A Peer-to-Peer Zone-Based Two-Level Link State Routing for Mobile Ad Hoc Networks

Mario Joa-Ng, *Student Member, IEEE*, and I-Tai Lu, *Senior Member, IEEE*

***Abstract*—A new global positioning system (GPS)-based routing protocol for ad hoc networks, called zone-based hierarchical link state (ZHLS) routing protocol, is proposed. In this protocol, the network is divided into nonoverlapping zones. Each node only knows the node connectivity within its zone and the zone connectivity of the whole network. The link state routing is performed on two levels: local node and global zone levels. Unlike other hierarchical protocols, there is no cluster head in this protocol. The zone level topological information is distributed to all nodes. This "peer-to-peer" manner mitigates traffic bottleneck, avoids single point of failure, and simplifies mobility management. Since only zone ID and node ID of a destination are needed for routing, the route from a source to a destination is adaptable to changing topology. The zone ID of the destination is found by sending one location request to every zone. Simulation results show that our location search scheme generates less overhead than the schemes based on flooding. The results also confirm that the communication overhead for creating and maintaining the topology in the proposed protocol is smaller than that in the flat LSR protocol. This new routing protocol provides a flexible, efficient, and effective approach to accommodate the changing topology in a wireless network environment.**

***Index Terms*—Ad hoc networks, global positioning system (GPS), hierarchical routing, link state, packet radio, routing, zone routing.**

I. Introduction

A mobile ad hoc network is a self-organizing and rapidly deployable network in which neither a wired backbone nor a centralized control exists. The network nodes communicate with one another over scarce wireless channels in a multihop fashion. The ad hoc network is adaptable to the highly dynamic topology resulting from the mobility of network nodes and the changing propagation conditions.

Various design choices for ad hoc networks are discussed in [1]. They are:

1) flat versus hierarchical architectures;
2) proactive versus reactive schemes.

In hierarchical architectures (e.g., the hierarchical spine routing protocol [2], [3]) the detail of the network topology is concealed by aggregating nodes into clusters and clusters into superclusters and so on [4]. Some nodes, such as cluster heads and gateway nodes, have a higher computation and communication burden than other nodes. Hence, the mobility management is complicated. The network reliability may also be affected due to single points of failure of these critical nodes. However, control messages may only have to propagate within a cluster. Thus, the multilevel hierarchy reduces the storage requirement and the communication overhead of large wireless networks. On the contrary, in flat architectures, all nodes carry the same responsibility. Flat architectures are not bandwidth efficient in large networks because control messages have to propagate globally throughout the network. The scalability gets worse when the number of nodes increases.

In proactive schemes, every node continuously maintains the complete routing information of the network. When a node needs to forward a packet, the route is readily available; thus, there is no delay in searching for a route. However, for a highly dynamic topology, the proactive schemes spend a significant amount of scarce wireless resource in keeping the complete routing information current. The proactive protocols such as the link state routing (LSR) protocol (open shortest path first) [5] and the distance vector routing protocol (Bellman–Ford) [5] were never designed to work in mobile networks [6]. They do not converge fast enough for the rapidly changing topology. Other distance vector routing protocols such as the destination-sequenced distance vector routing protocol [7] and the wireless routing protocol [8] were proposed to eliminate the counting-to-infinity and looping problems of the distributed Bellman–Ford algorithm. On the other hand, in reactive schemes (e.g., the ad hoc on-demand distance vector routing protocol [9], the temporally ordered routing algorithm [10], and the dynamic source routing protocol [11]), nodes only maintain the routes to active destinations. A route search is needed for every new destination. Therefore, the communication overhead is reduced at the expense of delay due to route search. Furthermore, the rapidly changing topology may break an active route and cause subsequent route search.

Haas and Pearlman proposed a hybrid reactive/proactive scheme called zone routing protocol (ZRP) [12]–[14]. Other ad hoc routing protocols can be found in [15]–[18]. In ZRP, each node proactively maintains the topological information within its routing zone (i.e., within a predefined distance) only. ZRP exploits the structure of the routing zone through a process known as bordercasting. Bordercasting allows a node to send messages to its peripheral nodes (nodes on the boundary of its routing zone) and prevents nonperipheral nodes from accessing the messages. Route discovery is efficiently done by

Manuscript received May 13, 1998; revised March 16, 1999. This work was supported in part by the NSF Graduate Research Traineeship CDA-92566881 and in part by ARO DAGG-55-98-1-0359.

M. Joa-Ng was with Polytechnic University, Farmingdale, NY 11735 USA. He is now with Telcordia Technologies, Inc., Red Bank, NJ 07701 USA (e-mail: mjoang@research.telcordia.com).

I.-T. Lu is with Polytechnic University, Farmingdale, NY 11735 USA (e-mail: itailu@rama.poly.edu)

Publisher Item Identifier S 0733-8716(99)04799-X.

bordercasting a route query to all the source's peripheral nodes, which in turn bordercast the query to their peripheral nodes and so on if the destination is not within their respective routing zones. Once the destination is found, a route reply is echoed back to the source. The ZRP path, which consists of a list of peripheral nodes between the source and the destination, is stored in the packet header or cached in the queried peripheral nodes. Any change in the peripheral nodes may render another route discovery.

We propose a "peer-to-peer" hierarchical routing protocol, zone-based hierarchical LSR protocol (ZHLS), which incorporates location information into a novel peer-to-peer hierarchical routing approach. The network is divided into nonoverlapping zones. Aggregating nodes into zones conceals the detail of the network topology. Initially, each node knows its own position and therefore zone ID through global positioning system (GPS). After the network is established, each node knows the low level (node level) topology about node connectivity within its zone and the high level (zone level) topology about zone connectivity of the whole network. A packet is forwarded by specifying the hierarchical address—zone ID and node ID—of a destination node in the packet header. Unlike other hierarchical protocols, there are no cluster heads in this protocol. The high level topological information is distributed to all nodes (i.e., in a peer-to-peer manner). This peer-to-peer characteristic avoids traffic bottleneck, prevents single point of failure and simplifies mobility management. Similar to ZRP, ZHLS is a hybrid reactive/proactive scheme. It is proactive if the destination is within the same zone of the source. Otherwise, it is reactive because a location search is needed to find the zone ID of the destination. However, unlike ZRP, ZHLS requires GPS and maintains a high level hierarchy for interzone routing. Location search is performed by unicasting one location request to each zone. Routing is done by specifying the zone ID and the node ID of the destination, instead of specifying an ordered list of all the intermediate nodes between the source and the destination. Intermediate link breakage may not cause any subsequent location search. Since the network consists of nonoverlapping zones in ZHLS, frequency reuse is readily deployable in ZHLS. The rest of this paper further elaborates this new routing protocol.

II. Zone Map

The network is divided into zones under ZHLS. A node knows its physical location by geolocation techniques such as GPS; then, it can determine its zone ID by mapping its physical location to a zone map, which has to be worked out at the design stage. The zone size depends on factors such as node mobility, network density, transmission power, and propagation characteristics. The partitioning can be based on simple geographic partitioning or on radio propagation partitioning. The geographic partitioning is much simpler and does not require any measurement of radio propagation characteristics, whereas the radio propagation partitioning is more accurate for frequency reuse. Radio propagation partitioning is preferable if a propagation measurement can be done at the design stage. However, some applications, such as emergency disaster rescue operation, tactical military communication, and law enforcement, do not permit such measurements. In such cases, a simple geographic partitioning has to be used.

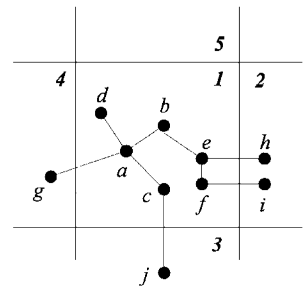


Fig. 1. Node level topology.

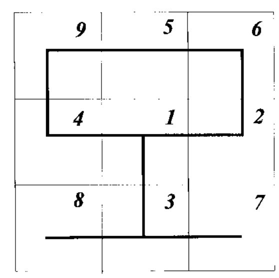


Fig. 2. Zone level topology.

III. Hierarchical Structure of ZHLS

Two levels of topology are defined in ZHLS: node level topology and zone level topology. If any two nodes are within the communication range, a physical link exists. The node level topology (Fig. 1) provides the information on how the nodes are connected together by these physical links. For example, in Fig. 1, if node a wants to send a data packet to node f, the data has to pass through a–b–e–f. If there is at least one physical link connecting any two zones, a virtual link then exists. The zone level topology (Fig. 2) tells how the zones are connected by these virtual links. For example, in Fig. 2, the virtual links between zone *4* and zone *3* are *4–1–3*. We will see later in Section VI how a node uses the node level topology to route a packet within a zone and how it uses the zone level topology to route a packet between the zones.

To facilitate this hierarchical LSR protocol, each node receives two types of link state packets (LSP's): node LSP's and zone LSP's. The node LSP of a particular node contains a list of its connected neighbors and is propagated locally within its zone. The zone LSP contains a list of its connected zones and is propagated globally throughout the network.

IV. Clustering

As mentioned in Section II, each node deploys a geolocation method to find its physical location and determines its zone ID by mapping its physical location to the zone map. Equipped with this zone ID, the node can start the intrazone clustering and then the interzone clustering procedures to build its routing tables.

A. Intrazone Clustering

Each node asynchronously broadcasts a link request. Nodes within its communication range in turn reply with link responses ⟨node ID, zone ID⟩. After all link responses are received, the node generates its node LSP that contains the node ID of its neighbors of the same zone and the zone ID of its neighbors of different zones. For example, in Fig. 1, nodes b, c, and d are node a's neighboring nodes, and zone *4* is its neighboring zone. It then propagates its node LSP locally throughout its zone via intermediate nodes. Since each node performs this procedure, a list of node LSP's, such as the one shown in Table I, can be stored in every node. However, nodes LSP's from other zones will not be stored because nodes LSP's are only propagated within their zone. The intrazone clustering procedure is depicted in Fig. 3(a)–(d).

After receiving all node LSP's of the same zone, each node will know the node level topology of that zone. The shortest path algorithm is used to build its intrazone routing table. Table II shows an example of the intrazone routing table of node a. Due to node mobility and channel fading, the previous procedure has to be performed periodically to detect and update any change in the physical links. If a node moves to another zone, its node LSP would be left in its old zone. So, a timer is set for each received node LSP, and any expired one will be deleted.

B. Interzone Clustering

Nodes may receive link responses from the nodes of their neighboring zones. These nodes are called gateway nodes. As shown in Fig. 1, nodes a, c, e, and f are gateway nodes of zone 1. Since node LSP's contain the zone ID's of the connected zones, each node will know which zones are connected to its zone. For example, based on the node LSP's in Table I, zones *2*, *3*, and *4* are connected zones of zone *1*. At the initialization stage, after making sure that all node LSP's are received, each node of the same zone generates the same zone LSP. The gateway nodes then broadcast the zone LSP throughout the network. Since every zone performs this procedure, a list of zone LSP's, identical to the one depicted in Table III, is stored by every node. So every node will know the zone level topology of the network.

Similar to the intrazone clustering, each node can determine its interzone routing table of the network from the zone LSP's. The interzone clustering procedure is depicted in Fig. 4(a)–(b). After each node receives all zone LSP's, the shortest path algorithm is used to find the shortest path in term of zone hops and build the interzone routing table. The interzone routing table of node a is shown in Table IV.

The previous procedure repeats periodically. However, the gateway nodes will not broadcast a zone LSP if its value is the same as the old one's. This takes advantage of the infrequent change in the virtual links and therefore reduces the amount of traffic. Moreover, unlike the node LSP's, no timer is set for zone LSP's. The zone LSP is updated only when any virtual link is broken or created.

Duplicate copies of zone LSP will not be forwarded. For example, a node receives two zone LSP's originated from different gateway nodes of the same zone. After forwarding the first one, the node will not forward the second one, as it is identical to the first one. Therefore, even though there may be more than one gateway node in a zone, only one zone LSP is

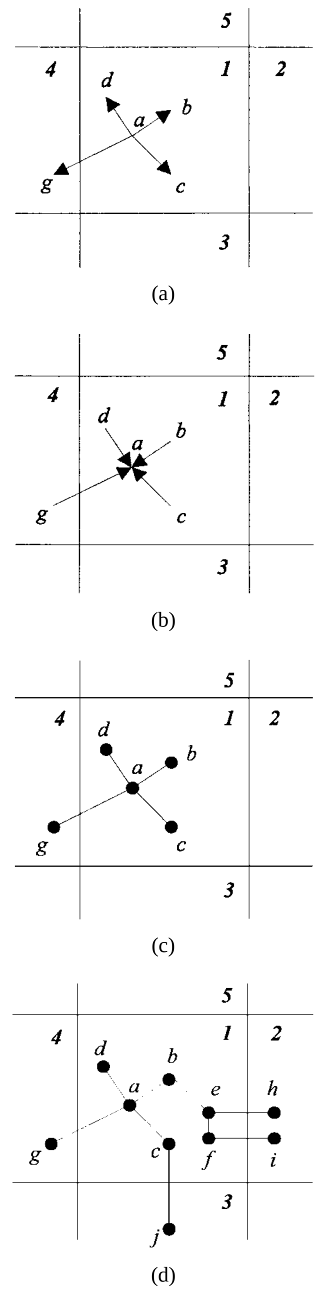


Fig. 3. Intrazone clustering procedure. (a) Node a broadcasts a link request to its neighbors. (b) Node a receives link responses from its neighbors. (c) Node a generates its own node LSP and broadcasts it throughout the zone. (d) All nodes perform the previous steps asynchronously.

TABLE I
NODE LSP's IN ZONE *1*

Source	Node LSP
a	*b,c,d,4*
b	*a,e*
c	*a,3*
d	*a*
e	*b,f,2*
f	*e,2*

TABLE II
INTRAZONE ROUTING TABLE OF NODE a

Destination	Next node
b	b
c	c
d	d
e	b
f	b
2	b
3	c
4	g

TABLE III
ZONE LSP's

Source	Zone LSP
1	2,3,4
2	1,6
3	1,7,8
4	1,9
5	6,9
6	2,5
7	3
8	3
9	4,5

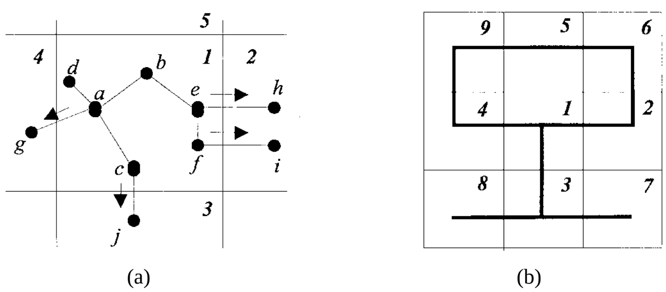


Fig. 4. Interzone clustering procedure. (a) Gateway nodes broadcast zone LSP's throughout the network. (b) Virtual links between adjacent zones are established.

TABLE IV
INTERZONE ROUTING TABLE OF NODE a

Destination zone	Next Zone	Next Node
2	2	b
3	3	c
4	4	g
5	4	g
6	2	b
7	3	c
8	3	c
9	4	g

generated from this zone. As the network spans a large area, zone LSP may not be received in the same order as they are sent. So, a time field is added to the zone LSP's, i.e., the zone LSP's are source sequenced. Since zone LSP's may be sent by more than one gateway node, the clocks of the nodes in the same zone have to be synchronized. The local synchronization is readily available if GPS is used. If the received zone LSP's are out of order, obsolete zone LSP's are deleted.

V. MATHEMATICAL ANALYSIS

A. Communication Overhead for Creating Topology

It is noteworthy to compare the communication overhead for creating the topology between the flat LSR protocol [5] and the ZHLS. Consider a network with N nodes. In LSR, each node will generate one LSP, and every other node has to forward it once. Therefore, the total amount of communication overhead generated by LSR S_{LSR} is

$$S_{\mathrm{LSR}} = N^2 \text{ messages.} \tag{1}$$

In ZHLS, the network is partitioned into M zones. Assuming that the nodes are distributed evenly throughout the network, each zone will have (N/M) nodes. The amount of communication overhead of node LSP's S_{node} becomes $(N/M)^2$ per zone or $M(N/M)^2 = N^2/M$ in the network. As each zone generates one zone LSP and every node has to forward all zone LSP's once, the amount of communication overhead of zone LSP's S_{zone} becomes NM. So, the total amount of communication overhead generated by ZHLS S_{ZHLS} is

$$S_{\mathrm{ZHLS}} = N^2/M + NM \text{ messages.} \tag{2}$$

It can be shown that S_{ZHLS} is always smaller than S_{LSR}, for $2 < M < N/2$. The number of zones will affect the communication overhead generated by ZHLS. When the number of zones M increases, S_{node} will decrease, and S_{zone} will increase. The minimum S_{ZHLS} is achieved when $dS_{\mathrm{ZHLS}}/dM = 0$ [it is a minimum value as $d^2S_{\mathrm{ZHLS}}/dM^2 = 2N^2/M^3 > 0$]. Therefore, the optimal number of zones to achieve the minimum S_{ZHLS} is

$$0 = \frac{dS_{\mathrm{ZHLS}}}{dM} = N - \frac{N^2}{M_{\mathrm{opt}}^2} \therefore M_{\mathrm{opt}} = \sqrt{N} \tag{3}$$

and the minimum S_{ZHLS} is

$$S_{\mathrm{ZHLS_{MIN}}} = 2N^{3/2} \text{ messages.} \tag{4}$$

A simulation in Section VII-A illustrates that the ZHLS routing protocol generates less communication overhead than the flat LSR protocol.

B. Communication Overhead Induced by Node Mobility

Locally propagated node LSP's are generated if the physical link between any two nodes creates or breaks due to node movement. On the other hand, globally propagated zone LSP's are generated only when the number of physical links connecting any two zones increases from zero or decreases to zero. The zone size of a network is chosen such that the average number of physical links connecting two zones is much higher than zero, i.e., the chance of having no physical links connecting two zones is small. Therefore, we expect that the transitions between the state of having no physical link and that of having physical links are infrequent, and the zone level topology is relatively robust to node movement compared to the node level topology. A simulation in Section VII-C shows how the number of physical links connecting two zones varies with various system parameters, such as number of nodes, communication range, and zone size. Assuming that, in the flat LSR protocol [5], the percentage of nodes generating LSP's in one clustering cycle due to changes in physical links is

p_{LSR}, the total amount of communication overhead induced by mobility in LSR K_{LSR} becomes

$$K_{\mathrm{LSR}} = N^2 p_{\mathrm{LSR}} \text{ messages/cycle.} \quad (5)$$

We also denote the percentage of nodes generating node LSP's in one cycle due to changes in physical links to be p_{node} and that of zones generating zone LSP's in one cycle to be p_{zone}. Therefore, the total amount of communication overhead induced by mobility in ZHLS K_{ZHLS} is

$$K_{\mathrm{ZHLS}} = N^2 p_{\mathrm{node}}/M + NMp_{\mathrm{zone}} \text{ messages/cycle.} \quad (6)$$

Since the zone level topology is more robust than the node level topology, $p_{\mathrm{zone}} < p_{\mathrm{node}} \cong p_{\mathrm{LSR}}$. Thus

$$K_{\mathrm{ZHLS}} < N^2 p_{\mathrm{LSR}}/M + NMp_{\mathrm{LSR}} < N^2 p_{\mathrm{LSR}} = K_{\mathrm{LSR}}.$$

Our simulation in Section VII-A illustrates that the communication overhead for maintaining the changing topology in ZHLS is much smaller than that in LSR. Thus, hierarchical routing reduces the overhead induced by mobility.

VI. Location Search and Routing Mechanism

In the current IP protocol, routing is designed to be hierarchical [5]. The network is partitioned into different subnetworks. Since the nodes in the IP network are stationary, each node is associated with a hierarchical IP address, which contains a fixed subnet ID. Similarly, in ZHLS, the network is partitioned into zones. However, the mobility of the nodes forbids us from associating them with fixed zone ID's. Therefore, a source needs to search for the zone ID of a destination node before any data transmission can start.

For example, node a wants to send data to node z (Fig. 5). Before sending data to node z, node a will check if node z exists in its intrazone routing table. If so, node a will route the data to node z according to its intrazone routing table. Otherwise, node z is in a different zone and node a will send a location request $\langle a,\ 1(a\text{'s zone ID}),\ z, X\rangle$ to every other zone X. Each intermediate node routes the location request destined for zone X to zone X according to its interzone routing table. The path from node a to zone X is adaptable to changing topology. A gateway node of each zone will receive the location request and check its intrazone routing table to see if node z exists in its zone. ZHLS does not limit one gateway node per zone. This avoids single point of failure. A gateway node in the same zone of node z will reply with a location response $\langle z,\ 5\ (z\text{'s zone ID}),\ a,\ 1\rangle$. As we will show in Section VII-E, this search incurs a much smaller amount of overhead than a corresponding search—flooding—in the dynamic source routing protocol (DSR) [11] and the ad hoc on demand distance vector routing protocol (AODV) [9].

The zone ID (5) and the node ID (z) are then specified in the data header. Node a will route the data via node g to zone *5* according to its interzone routing table (Table IV). All intermediate nodes, except those in zone *5*, route the data to zone *5* according to their interzone routing tables. When the data reaches zone *5*, the intermediate nodes will instead use their intrazone routing tables to route the data to node z.

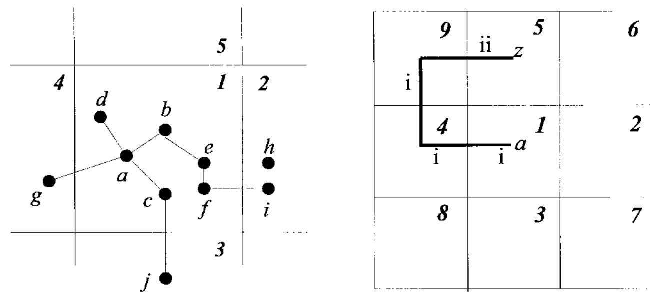
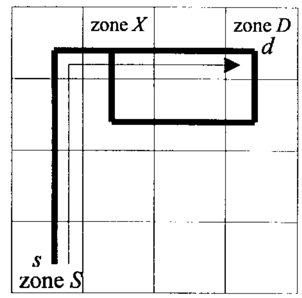
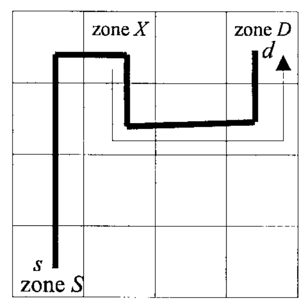


Fig. 5. Routing path i uses the interzone routing table, and path ii uses intrazone routing table.



(a)



(b)

Fig. 6. Virtual backbone changes. (a) Time t1. (b) Time t2.

Even if the node level or the zone level topology changes during the data transmission, routing can still be done properly. For example, the zone level topologies at times t1 and t2 are shown in Fig. 6(a) and (b), respectively. Nodes in zone X can still route the data to node d even though one of the virtual paths between zone X and zone D (zone ID of node d) is broken at the time of transmission. Moreover, the packet is sent properly even if node s has slightly outdated interzone information because only zone ID and node ID of a destination are needed for routing; the route is adaptable to dynamic topology. On the contrary, in the DSR protocol [11], a subsequent search has to be performed to find a route again whenever the current route is broken due to node mobility.

It is possible that more than one cluster exists within a zone even if the zone size is chosen according to the typical transmission range of a node. For example, there may be a large obstacle such as a hill, a building, etc., in the zone that blocks radio communication. As shown in Fig. 7, there are two clusters in the same zone. Every node will receive two zone LSP's from zone *1*. To identify them, one additional field, the smallest node ID, is added to the zone LSP. In Fig. 7, every

TABLE V
COMPARISON OF COMMUNICATION OVERHEAD GENERATED IN ZHLS AND THAT IN LSR UNDER DIFFERENT NUMBER OF NODES N. THE NETWORK IS PARTITIONED INTO NINE ZONES ($M = 9$) IN THE CASE OF ZHLS

No of nodes N	ZHLS						LSR	
	Zone LSP S_{zone}		Node LSP S_{node}		Total LSP S_{ZHLS}		LSP S_{LSR}	
	simulation result	predicted result NM	simulation result	predicted result N^2/M	simulation result	predicted result $NM+N^2/M$	simulation result	predicted result N^2
100	879	900	1136	1111	2015	2011	10000	10000
200	1765	1800	4576	4444	6341	6244	40000	40000
300	2669	2700	10248	10000	12917	12700	90000	90000
400	3558	3600	18200	17778	21758	21378	160000	160000
500	4454	4500	28262	27778	32716	32278	250000	250000

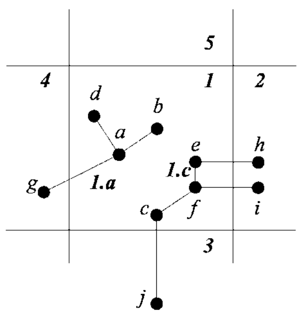


Fig. 7. More than one cluster in one zone.

node receives two zone LSP's—LSP *1.a* and LSP *1.c*—with different zone connectivity information from zone *1*, i.e., zone *1* is split into zones *1.a* and *1.c*. The rest of the processing will be the same except that the zone field will have one more subfield.

VII. SIMULATION

Hierarchical approach reduces the communication and the storage requirements significantly [19], [20]. In this section, we will compare the amount of communication overhead for building the topology in the ZHLS and the amount in the flat LSR [5]. Since hierarchical routing is used in ZHLS, the path to a destination may be suboptimal. We will study the impact on path length when ZHLS is used. Also, it is crucial to know whether the zone level topology is robust under a dynamic topology. If the zone level topology rarely changes, a small number of globally propagated messages are generated. In addition, hierarchical approach reduces the amount of communication overhead induced by node mobility. We will compare the amount of communication overhead induced by node mobility generated in ZHLS and the amount generated in LSR. Finally, we will investigate the amount of overhead generated in the location search scheme of ZHLS and compare the amount with that generated in the DSR [11] and the AODV [9].

A. Communication Overhead for Creating Topology

A Maisie [21], [22] simulation is developed to count the amounts of communication overhead for creating the topology in ZHLS and in LSR. First of all, we study how the communication overhead varies with the number of nodes N. The nodes are randomly located inside a square of length 99 units. Each node has a communication range l of 20 units. In the case of ZHLS, the network is partitioned into nine square zones ($M = 9$), each of length 33 units. We run the simulation from $N = 100$ to $N = 500$. Table V shows the simulation results and the results predicted in (1) and (2). It illustrates that the simulation results match with the predicted results. We also observe that $S_{\rm zone}$ varies linearly with N, whereas $S_{\rm node}$ and $S_{\rm LSR}$ vary with N^2. Most importantly, the total amount of communication overhead generated in ZHLS is much smaller than that generated in LSR.

We now study the effect of the number of zones M on the communication overhead generated in ZHLS. We have shown that $S_{\rm node} = N^2/M$ and $S_{\rm zone} = NM$. When M increases, $S_{\rm node}$ will decrease whereas $S_{\rm zone}$ will increase. According to (3) and (4), the minimum of $S_{\rm ZHLS}$ is $2N^{3/2}$ when the number of zones is $\sqrt{N}$. In our simulation, the network has 500 nodes, each with a communication range of 20 units. The network is within a square of 100 units. Table VI shows how $S_{\rm zone}$, $S_{\rm node}$, and $S_{\rm ZHLS}$ vary with different number of zones M. With $N = 500$, $M_{\rm opt} = \sqrt{N} \approx 22$. Our simulation results show that $M_{\rm opt}$ is 25.

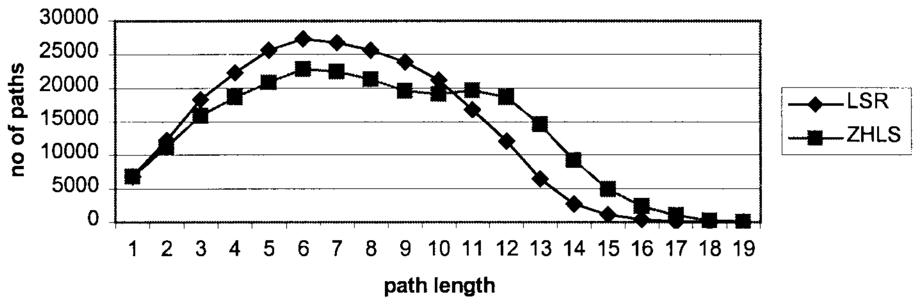
B. Path Length

Hierarchical routing may give a suboptimal path between two nodes, and so the length of ZHLS path may be higher than that of LSR path. In this section, we study how the path length is affected when ZHLS is used.

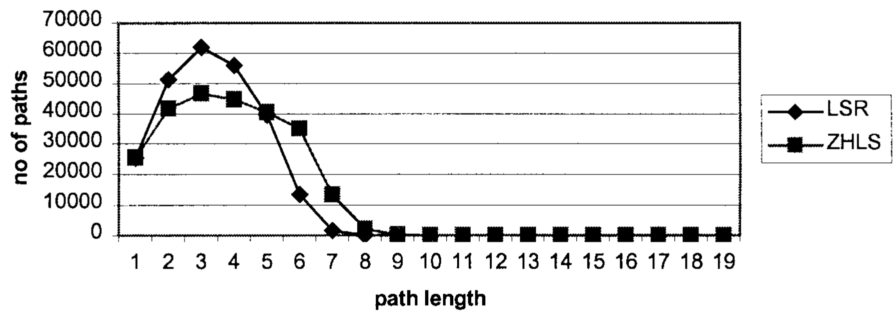
The Maisie [21], [22] simulation is added with the functionality of finding both the ZHLS path and the LSR path between every pair of nodes in the network and counting the corresponding path lengths. In the simulation, the network is within a square of 99 units and is partitioned into nine zones in ZHLS. Five hundred nodes are randomly located inside the network. Fig. 8 shows the distributions of path lengths between every pair of nodes in the network for ZHLS and for LSR under various communication ranges l. Our results show that a suboptimal path is rendered in ZHLS. The average path lengths of ZHLS and LSR in the network are shown in Table VII. The average path increases by about 15% when ZHLS is used. We run the simulation again for a network with 100

TABLE VI
THE AMOUNTS OF ZONE LSP OVERHEAD S_{zone}, NODE LSP OVERHEAD S_{node}, AND TOTAL OVERHEAD S_{ZHLS} UNDER DIFFERENT NUMBER OF ZONES M. THE NUMBER OF NODES N IS 500

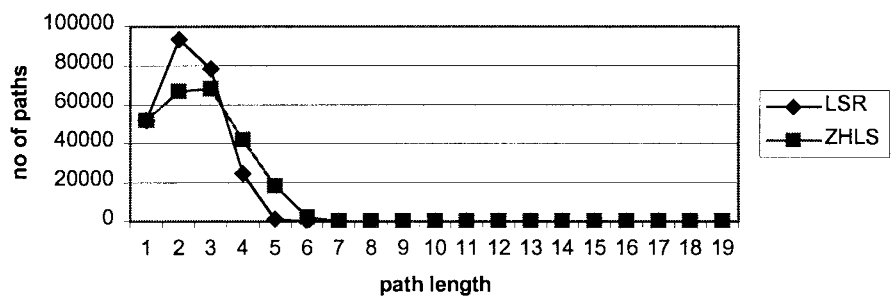
No of zones *M*	Zone LSP S_{zone}		Node LSP S_{node}		Total overhead S_{ZHLS}	
	simulation result	predicted result *NM*	simulation result	predicted result N^2/M	simulation result	predicted $NM+N^2/M$
4	1786	2000	62958	62500	64744	64500
9	4454	4500	28262	27778	32716	32278
16	7992	8000	16130	15625	24122	23625
25	12500	12500	10600	10000	23100	22500
36	18000	18000	7304	6944	25304	24944
49	24500	24500	5526	5102	30026	29602



(a)



(b)



(c)

Fig. 8. Distribution of path lengths between every pair of nodes in the network. For a network of 500 nodes, there are 250 000 paths. (a) Communication range = 10 units. (b) Communication range = 20 units. (c) Communication range = 30 units.

nodes. Table VIII shows the simulation results. The average path increases by around 13% when ZHLS is used. Therefore, the impact on the path length by ZHLS is not significant.

Another interesting result is that the average path length of ZHLS increases only slightly from 3.81 to 4.39 (for $l = 20$) and from 2.65 to 2.81 (for $l = 30$) when the number of nodes decreases from 500 to 100. Thus, the number of nodes has a small impact on the path length.

C. Stability of Zone Level Topology

We write a simple MATLAB program to study how the number of physical links connecting two neighboring zones varies with the number of nodes N and the communication range l. The network is within a rectangle of size 100×200 units and is divided into two square zones, each of length 100 units. N nodes are placed randomly within the network. A

TABLE VII
AVERAGE PATH LENGTHS OF ZHLS AND LSR IN A NETWORK OF 500 NODES

Communication Range	average LSR path length (in hops)	average ZHLS path length (in hops)	% of increase
10	7.06	7.94	12.5%
20	3.32	3.81	14.8%
30	2.31	2.65	14.7%

TABLE VIII
AVERAGE PATH LENGTHS OF ZHLS AND LSR IN A NETWORK OF 100 NODES

Communication Range	average LSR path length (in hops)	average ZHLS path length (in hops)	% of increase
20	3.91	4.39	12.3%
30	2.48	2.81	13.3%

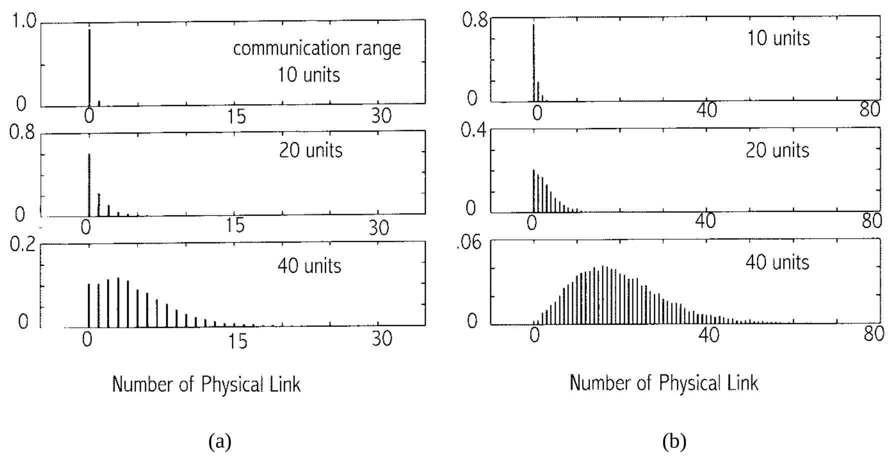


Fig. 9. Probability density function of physical links connecting two zones. The network is within a rectangle of size 100 × 200 units and is divided into two square zones, each of length 100 units: (a) 20 nodes. (b) 40 nodes.

physical link connects any two nodes if the distance between them is less than l units. Then, we count the number of physical links (m) connecting nodes of different zones. We repeat the process 20 000 times. From the result, we can form the probability density function (pdf) of having m physical links between two zones.

We run the simulation for $N = 20$ and $N = 40$ and for $l = 10$, $l = 20$, and $l = 40$ units. Fig. 9 illustrates how the pdf of m physical links depends on the number of nodes N and the communication range l. It is observed that higher node density and higher communication range can increase the chance of having large number of physical links connecting the zones, i.e., the probability of having no physical link connecting the zones is small. We will see how the number of nodes and the node mobility affect the stability of zone layer topology in the next section.

D. Mobility Effect

The Maisie [21], [22] simulation is extended to support mobility. It counts the amount of communication overhead induced by node mobility in ZHLS (K_{ZHLS}) and that in LSR (K_{LSR}). A network of N nodes is within a square of length 99 units. Each node has a communication range l of 20 units and node mobility ν of 4 units/cycle. When a node locates at (x, y) at time tth cycle, it will move to $(x+\Delta x, y+\Delta y)$ at time $t+1$th cycle, where Δx and Δy are uniformly distributed within $[-\nu, +\nu]$. In the case of ZHLS, the network is partitioned into $M = 9$ square zones, each of length 33 units. First of all, we compare the impacts on the topology in ZHLS and in LSR induced by node mobility as well as the associated overhead under different number of nodes. The simulation is run for ten cycles. The average number of LSP's and the average amount of communication overhead per cycle in ZHLS and LSR are calculated. Table IX shows that the communication overhead to maintain the changing topology in ZHLS (K_{ZHLS}) is smaller than that in LSR (K_{LSR}). This supports our claim that hierarchical approach reduces the amount of overhead of dynamic changing topology. Moreover, the percentage of zones generating zone LSP's due to node mobility (p_{zone}) is always smaller than that of nodes generating node LSP's (p_{node}) as well as that of nodes generating LSP's in LSR

TABLE IX
COMPARISON OF COMMUNICATION OVERHEAD GENERATED IN ZHLS AND THAT IN LSR UNDER DIFFERENT NUMBER OF NODES N. THE NETWORK IS PARTITIONED INTO NINE ZONES ($M = 9$) IN THE CASE OF ZHLS. THE MOBILITY LEVEL IS 4 UNITS/CYCLE

No of nodes N	ZHLS							LSR		
	Zone			Node			Total			
	No of LSPs n_{zone}	% of zones generating LSPs p_{zone} n_{zone}/M	Overhead K_{zone}	No of LSPs n_{node}	% of nodes generating LSPs p_{node} n_{node}/N	Overhead K_{node}	Overhead K_{ZHLS}	No of LSPs n_{LSR}	% of nodes generating LSPs p_{LSR} n_{LSR}/N	Overhead K_{LSR}
100	5.4	60.00%	527	82.4	82.40%	946	1473	86.7	86.70%	8670
200	1.6	17.78%	316	195.7	97.85%	4432	4748	196.7	98.35%	39340
300	0.8	8.89%	238	298.0	99.33%	10096	10334	298.1	99.37%	89430
400	0.4	4.44%	160	399.7	99.93%	18051	18211	399.9	99.98%	159960
500	0	0%	0	499.8	99.96%	28095	28095	499.9	99.98%	249950

TABLE X
COMPARISON OF COMMUNICATION OVERHEAD GENERATED IN ZHLS AND THAT IN LSR UNDER DIFFERENT LEVELS OF NODE MOBILITY ν. THE NETWORK CONTAINS 100 NODES AND IS PARTITIONED INTO NINE ZONES

Mobility level ν units/cycle	**Zone LSP**			**Node LSP**		
	No of LSPs n_{zone}	**% of zones generating LSPs p_{zone} n_{zone}/M**	**Overhead K_{zone}**	**No of LSPs n_{node}**	**% of nodes generating LSPs p_{node} n_{node}/N**	**Overhead K_{node}**
2	2.90	32.22%	284	62.76	62.76%	724
4	4.40	48.89%	433	87.26	87.26%	1033
6	4.84	53.78%	475	93.10	93.10%	1091
8	5.24	58.22%	513	96.58	96.58%	1111
10	5.92	65.78%	579	98.44	98.44%	1158

(p_{LSR}); thus, the zone level topology is relatively robust to mobility. When the number of nodes N increases, p_{zone} decreases, whereas p_{node} and p_{LSR} increase. As N reaches 500, the zone level topology is not affected by the node movement. The virtual links do not break because the network is so dense that the breakage of interzone physical links caused by node movement is compensated by the creation of interzone physical links due to node movement.

We proceed to study how the mobility level affects the amount of communication overhead in ZHLS. The same network of 100 nodes is used. It is within a square of length 99 units and is partitioned into nine square zones, each of length 33 units. We run the simulation for a period of 50 simulation cycles and find the average number of LSP's and the average communication overhead per cycle in ZHLS under different levels of mobility ν. Table X shows that both types of LSP's—node and zone—increase with increasing mobility level for a network of 100 nodes. However, the zone level topology is more robust than the node level topology, as the percentage of zones generating zone LSP's is much smaller than the percentage of nodes generating node LSP's.

E. Location Search Overhead

In this section, we attempt to compare the amount of overhead generated for searching a destination node in ZHLS and that in DSR and AODV. In ZHLS, a source node searches for a destination node by broadcasting a location request to every other zone (broadcasting on zone level topology). In DSR and AODV, a source node searches by flooding. In our simulation, N nodes are randomly located inside a square of length 99 units. Each node has a communication range l of 20 units. The network is partitioned into nine zones. We run a location search for every pair of source–destination and find the average communication overhead per search in ZHLS. The overhead for route search in DSR and AODV is always N because flooding is used.

Table XI depicts how the search overhead varies with the number of nodes in ZHLS and in DSR and AODV. The results confirm our claim that broadcasting on zone level topology, which is used by ZHLS, saves more bandwidth than flooding used by DSR and AODV. It is interesting to find that the average overhead in ZHLS (L_{ZHLS}) is unchanged even if the number of nodes increases. As shown in Section VII-B, the *average path length* changes only slightly with increasing number of nodes. Since $L_{\mathrm{ZHLS}} = (M - 1) \cdot$ average path length, L_{ZHLS} is scalable when the number of nodes increases.

VIII. CONCLUSION

We have presented a new routing protocol for mobile ad hoc networks, called ZHLS. The main idea is to use the hierarchical routing approach in a peer-to-peer way for large mobile wireless networks.

Scarce wireless resource is an important factor in designing a routing protocol for mobile ad hoc networks. Simulation data has shown that, for a network of size N, the amount of communication overhead in the proposed ZHLS is of the order

TABLE XI
COMPARISON OF SEARCH OVERHEAD GENERATED IN ZHLS AND THAT IN DSR AND AODV UNDER DIFFERENT NUMBER OF NODES N. THE NETWORK IS INSIDE A SQUARE OF LENGTH 99 UNITS AND IS PARTITIONED INTO NINE ZONES. EACH NODE HAS A COMMUNICATION RANGE OF 20 UNITS

No of nodes N	Average overhead for location search in ZHLS L_{ZHLS}	Overhead for route search in DSR and AODV $L_{flooding}$
100	27.72	100
200	24.30	200
300	23.28	300
400	22.73	400
500	22.18	500

of $N^{3/2}$, whereas that in the flat LSR protocol [5] is of the order of N^2. It is clear that ZHLS is more bandwidth efficient and scales better than LSR.

Node mobility is another important design consideration. In ZHLS, node mobility does not usually have a global effect, and not many globally propagated control messages are generated. On the other hand, in a flat protocol, node mobility has a global effect and creates a significant amount of globally propagated control messages. Simulation results have asserted that the communication overhead induced by mobility is much smaller in ZHLS than that in LSR. Also, simulation results have confirmed that the chance of a virtual link breakage in ZHLS is smaller than that of a physical link breakage in a flat protocol. So, the zone level topology in ZHLS is relatively stable.

Unlike other hierarchical routing protocols [3], [15], our peer-to-peer hierarchical protocol does not designate any node as cluster head. As a result, a single point of failure and traffic bottleneck can be avoided. Mobility management in ZHLS is simple as all nodes play the same role in the network.

Owing to the bandwidth constraint, most ad hoc routing protocols [9]–[11] are reactive; each node only maintains the routing information of active destinations. Therefore, path search is necessary for all reactive schemes. In reactive schemes with a flat architecture [9]-[11], flooding is the only way to search for a route to a destination. On the other hand, in ZHLS that has a hierarchical architecture, broadcasting on the zone level topology is used for location searches. Simulation results demonstrate that broadcasting on the zone level topology in ZHLS generates a smaller amount of communication overhead than flooding in a flat protocol. Moreover, in the flat and reactive protocols [9]–[11], path containing the nodes between a source and a destination is needed for routing. Any intermediate link breakage will invalidate the path and render subsequent search. On the contrary, in ZHLS, only the zone ID and the node ID of a destination are needed for routing. The actual routing path is adaptable to the changing topology, and a subsequent search is not required as long as the destination does not hand off to another zone. Handoff management has been included in ZHLS to alleviate the handoff effect [23].

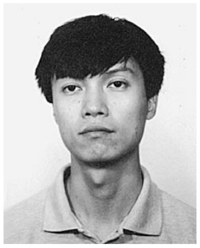
In ZHLS, the network is divided into zones as in cellular networks. Frequency reuse commonly used in cellular networks is readily deployable in our protocol to lessen the bandwidth constraint. Furthermore, handoff concept is borrowed from the cellular networks to design our handoff management [23].

REFERENCES

[1] Z. J. Haas, "Panel report on ad hoc networks—MILCOM'97," *ACM Mobile Comput. Commun. Rev.*, vol. 2, pp. 15–18, Jan. 1998.
[2] B. Das and V. Bharghavan, "Routing in ad-hoc networks using minimum connected dominating sets," in *Proc. IEEE ICC'97*.
[3] B. Das, R. Sivakumar, and V. Bharghavan, "Routing in ad hoc networks using a virtual backbone," in *Proc. IEEE IC3N'97*, pp. 1–20.
[4] G. S. Lauer, "Packet-radio routing," *Routing in Communications Networks*, M. E. Steenstrup, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1995, pp. 375–379.
[5] R. Perlman, *Interconnections: Bridges and Routers*. Reading, MA: Addison-Wesley, 1992, pp. 149–152, 205–233.
[6] J. P. Macker and M. S. Corson, "Mobile ad hoc networking and the IETF," *ACM Mobile Comput. Commun. Rev.*, vol. 2, pp. 9–14, Jan. 1998.
[7] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *ACM Comput. Commun. Rev.*, vol. 24, pp. 234–244, Oct. 1994.
[8] S. Murthy and J. J. Garcia-Luna-Aceves, "An efficient routing protocol for wireless networks," *ACM Mobile Networks Applicat. J.*, vol. 1, no. 2, pp. 183–197, 1996.
[9] C. E. Perkins, "Ad hoc on demand distance vector (AODV) routing," Internet Draft, Nov. 1997.
[10] V. D. Park and M. S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proc. IEEE INFOCOM'97*, Kobe, Japan.
[11] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Computing*, T. Imielinski and H. Korth, Eds. Norwell, MA: Kluwer, 1996.
[12] M. R. Pearlman and Z. J. Haas, "Determining the optimal configurations for the zone routing protocol," this issue, pp. 1395–1414.
[13] Z. J. Haas and M. R. Pearlman, "The performance of query control scheme for the zone routing protocol," in *Proc. ACM SIGCOMM'98*, Vancouver, Canada.
[14] Z. J. Haas, "The zone routing protocol (ZRP) for ad hoc networks," Internet Draft, Nov. 1997.
[15] M. Gerla and J. T. Tsai, "Multicluster, mobile, multimedia radio network," *ACM Wireless Networks*, vol. 1, no. 3, pp. 255–265, 1995.
[16] D. J. Baker, A. Ephremides, and J. A. Flynn, "The design and simulation of a mobile radio network with distributed control," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 226–237, Jan. 1984.
[17] A. Ephremides, J. E. Wieselthier, and D. J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling, *Proc. IEEE*, vol. 75, no. 1, pp. 56–73, Jan. 1987.
[18] J. Sharony, "An architecture for mobile radio networks with dynamically changing topology using virtual subnets," *ACM Mobile Networks Applicat. J.*, vol. 1, no. 1, pp. 75–86, 1996.
[19] L. Kleinrock and F. Kamoun, "Hierarchical routing for large networks: Performance evaluation and optimization," *Comput. Networks*, vol. 1, no. 3, pp. 155–174, 1977.
[20] J. Behrens and J. J. Garcia-Luna-Aceves, "Hierarchical routing using link vectors," in *Proc. IEEE INFOCOM'98*, San Francisco, CA.
[21] UCLA Parallel Computing Lab, "Maisie user manual release 2.2," Dec. 1995.

[22] J. Short, R. Bagrodia, and L. Kleinrock, "Mobile wireless network system simulation," in *Proc. ACM Mobile Commun. Network Conf.*, Berkeley, CA, Nov. 1995.

[23] M. Joa-Ng, "Routing protocol and medium access protocol for mobile ad hoc networks," Ph.D. dissertation, Dept. Elect. Eng., Polytechnic Univ., Brooklyn, NY, 1999.



Mario Joa-Ng (S'98) received the B.Sc. degree in physics and the M.Sc. degree (with distinction) in electrical engineering from the University of Hong Kong, in 1986 and 1994, respectively, and the Ph.D. degree in electrical engineering from Polytechnic University, Brooklyn, NY, in 1999.

He is currently a Research Scientist at Telcordia Technologies, Red Bank, NJ. He has held summer positions at NTT Multimedia Laboratories and Hughes Research Laboratories. His main research interests lie in mobile ad hoc network architecture, ad hoc routing protocol, wireless medium access control, and simulation and performance evaluation of wireless networks.

Dr. Joa-Ng was awarded an NSF Graduate Research Traineeship and was a recipient of the NSF Summer Institute '97. He is a student member of ACM and a member of Tau Beta Pi.

I-Tai Lu (M'87–SM'95) received the B.S.E.E. degree from National Chiao-Tung University in 1976, the M.S.E.E. degree from National Taiwan University in 1978, and the Ph.D.E.E. degree from Polytechnic University, Brooklyn, NY, in 1985.

From 1978 to 1980, he was an instructor in the Department of Communication Engineering, National Chiao-Tung University. Presently, he is an Associate Professor of Electrical Engineering, Polytechnic University. His current research interests are wireless communications, underwater and structural acoustics, electromagnetic waves, inverse problems, and signal processing. He has published over 45 journal papers, and 35 proceeding papers and/or book chapters, given over 30 invited lectures, and conducted over 100 conference/workshop/seminar presentations.

Dr. Lu has received the NSF Engineering Initiation Award, a Torchlight's Excellent Youth Award, a Ministry of Education Award, an Academic Excellency Award from the Alumni Association of National Chiao Tung University, the County Magistrate's Award, and awards from the Electrical Engineering Foundation of National Taiwan University, Dai-ichi Kangyo Bank Ltd., and Hong's Foundation for Education and Culture. He is a member of Sigma Xi, Commission B of URSI, the Institute for Theory and Computation of the Electromagnetic Academy, Who's Who Among Asian Americans, Who's Who in Science and Engineering, the Acoustical Society of America, and the Optical Society of America.