multicast tree construction is \( O(1) \) with respect to the network size and the group size.

**Proof:** Multicast session initiation and termination include a flooding of a NEW_SESSION message and a flooding of END_SESSION message, so the cost for multicast session initiation and termination is:

\[
Cost_{init \_end} = \frac{1}{NT}(2 \times N) = O(1).
\]
To analyze the cost for the joining and leaving process, we consider the worst case that all the zones need to join the multicast tree and become tree zone and all the members to join the tree are not zone leaders. In this case, group members need to first send a JOIN_REQ message to their leaders in local zones, and each leader needs to send a JOIN_REQ message towards the root zone to join the multicast tree. The JOIN_REQ message from a zone will be intercepted by an upstream zone leader. The distance between an upstream and a downstream zone leader is shorter than $2\sqrt{2r} + r_t$, where $r_t$ is the transmission range. According to [22], the average number of hops of the greedy forwarding path between the source and the destination is $\frac{d}{2}$, where $d$ is the average distance between the source and the destination and $\bar{d}$ is the average forwarding progress made towards the destination in the course of one transmission. Parameter $\bar{d}$ depends on $r_t$ and the average number of nodes within the transmission range, and is constant with respect to the group size and the network size. Hence,

$$\text{Cost}_{join} \leq \frac{1}{NT} (M_a + M_z 2\sqrt{2r} + r_t \bar{d}),$$

and as $M_a \leq N$ and $M_z \leq N$, $\text{Cost}_{join} \leq O(1)$. Since the JOIN_REPLY message follows the reverse direction of JOIN_REQ message, and the leaving process is similar to the joining process, $\text{Cost}_{reply} = \text{Cost}_{join} \leq O(1)$ and $\text{Cost}_{leave} = \text{Cost}_{join} \leq O(1)$. Therefore, the cost for multicast tree construction is:

$$\text{Cost}_{tree} = \text{Cost}_{init\_end} + \text{Cost}_{join} + \text{Cost}_{reply} + \text{Cost}_{leave} = O(1).$$

This indicates that the per-node control overhead involved in multicast tree construction remains relatively constant with respect to network size and group size.

### 4.3 Cost for tree maintenance

The cost involved in multicast tree maintenance include the handling of zone crossing of multicast members, the tree reconstruction when there is an empty zone, and the tree branch maintenance.

![Fig. 5: The moving distance of a mobile node in a zone.](image)

**Lemma 2:** Assume that a node keeps the same moving direction in a zone. The average moving distance of the mobile nodes in a zone is $\frac{\pi r}{4}$.

**Proof:** The moving distance $d$ of a node in a zone is the length of its moving trail in the zone square. For example, in Fig. 5, line $a$ is such a moving trail. Suppose the angle formed by the moving trail and the bottom side of the zone square is $\theta$. Due to the symmetry of the square, we only need to consider the case when $\theta \in [0, \pi]$. As illustrated in Fig. 5, all the possible moving trails with angle $\theta$ are located between two parallel lines $b$ and $c$, where $b$ and $c$ are tangent to the zone with angle $\theta$. Line $l$ is perpendicular to $b$ and $c$ and intersects $b$ at point $A$, $a$ intersects $l$ at $B$. Suppose the distance between $A$ and $B$ is $z$, the length of a moving trail is decided by its angle $\theta$ and the distance $z$. Therefore, we can obtain the average distance of a node moving in a zone as

$$d = \int_0^\pi \int_0^\frac{\pi}{2} \left(2f_0^\frac{z}{\sin \theta} \frac{\sin \theta}{\sin \theta} dz + \int_0^\frac{\pi}{2} (\cos \theta - \sin \theta) \frac{r}{\cos \theta} dz \right) d\theta$$

$$= \frac{\pi r}{4}.$$  

(3)

**Lemma 3:** The per node cost of multicast tree maintenance is $O(1)$ with respect to the network size and the group size.

**Proof:** The cost for the tree maintenance is composed of the cost for handling zone crossing of member nodes, the cost in adapting the tree structure in the presence of empty zones, and the cost in maintaining tree branches, as presented in Section 3.6.1, 3.6.2 and 3.6.4. We first analyze these costs separately.

The first component includes the cost of handling the moving between different zones due to non-leader members and due to leaders. For a non-leader member, it must send a JOIN_REQ message to the new zone leader, which will send a JOIN_REPLY, and a LEAVE message to the old leader. Its distances to the new and the old zone leaders are shorter than $\sqrt{2r}$ and $2\sqrt{2r}$ respectively. By Lemma 2, the average frequency of a node moving between different zones is $\frac{\pi r}{4}$. A leader will send its multicast table to the new selected leader when moving out of the zone. Hence, the cost due to member nodes moving across zones is:

$$\text{Cost}_{moving} = \text{Cost}_{non\_zLdr} + \text{Cost}_{zLdr} \leq \frac{M_n 4v(1 + 1 + 2) \sqrt{2r}}{\pi r} + \frac{M_z 4v \sqrt{2r}}{\pi r \bar{d}} = O(1).$$  

(4)

Due to node movement, a zone may become empty. The empty zone may be a tree zone or the root zone. When the upstream zone of a tree zone is to be empty, the moving away leader will hand the tree zone over to its own upstream zone, which will send a JOIN_REPLY message to the new zone leader, and a LEAVE message to the old leader. Its distances to the new and the old zone leaders are shorter than $\sqrt{2r}$ and $2\sqrt{2r}$ respectively. By Lemma 2, the average frequency of a node moving between different zones is $\frac{\pi r}{4}$. A leader will send its multicast table to the new selected leader when moving out of the zone. Hence, the cost due to member nodes moving across zones is:

$$\text{Cost}_{empty\_zone} = \text{Cost}_{tree\_zone} + \text{Cost}_{root\_zone} \leq \frac{4vT}{\pi r} (\text{Cost}_{join} + \text{Cost}_{reply}) + \frac{N 4v}{\pi r} = O(1).$$  

(5)

The cost for tree branch maintenance should be also less than the cost of joining process with frequency $\frac{T}{\text{Intval}_{active}}$.

$$\text{Cost}_{active} \leq \frac{T}{\text{Intval}_{active}} (\text{Cost}_{join} + \text{Cost}_{reply}) = O(1).$$  

(6)

Therefore, the total per node cost for tree maintenance is:

$$\text{Cost}_{maintain} = \text{Cost}_{moving} + \text{Cost}_{empty\_zone} + \text{Cost}_{active} \leq O(1).$$  

(7)

### 4.4 Cost for the Protocol

We summarize the per node cost of the protocol and validate our quantitative analysis through simulations.
4.4.1 Quantitative analysis on the per node cost

Theorem 1: The EGMP control overhead as the average number of control message transmissions per node every second has a complexity of $O(1)$ with respect to the network size and the group size.

Proof: The overhead of the protocol is generated from the tree construction and maintenance and the periodic beaconing in the underlying geographic unicast routing protocol. By Lemma 1, Lemma 3 and Eq. 2, the cost of the protocol, i.e., the number of transmissions of control messages per node every second with respect to the network size and the group size is:

$$\text{Cost}_{\text{protocol}} = \text{Cost}_{\text{tree}} + \text{Cost}_{\text{maintain}} + \text{Cost}_{\text{unicast}} = O(1).$$

4.4.2 Validation of the cost analysis by simulation

We validate our quantitative analysis on the protocol cost through simulations. The simulation settings and protocol parameters were set as those in Section 5. We studied the protocol cost, i.e., the average number of transmissions of control messages by each node per second, with network size varied from $1500m \times 1500m$ with 156 nodes to $3900m \times 3900m$ with 1056 nodes and the group size varied from 10 members to 200 members.

Fig. 6 (a) and (b) validate our quantitative analysis on the protocol cost. The protocol cost keeps almost constant between 0.3 and 0.4 with different network sizes and group sizes.

The above analysis results indicate that when the network size and the group size increase, the control overhead placed on each node per second by the protocol will remain relatively constant. Next, we will further demonstrate the scalability and efficiency of the protocol by simulation studies.

5 Performance Evaluation

We implemented the EGMP protocol using Global Mobile Simulation (GloMoSim) [24] library, and compare it with ODMRP [8] which is widely used and considered to be robust over a dynamic network, and the geographic multicast protocol SPBM [20] which is designed to improve the scalability of position-based multicast. The SPBM is a quad-tree-based protocol as introduced in Section 2. ODMRP is a mesh-based on-demand non-geographic multicast protocol, and takes a soft-state approach to maintain multicast group members. A multicast source broadcasts a Join-Query messages to the entire network periodically. An intermediate node stores the source ID and the sequence number, and updates its routing table with the node ID (i.e., backward learning) from which the message was received for the reverse path back to the source. A receiver creates and broadcasts a Join Reply to its neighbors, with the next hop node ID field filled by extracting information from its routing table. The neighbor node whose ID matches the next-hop node ID of the message realizes that it is on the path to the source and is part of the forwarding group. It then broadcasts its own Join Table built upon matched entries. This whole process constructs (or updates) the routes from sources to receivers and builds a mesh of nodes, the forwarding group.

The simulations were run with 400 nodes randomly distributed in an area of $2400m \times 2400m$. The nodes moved following the modified random waypoint mobility model [25]. The moving speed of nodes are uniformly set between the minimum and maximum speed values which are set as as 1 m/s (with pause time as 100 seconds) and 20 m/s respectively except when studying the effect of mobility. We set the MAC protocol and radio parameters according to the Lucent WaveLAN$^TM$ card, which operates at a data rate 11Mbps and radio frequency 2.4GHz with a nominal transmission range 250m. IEEE 802.11b was used as the MAC layer protocol. Each simulation lasted 500 simulation seconds. Each source sends CBR data packets at 8 Kbps with packet length 512 bytes. The CBR flows start at around 30 second so that the group membership management has time to initialize and stop at 480 second. By default, there is one source, and one multicast group with 100 members. A simulation result was gained by averaging over six runs with different seeds.

5.1 Parameters and Metrics

Table 4 lists the simulation parameters of EGMP with beacon interval set as [13]. The simulations for ODMRP are based on the codes carried with the simulator, with the parameters set as in [9]. We fixed several bugs in the GloMoSim codes which would prevent a forwarding group node from sending JOIN TABLES. The improvement doubles the delivery ratio and reduces the control overhead of ODMRP. Additionally, we implemented SPBM in GloMoSim according to [20] and the ns2 codes provided by the authors with the same parameter settings except that the square size was set to 150 m so that the nodes in a square are within each other’s transmission range. The number of levels of the quad-tree changes accordingly with the square size and the network size we used. For packet forwarding in SPBM, different from the scheme described in [20], we used the square center as the destination position, which improves the delivery ratio of SPBM.

We focus on the studies of the scalability and efficiency of the protocol under the dynamic environment and the following metrics were used for the multicast performance evaluation:

1) Packet delivery ratio: The ratio of the number of packets received and the number of packets expected to receive. Thus for multicast packet delivery, the ratio is equal to the total number of received packets over the number of originated packets times the group size.

2) Normalized control overhead: The total number of control message transmissions divided by the total number of received data packets. Each forwarding of the control message was counted as one transmission. Different
We first compare the performance of ODMRP, SPBM and EGMP with the variation of moving speed and node density, we then study the scalability of the three protocols with the change of group size and network size.

5.2 Simulation Results

We first compare the performance of ODMRP, SPBM and EGMP with the variation of moving speed and node density, we then study the scalability of the three protocols with the change of group size and network size.

5.2.1 Effect of moving speed

It is critical and challenging for a multicast routing protocol to maintain a good performance in the presence of node mobility in an ad hoc network. We evaluate the protocol performance by varying maximum moving speed from 5m/s to 40m/s.

From Fig. 7(a), the delivery ratios of all the protocols reduce as mobility increases, while the delivery ratio of ODMRP drops much faster. At higher moving speed, as it is more difficult to track the group membership and maintain the routing path, more packets are lost. Although the mesh structure used in ODMRP is more robust than the general tree structure, the mesh structure built through back learning is likely to become invalid as nodes move. In all the mobility cases, the geographic multicast protocols EGMP and SPBM have much higher delivery ratios. This is as expected. As geographic forwarding only requires local topology information and is more robust to the network topology change, both protocols achieve more reliable packet delivery by taking advantage of geographic unicast for data packet transmissions. The use of virtual zone in EGMP further increases its membership management reliability and stability.

The multicast and the total control overhead of EGMP are seen to be lower than those of ODMRP and SPBM at different moving speeds (Fig. 7(b)). SPBM is seen to have more than six times overhead of the other two protocols due to its use of periodic local and network-wide flooding in its membership management. In EGMP, the membership management is based on the efficient local zone structure and network-wide tree structure without involving periodic network-wide flooding, which greatly improves the efficiency and scalability of the protocol. As expected, the control overheads of all the protocols increase at higher mobility. In EGMP, when nodes move faster, there are more frequent zone leader changes and zone crossings, which triggers more rejoining processes. The increase of the normalized control overhead of all the protocols is also due to the reduced packet receiving under high mobility. The overhead due to the beacon messages in underlying unicast is seen to be smaller in SPBM, while the beacons contribute to most of the control overhead of EGMP. The reason is as follows, GPSR [13], the underlying unicast protocol of SPBM and EGMP, adopts promiscuous beacons to reduce the beaconing overhead, i.e., data packets also serve as beacons by carrying the position information of the forwarders.
The sending of the next beacon message will be delayed once a node forwards a packet carrying its position. To further reduce beaconing overhead, in our simulation, some multicast control messages in SPBM and EGMP also carry the positions of the forwarding nodes. As SPBM has a larger number of multicast control messages and a higher number of data packet transmissions (Fig. 7 (b) and (c)), more beacon messages in its underlying geographic unicast are suppressed.

Both EGMP and SPBM have higher data packet transmission overheads than ODMRP as seen in Fig. 7(c). Instead of forwarding packets through hop-by-hop broadcasting as that in ODMRP, the multicast forwardings in EGMP and SPBM are through hierarchical structures, leading to more data transmissions. However, between the two geographic multicast protocols, EGMP has a much lower packet transmission overhead as a result of the stronger aggregation of packet forwardings through the efficient tree-based transmissions. On the other hand, the packet transmission overhead of SPBM is seen to rise quickly when moving speed increases. At a higher node moving speed, the membership change of a low-layer square in SPBM cannot be distributed quickly to the upper layer, and a lot of packets are forwarded to the squares with outdated membership information, which leads to a higher packet transmission overhead.

In EGMP, when a node wants to join a group, it will start the joining process immediately, and with the efficient tree structure assumed, the nodes can join the multicast structure very fast as shown in Fig. 7(d). SPBM is seen to have the largest joining delay most of the time. As described in Section 2, with the use of periodic multi-level membership update mechanism, it may take a long time for a bottom level square of SPBM to distribute its membership change to the upmost level. In ODMRP, the mesh structure is built on the source’s demand, and a source sends out a JOIN QUERY message periodically to refresh the mesh structure. If the nodes want to join a group, they need to wait until the next mesh refreshing period. The refreshing interval is set as 3 seconds according to [9]. From the figure, the average joining delay of ODMRP drops with the increase of mobility, as the higher moving speed helps a member to connect to the source more quickly in the non-geographic mesh structure.

In summary, both geometric multicast protocols EGMP and SPBM have similar delivery ratios and are more robust to the mobility than ODMRP, and achieve more than 20% higher delivery ratios than ODMRP at the highest mobility tested. EGMP has the minimum control overhead and group joining delay under all the mobility. The control overhead of ODMRP and EGMP are comparable, while the overhead of SPBM is about six times their overhead. Similarly, the joining delay of SPBM is also six times that of EGMP. The joining delay of ODMRP reduces with the increase of mobility, and is still three times that of EGMP at the highest mobility. The increase of mobility also leads to the significant increase of transmissions of SPBM, as the membership change of a low-layer square in SPBM cannot be distributed quickly to the upper layer which results in outdated membership information and higher packet transmission overhead.

Fig. 9: Performance vs. group size.

Fig. 10: Performance vs. network size.

### 5.2.2 Effect of node density

Geographic routing is sensitive to the node density and performs better in a dense network. Node density is also closely related to the performance of zone-based protocols. When the node density is low, there will be more empty zones, which will negatively affect the performance. In EGMP, specific design has been made to minimize the impact of empty zone on the performance of multicasting.

As expected, both EGMP and SPBM have higher delivery ratios at a higher node density (Fig. 8 (a)). The delivery ratios of all three protocols are lower when the network is sparse and more group members are disconnected from the network graph, while the delivery ratios of EGMP and SPBM go up quickly as the network density increases. However, when the node density is higher than 50 nodes/km², the increase of delivery ratio becomes slower, as there are more collisions among nodes and hence more packet loss.

In Fig. 8(b), the control overhead of SPBM rises quickly with the increase of node density as more nodes are involved in the periodic multi-level flooding for the membership
management. When the network is sparse, there is a smaller number of periodic beacons and thus a smaller unicast control overhead in both EGMP and SPBM. When the network is very sparse, EGMP has a slightly higher control overhead than that of ODMRP, as the underlying geographic unicast routing more frequently uses recovery forwarding to recover from the local void when a node cannot find a neighbor closer to the destination. A zone also has a higher probability of being empty, which results in a higher tree maintenance overhead.

In geographic routing, the likelihood of using recovery forwarding drops with the increase of network density, and the data packet transmission overhead in SPBM reduces accordingly (Fig. 8(c)). However, the transmission overheads of both EGMP and ODMRP increase, and the overhead of ODMRP increases faster. In EGMP, the transmission follows a zone-based tree structure. Whenever a data packet reaches an on-tree zone, it will be forwarded to the leader first, and more on-tree zones will lead to more data packet forwarding and hence a higher packet transmission overhead. In a dense network, the number of empty zones reduces and there is more opportunity for a tree branch to be built between two neighboring zones, which increases the number of on-tree zones and forwarding overhead. The packet transmission overhead of ODMRP goes up because there are more nodes in the forwarding mesh.

For all the protocols, the disconnected topology graph in a sparse network leads to a longer joining delay (Fig. 8(d)). The followed slight increase of the joining delay at high node density is due to more transmission collisions. Such increase is more obvious for SPBM since its higher control overhead results in more collisions.

Overall, all the protocols perform better in a denser network. EGMP and SPBM have consistently higher delivery ratios than that of ODMRP. SPBM has a significantly higher control overhead and joining delay in a dense network as a result of its periodic multi-level flooding of membership management message, while EGMP remains to have the lowest delay as it allows group members to join and leave the group immediately on demand. SPBM has more transmissions in a sparse network due to the more frequent use of recovery forwarding of the underlying geographic unicast protocol, while the transmissions of both EGMP and ODMRP increase at a higher node density, as EGMP has more on-tree zones and ODMRP has more nodes in the forwarding mesh.

5.2.3 Effect of the group size

Next we evaluate the protocol performance with the group size varied from 10 members to 200 members.

Fig. 9 demonstrates that EGMP can scale to a large group size and perform well with various group sizes. When the group size increases, the delivery ratios of ODMRP and SPBM rise. When more nodes join the multicast group, the higher redundancy in the mesh structure of ODMRP will provide more robust delivery paths, while the membership of the squares in SPBM becomes more stable. In EGMP, when having a larger group size, a temporary disconnection of the tree structure may affect a larger number of members with more packet losses, resulting in a slight reduction in the delivery ratio. Overall, EGMP shows a more stable performance with different group sizes, while ODMRP and SPBM rely on a higher number of group members to maintain more stable multicast infrastructure and can perform poorly at a small group size.

In Fig. 9(b), ODMRP and SPBM are seen to have very high multicast control overheads when the group size is small, as in ODMRP all the mobile nodes are involved in the periodic flooding of JOIN QUERY message while in SPBM the proactive control message flooding causes more unnecessary overhead. While in EGMP, the multicast overhead remains very low at different group sizes and reduces as the group size increases. When more group members join the multicast group, there will be more control state aggregation for membership management by a zone leader and more member zones can share a tree branch. This confirms the efficiency of our zone-based tree structure. When the group size increases, the normalized control overheads of all the protocols drop accordingly with more data packets delivered. When the group size of EGMP is smaller, the ratio of the multicast overhead to the total control overhead is also smaller.

In Fig. 9(c), the data packet transmission overheads of all the protocols reduce when the group size increases as a result of the higher aggregations of packet transmissions. ODMRP has a high packet transmission overhead when the group size is small, since as mentioned above its periodic network-wide flooding of data packets generates relatively high normalized packet transmission overhead.

In Fig. 9(d), the change of group size has different impacts on the joining delay of the three protocols. For EGMP, as the group size increases, the joining delay of all the protocols drop accordingly with more data packets delivered. When the group size of EGMP is smaller, the ratio of the multicast overhead to the total control overhead is also smaller. In contrast, SPBM and ODMRP have much lower delivery ratios when the group sizes are small because SPBM has less stable membership and ODMRP has less robust mesh paths. SPBM and ODMRP also have much higher normalized control overheads at smaller group sizes due to their uses of periodic flooding of membership changes in the quad-tree. Relying on the periodic JOIN QUERY message to refresh the mesh structure for a node to join a group, the size does not have a significant impact on the joining delay of ODMRP.

In summary, EGMP has high delivery ratios for all the group sizes, and it does not incur unnecessary control overhead when there is no member management needed in a zone. In contrast, SPBM and ODMRP have much lower delivery ratios when the group sizes are small because SPBM has less stable membership and ODMRP has less robust mesh paths. SPBM and ODMRP also have much higher normalized control overheads at smaller group sizes due to their uses of periodic flooding messages regardless of the group size. The data transmission overheads for all the protocols reduce as the group size increases due to the aggregations of packet transmissions. The group size has little impact on the joining delay of EGMP, while SPBM has a significantly higher joining delay when the network is sparse.
5.2.4 Effect of the network size
To study the scalability of the protocol with network size, we varied the network range from $1500m \times 1500m$ to $3900m \times 3900m$. The node density is kept as before, thus the total number of nodes is varied from 156 nodes to 1056 nodes. Since the periodic local and network-wide message flooding in SPBM saturates the memory faster, we run simulations on SPBM with the network size increasing up to only $3300m \times 3300m$ with 756 nodes.

EGMP has a better scalability to the network size than ODMRP and SPBM as demonstrated in Fig. 10. The delivery ratios of ODMRP and SPBM drop faster than that of EGMP with the increase of network size. When the network size reaches $3900m \times 3900m$ with 1056 nodes, the difference between the delivery ratios of ODMRP and EGMP is more than 55%.

As expected, all the protocols have higher control overheads at a larger network size (Fig. 10 (b)). For ODMRP, more nodes are involved in the periodic JOIN QUERY message flooding. For EGMP, a larger network range leads to longer paths for the control messages at the upper tier. For the geographic-based unicast routing, more beacons will be generated with a larger number of network nodes. The control overhead of SPBM, however, is seen to rise much more sharply than those of EGMP and ODMRP, as a result of the increase of quad-tree levels in SPBM and the corresponding increase of periodic multi-level message flooding. As the network size increases, due to the longer packet forwarding paths, the total number of data packet transmissions of all the protocols also goes up (Fig. 10 (c)). Compared to EGMP, SPBM has more than double the packet transmissions in all the network sizes, and the difference becomes more evident at larger network sizes.

All three protocols also have longer joining delay when the network size increases as in Fig. 10 (d). The joining delay of ODMRP is significantly impacted by the network size, as both its periodic network-wide flooding of JOIN QUERY and its broadcast-based packet forwarding will not perform well. More data collisions during the flooding will result in a longer waiting time for a group member to receive the first data packet from the source, and a larger number of packet loss as confirmed by the low delivery ratio in Fig. 10 (a). For SPBM, with the increase of the number of the quad-tree levels, the membership change of a node may need to go through more levels to send out leading to a longer joining delay. The joining delay of EGMP only rises slightly, as a new member may need to connect to a farther away tree branch.

In summary, EGMP performs much better than SPBM and ODMRP in a large network, and has a significantly higher delivery ratio, lower control overhead, and lower joining delay due to its virtual-zone-based geometric membership management and transmission infrastructures.

6 Conclusions
There is an increasing demand and a big challenge to design more scalable and reliable multicast protocol over a dynamic ad hoc network (MANET). In this paper, we propose an efficient and scalable geographic multicast protocol, EGMP, for MANET. The scalability of EGMP is achieved through a two-tier virtual-zone-based structure, which takes advantage of the geometric information to greatly simplify the zone management and packet forwarding. A zone-based bi-directional multicast tree is built at the upper tier for more efficient multicast membership management and data delivery, while the intra-zone management is performed at the lower tier to realize the local membership management. The position information is used in the protocol to guide the zone structure building, multicast tree construction, maintenance, and multicast packet forwarding. Compared to conventional topology-based multicast protocols, the use of location information in EGMP significantly reduces the tree construction and maintenance overhead, and enables quicker tree structure adaptation to the network topology change. We also develop a scheme to handle the empty zone problem, which is challenging for the zone-based protocols. Additionally, EGMP makes use of geographic forwarding for reliable packet transmissions, and efficiently tracks the positions of multicast group members without resorting to an external location server.

We make a quantitative analysis on the control overhead of the proposed EGMP protocol and our results indicate that the per-node cost of EGMP keeps relatively constant with respect to the network size and the group size. We also performed extensive simulations to evaluate the performance of EGMP. Compared to the classical protocol ODMRP, both geographic multicast protocols SPBM and EGMP could achieve much higher delivery ratio in all circumstances, with respect to the variation of mobility, node density, group size and network range. However, compared to EGMP, SPBM incurs several times of control overhead, redundant packet transmissions and multicast group joining delay. Although SPBM is designed to be scalable to the group size, it has very low packet delivery ratio when the group size is small without a stable membership in each level of quad-tree square, and cannot perform well under a large network size due to the use of multi-level network-wide flooding of control messages. ODMRP takes advantage of broadcasting to achieve more efficient packet forwarding, but the transmissions are much more unreliable due to its difficulty of maintaining forwarding mesh under mobility, which leads to a lower packet delivery ratio. The multicast group joining delay of ODMRP is also much higher than that of EGMP.

Our results indicate that geometric information can be used to more efficiently construct and maintain multicast structure, and to achieve more scalable and reliable multicast transmissions in the presence of constant topology change of MANET. Our simulation results demonstrate that EGMP has high packet delivery ratio, and low control overhead and multicast group joining delay under all cases studied, and is scalable to both the group size and the network size. Compared to the geographic multicast protocol SPBM, it has significantly lower control overhead, data transmission overhead, and multicast group joining delay.

References


Xiaojing Xiang (M’06) received her BS and MS degrees in computer science from Nanjing University, Nanjing, China, and her PhD degree in computer science and engineering from the State University of New York at Buffalo, Buffalo, New York.

She is currently with Microsoft Corporation, Redmond, Washington. Her research interests include protocol design and analysis in mobile ad hoc networks, architecture design for service provisioning, routing and cross-layer protocol design in computer networks, pervasive computing and communications, as well as next generation Internet technologies.

Xin Wang (M’1 / ACM’4) received her BS and MS degrees in telecommunications engineering and wireless communications engineering from Beijing University of Posts and Telecommunications, Beijing, China, and her PhD degree in electrical and computer engineering from Columbia University, New York, NY.

She is currently an assistant professor in the department of Electrical and Computer Engineering of the State University of New York at Stony Brook, Stony Brook, New York. Before joining Stony Brook University, she was a member of Technical Staff in Bell Labs Research, Lucent Technologies, New Jersey and an assistant professor in the department of Computer Science and Engineering of the State University of New York at Buffalo, Buffalo, New York.

Her research interests include wireless networks and communications, mobile and distributed computing, and wireless services.

Yuanyuan Yang received the BEng and MS degrees in computer science and engineering from Tsinghua University, Beijing, China, and the MSE and PhD degrees in computer science from Johns Hopkins University, Baltimore, Maryland. She is a Professor of Computer Engineering and Computer Science at Stony Brook University, New York, and the Director of Communications and Devices Division at New York State Center of Excellence in Wireless and Information Technology (CE-WIT). Dr. Yang’s research interests include wireless networks, optical networks, high speed networks, and parallel and distributed computing systems. Dr. Yang has published over 200 papers in major journals and refereed conference proceedings and holds six U.S. patents in these areas. She is currently an Associate Editor for IEEE Transactions on Computers and a Subject Area Editor for Journal of Parallel and Distributed Computing. She has served as an Associate Editor for IEEE Transactions on Parallel and Distributed Systems.