Route Discovery and Reconstruction in Mobile Ad hoc Networks

Robin Chellappa+, Andrew Jennings++
School of Electrical and Computer Engineering
RMIT University
Melbourne, Australia
+robin.chellappa@ieee.org, ++ajennings@rmit.edu.au

Nirmala Shenoy
Department of Information Technology
Rochester Institute of Technology
New York, U.S.A
ns@it.rit.edu

***Abstract*-The dynamic topology of ad hoc networks makes route discovery and reconstruction a non-trivial task. In this work we propose the use of mobility prediction to assist in ad hoc routing. A method of route discovery and route reconstruction using the sectorized mobility prediction technique is presented. Simulation and analytical results of the proposed routing strategy find it to offer considerable reduction in the amount of routing traffic generated during the route discovery phase.**

I. INTRODUCTION

Ad hoc networks are self-organizing and self-configuring multi-hop networks with dynamic topologies effected by node mobility. All nodes are capable of movement and can be connected dynamically in an arbitrary manner. Nodes of the network function as mobile routers that discover and maintain routes to other nodes in the network. In order to define a semblance of grouping in ad hoc networks and enhance network manageability and energy economy the concept of clustering has been introduced. The cluster head of each cluster has a bi-directional link with each of its member nodes. The clusters are dynamic groupings and hence mobile nodes frequently leave their current cluster of membership and join a neighbouring cluster. As routing is the process of choosing a path over which data is to be sent, node mobility and cluster changes make routing in ad hoc networks a non-trivial task. User mobility can be modeled in a continuous manner based on location tracking techniques such as the Global Positioning System (GPS). Ad hoc nodes though may not support GPS due to energy restriction of the nodes. Ad hoc networks operate on finite energy reserves, which requires that node mobility be simplified to help reduce location updates. We argue that an exact geographic position of the user is not necessary. An accurate position of the user relative to its present and future connection points is preferred.

In this paper we describe the sector-cluster concept that simplifies user mobility and tracking. We propose a prediction based on-demand routing protocol that is built on Location Aided Routing (LAR)[1], called the Prediction based LAR (P-LAR). We introduce a method of node mobility prediction to reduce the search space required for route discovery, which in turn reduces the route discovery overhead. In Section II we describe related work with the proposed approach developed in Section III. In Section V we present our simulation results supported by our analytical results in Section IV and the conclusions in Section VI.

II. RELATED WORK

As ad hoc routing protocols need to adapt quickly to frequent and unpredictable topology changes conventional routing strategies are inefficient. The amount of routing related traffic is to be minimised and so frequent update of routing tables is wasteful of the limited wireless bandwidth [2]. While a lot of research has addressed the routing problem in ad hoc networks [3] not all algorithms take into account the physical location of a destination node. Location information can be used for directional routing not possible with basic flooding algorithms and results in reduction of routing overhead.

Geographic Addressing and Routing (GeoCast) [4] forwards routing messages to all nodes in a specified geographic area rather than to node addresses. Making use of location information a destination address is expressed as a point (in terms of latitude and longitude), a circle (with centre and radius) or a polygon (as a list of points). Where the destination is a circle or a polygon every node in the area will receive the message. The Distance Routing Effect Algorithm for Mobility (DREAM) [5] is a proactive routing protocol using location information aimed at providing distributed, loop-free, multipath routing. Realising the principles of distance effect (i.e., the greater the distance between two nodes the slower their relative mobility appears) and mobility rate (i.e., the faster a nodes moves more frequent is the required position update) it minimizes routing overhead. Greedy Perimeter Stateless Routing (GPSR)[6] uses only neighbour location information in forwarding data packets. It requires only a small amount of per-node routing state, has a low routing message complexity, and works best for dense wireless networks. Beacon messages are broadcast at each node to inform neighbours of its position which results in minimized one-hop-only topology information at each node.

Location Trace Aided Routing (LOTAR) [7] is based on an on-demand scheme and utilizes location information to aid in routing. It is suited to support real time traffic, as it is able to keep low packet loss rates even in a high mobility environment. Location information is used to limit the search area, predict the route lifetime and handover a flow to a better route. Location Aided Routing (LAR) defines the Expected Zone and the Request Zone to aid in route discovery and limit the flooding area. The expected zone of a node D with respect to a node S is the region in which node S expects to find node D at a time t_1. The request zone is defined so that a node forwards the route request only if it belongs to the request zone. The request zone must include the expected zone to ensure that the request would reach the destination. In some cases when S is outside the expected zone it is necessary to include areas outside the expected zone. S and D are both required to belong to the request zone. If a path cannot be found within a predefined time period then the entire network space is included in the following route request. The probability of finding a path increases as the request area increases. However the route discovery overhead also increases with the size of the request zone. In LAR scheme 1, the request zone is defined as the smallest rectangle that contains the current location of the source and the expected zone of the destination such that the sides of the rectangle are parallel to the co-ordinate axes. The source node determines the four corners of the rectangle and includes their co-ordinates in the route request message. A receiving node is thus able to determine if it is in the request zone and forwards or discards the route request accordingly. In LAR scheme 2, the request zone is not specified explicitly by the route request message. The source node S includes two pieces of information in its route request-the distance of S from the destination D ($DIST_S$), and the positional co-ordinates (X_d, Y_d) of node D. When a node I receives the route request it is able to calculate its distance from the destination and forwards the message only if its position is closer or not much farther from D than node S. The route message is forwarded in a node J provided it is closer to D than the forwarding node to J.

III. PREDICTION BASED LOCATION AIDED ROUTING

In this section we introduce the Prediction based Location Area Routing that combines location information with a mobility prediction technique to reduce the search space for route construction. The mobility prediction technique employs a sector-cluster concept that gives a discrete approximation to the user location. This also eliminates the need for continuous location updates. The rationale for the method is that maximum information of the location (present & future) of the destination node can reduce the area of the request zone to a minimum. This method of routing is best suited for cluster based ad hoc networks.

A. *Preliminaries*

1) *Sector-Cluster Concept:*

In cluster based ad hoc networks we seek to define the location of the user with respect to its position with that of the cluster head. The cluster head has complete knowledge of each of its member nodes. The region of the cluster is divided into sectors as in Fig. 1. We introduce two types of sectors depending on whether or not the sector is adjacent to a neighbouring cluster. C-type cluster-sectors C_1, C_2, C_3 have as neighbours, clusters that are accessible through their cluster gateway nodes. It is only from C-type clusters that cluster change is possible. Nodes in S-type cluster-sectors S_1, S_2, S_3 are not candidates for cluster change as there are no adjacent clusters present. In between two sectors we also introduce the boundary sectors. Boundary sectors aid in prediction of sector changes within the cluster and is the region between adjacent clusters where the probability of sector change is high. The width of the boundary sector can be user dependant with highly mobile users having a wider boundary sector than less mobile users. We further define the optimal region (dotted circle) as the region of the cluster in which the nodes are in the closest proximity to each other. The purpose of this region is to eliminate the involvement of far lying nodes in the first stages of route construction. The optimal region is expanded until a route is successfully constructed. There is also defined the Hi-CC region-which is the region where there is a high probability of cluster change. The Hi-CC region is the area of the cluster (black outer circle) where the node is tracked for cluster change prediction purposes. By predicting cluster changes routes can be reconstructed prior to existing route failure.

2) *Cell-Sector Numbering scheme for user movement prediction:*

In cellular networks it is possible to predict regular user movements based on a mobility history base (MHB). This is not feasible in an ad hoc network as the movements of users are variant and dependant on the current purpose of the network. They do not follow regular day-to-day paths. While it is possible to obtain accurate tracking by making the prediction process highly complex the proposed method is computation non-intensive and introduces minimal amounts of additional traffic on the wireless link. The cluster sector-numbering scheme is able to predict the next cluster change depending on the user's current cluster and direction of travel in the Hi-CC region. Prediction is cluster-sector-wise and not physical location-wise as the need is only to predict the next connection point of the mobile user. The Cluster-Sector numbering scheme is as in Fig. 2. The numbering scheme is only for prediction purposes and sits on top of any other cluster numbering scheme that may be in use.

The user's current cluster is always identified as the reference cluster 0, i.e., if the user moves from cluster 0 in the figure to cluster 5 then cluster 5 becomes cluster 0 for mobility prediction purposes. Each sector of the resident cluster is then identified using 0i|aj where '0' is the reference cluster and 'a' denotes the neighbouring cluster to which the user can handoff to from this particular sector of the reference cluster. Re-referencing of a neighbour cluster is only done if the distance from the original reference cluster sector to the present resident sector is at the least 2 cluster-sector crossings. The system is robust enough to handle oscillating users between two sectors of neighbouring clusters without any re-referencing.

B. *Determination of routing request Area for route discovery*

Maximum information about the movement of the user restricts the size of the request area to a minimum. We make use of the expected zone & request zone concepts of Location Aided Routing but our definition of these zones is quite different and based on the sectorized cluster structure. We define the request zone making use of a futuristic definition of the user locations from a network level.

The Expected Zone is defined as the current cluster of the user and the next predicted cluster sector if applicable. The Request Zone includes the sector(s) of the expected zone to ensure that a path is constructed even if the user has changed its position of network connection. The request zone determines which of the nodes in the network forward the route request thus reducing the number of routing request messages generated. There is a however a trade-off between the size of the request zone and the probability of finding a routing path to the destination. We explain the working of the proposed method in the three scenarios outlined below.

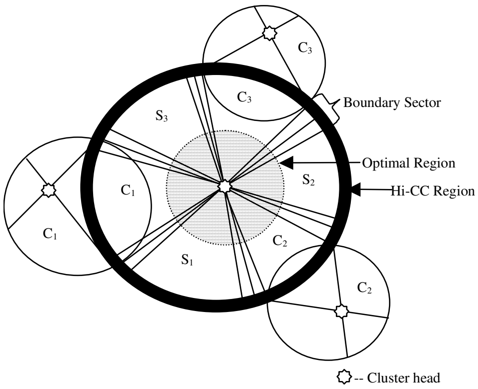


Fig. 1. The cluster-sector concept with S-type and C-type clusters, optimal region and boundary sectors.

$$\text{Cluster } C \in \{C_i \mid i = 1,2...n\}, \text{where,}$$
$$\text{n is the no.of adjacent clusters.}$$
$$\text{Each } C_i \text{is a C - type cluster - sector.}$$
$$\text{Cluster } C \in \{S_j \mid j = 1,2...N-n\}, \text{where,}$$
$$S_J \text{ is a S - type cluster - sector \&}$$
$$\text{N is the total no.of sectors in the cluster}$$

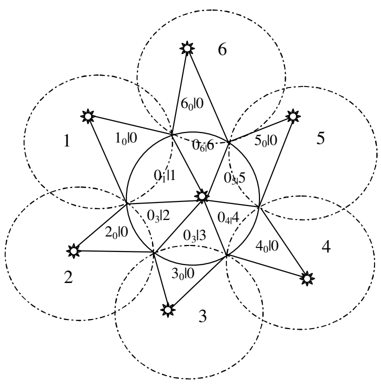


Fig. 2. The cluster-sector numbering scheme for mobility prediction.

Definition :
Each sector of cluster 0 takes a value $0_i \mid i$
where $i = 1,2,...n$
n = no.of adjacent clusters of 0
An adjacent sector takes the value $(n \bmod i + n)_0 \mid 0$

1) *Source and destination nodes are within the same cluster:*

As the source and destination nodes are located within the same cluster the request zone can be restricted to the cluster-sectors within their cluster of membership depending on whether or not the destination is a candidate for cluster change. If the user is a candidate for cluster change, the cluster sector of the neighbouring cluster is included in the request zone as well. If the destination node is located in one of the boundary sectors and is a candidate for sector change then both adjacent cluster-sectors are included in the request zone. The request zone is determined as in Fig. 3.

In order to construct a route from S to D we include the sectors of residence of S and D and the optimal region denoted by r (i) of the cluster as the initial route request area. If a route request reply is not received within a certain time period then we increase the optimal region of the cluster indicated by r (j). The optimal region is increased until a route request reply message is received. It is only in sectors of S (sector 2) and D (sector 5) that routes are forwarded outside the optimal region circle. All nodes within and without the optimal circle employ the distance rule in that a route request is forwarded only if the node is closer in distance to the destination node than the previous node (non forwarding nodes u & m).

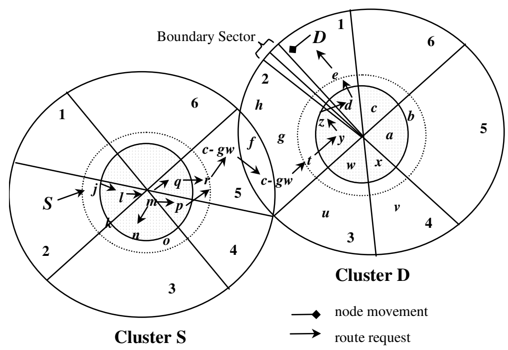


Fig. 4. Determination of routing request area when S and D are in neighbouring clusters.

As the destination D is mobile it is possible that it might be in a location different to the one it was when a route request was generated. If the node was in a boundary sector then the route request area includes both current and predicted sectors. We include both sectors to accommodate possibly oscillating destinations. If the destination node is out of the boundary sector prior to the route discovery procedures being completed the route request area is reduced to include only the new sector of residence.

2) *Source and destination nodes are in neighbouring clusters:*

If the source and destination nodes are in neighbouring clusters then the Request Zone will include sectors of both the clusters. The determination of the request zone must take into account the number of routing packets generated as well as the mobility of the destination and cluster gateway nodes. The request zone is determined as in Fig. 4.

The route discovery proceeds along very similar lines to the previous scenario. However since there are two clusters involved the cluster gateway (*c-gw*) nodes are employed for communication between the neighbouring clusters. In cluster of S the route request message is forwarded from S in sector 2 through the optimal circle on to the *c-gw* node in sector 5. The forwarding nodes employ the distance rule with respect to the *c-gw* node and not with respect to the final destination D. The cluster head of each cluster only has knowledge of the nodes in its cluster and hence forwarding with respect to the position of D is not efficient, as it would involve inter-cluster-head messaging. From the *c-gw* in sector 5 of cluster S the route request is forwarded to *c-gw* of cluster D. Forwarding of the route request within cluster D then reduces to the previous scenario.

3) *Source and destination nodes are in non-neighbouring cluster:*

When the source and destination nodes are in non-neighbouring clusters a route needs to be constructed that traverses the intermediate clusters to the destination node. The request zone would include the sectors of all intermediate clusters. The request zone is determined as in Fig. 5.

In cluster S and all intermediate clusters the routing request is forwarded in the optimal regions that have been defined. The distance rule for all the forwarding nodes in cluster S and cluster 2 is applied with respect to their respective *c-gw* nodes. In cluster S the route request is forwarded from S in sector 2 through the optimal circle to *c-gw1* in sector 5. The cluster gateway node of S forwards the request on to *c-gw2* of cluster 2. In cluster 2 the request is forwarded through the optimal circle on to *c-gw3*, which forwards the request onto *c-gw4* of cluster D. In cluster D the request is forwarded as in the first scenario.

C. Route Reconstruction

Due to the dynamic nature of the ad hoc topology there is a high likelihood of nodes moving out of range of each other thus causing a route to break. The destination node D to which a route has been constructed might move out of range of the last hop node. We employ mobility prediction to set up an alternative route before the existing route fails thus minimising the time for route reconstruction. Routes are only partially reconstructed from an anchor node that is identified depending on the sector that D is moving into. In most cases there does not arise a need to initiate the route discovery mechanism direct from the source node again. When D enters a boundary sector it initiates a sector_change message that is forwarded along its path to S. The sector change message contains the predicted sector that D is moving into. Forwarding of the message is stopped if it reaches a node that is in the same sector as the predicted sector of change. This node then acts as the anchor node for route reconstruction and sends a route_reconstruct_request message again employing the distance and optimal region rules of route discovery. Once the request message is received at D it sends a route_reconstruct_reply. The route_reconstruct_request is retransmitted if a route_reconstruct_reply is not received within a timeout period. The timeout period can be made quite short to ensure that a reply is received from D at the moment of its arrival in the new sector. There does exist a trade-off between the latency in reconstructing the route and message overhead that is generated. Optimisation of the width of the boundary sectors is essential to ensure that sector change messages are not generated too early while at the same time giving the anchor node enough time to ensure that a route can be reconstructed with minimum latency. In Fig. 6 node p acts the anchor node and the route is reconstructed in sector 6 of the cluster as D changes it sector of residence from sector 5 to sector 6.

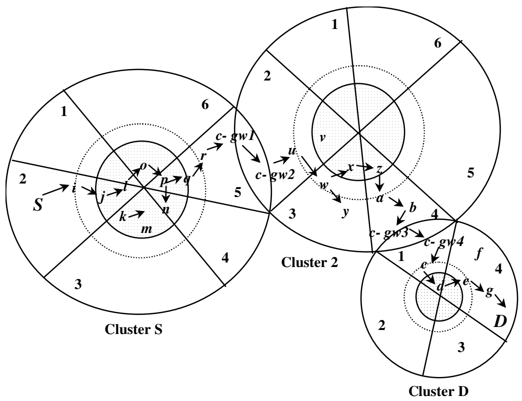


Fig. 5. Determination of routing request area when S and D are in non-neighbouring clusters.

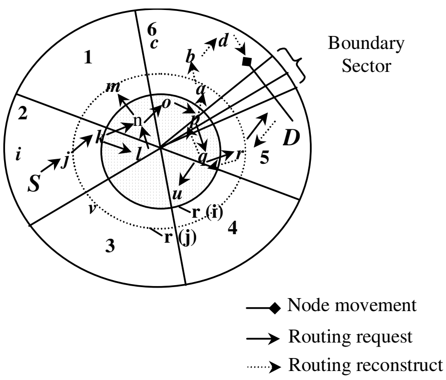


Fig. 6. Route reconstruction with mobility prediction.

IV. Mathematical Results

To calculate the probability that a node will become a contributor to the routing traffic we make use of the jumping rabbit problem [8]. The problem defines a rabbit running across a field executing a "random walk" with jumps of varying magnitudes and falling into a ditch. Our point of interest is the probability that the rabbit falls into the ditch after a certain number of moves. Problems for which a certain critical level exists can be defined by the jumping rabbit problem. In our case, for a node to become a contributor to the routing traffic it should be found in the routing request zone. We therefore define our critical level as a position threshold – the node entering the route request zone. We restrict our comparisons to the LAR-distance based method and the LAR-prediction based method.

It is typical of ad hoc networks that nodes tend to enter and leave clusters thus causing the size of the cluster to vary. This property is to be incorporated into the probability function. We assume that at all times the movement of the cluster head is within the cluster. Also the size of the cluster is bound to vary as users enter and leave the cluster. The probability that a mobile node will contribute to the routing traffic is the probability that the node is found in the routing area as the cluster size changes.

From previous work in [9], the probability that the routing area starts at a position m_{ra} after X_{cs} cluster size changes can be given as,

$$P_{X_{cs}}(m_{ra}) = \frac{X_{cs}!\, p_{cs}^{1/2(X_{cs}+m_{ra})}(1-p_{cs})^{1/2(X_{cs}-m_{ra})}}{[1/2(X_{cs}+m_{ra})]!\,[1/2(X_{cs}-m_{ra})]!} \qquad (1)$$

where,

p_{cs} is the prob. of cluster size increase.

Probability that a node is found in the routing area at a position *m* after *x* moves so as to be involved in the route setting up process is given as in (2).

$$P_x(m) = \sum_{m=0}^{S_{route\_area}} p.\frac{(X-1)!\; p^{1/2(X-1+m)}(1-p)^{1/2(X-1-m)}}{[1/2(X-1+m)]!\,[1/2(X-1-m)]!} . \frac{X_{cs}!\, p_{cs}^{1/2(X_{cs}+m_{ra})}(1-p_{cs})^{1/2(X_{cs}-m_{ra})}}{[1/2(X_{cs}+m_{ra})]!\,[1/2(X_{cs}-m_{ra})]!} \qquad (2)$$

such that, $m \sim m_{ra} \le S_{route\_area}$

where S_{route_area} is the size of routing area.

A. LAR - distance based method

Using (2) we can calculate the probability that 'n' nodes are involved in the routing process. This would mean that in an N node network there are n nodes located at distances $x_i, x_{i+1} \ldots x_n$ such that, $x_i \ge x_{i+1} \ldots \ge x_n$. The probability that n nodes are involved in the routing process is given as,

$$P(n) = \sum_{m=x_i}^{x_n} P_x(m) \qquad (3)$$

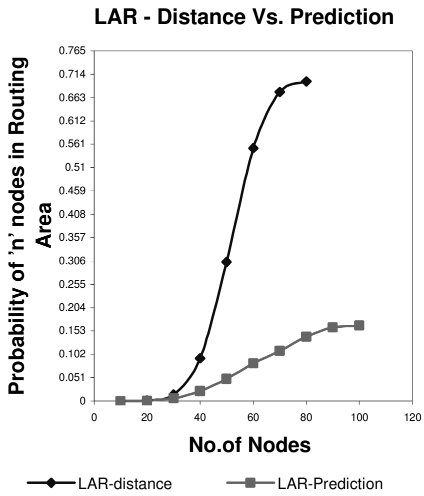


Fig. 7. Analysis of LAR-distance and LAR-prediction.

B. LAR - prediction based method

For simplicity we assume that in the routing area each node is able to communicate to a node in the neighbouring cluster sector. If the routing area is defined as *p* positions then the probability of finding a node at positions $1, 2 \ldots p$ i.e. in the routing area is given as,

$$P(ra) = \sum_{m=1}^{p} P_x(m) \qquad (4)$$

The probability of finding *n* nodes in the routing area is then given as,

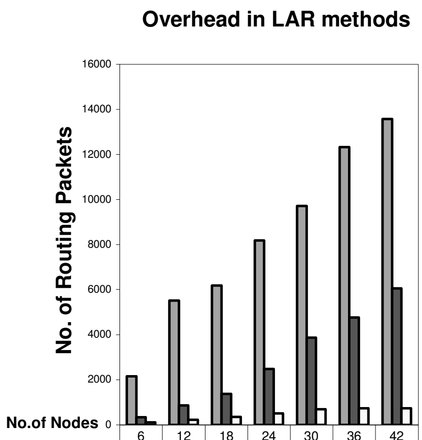
$$P(n) = \binom{N}{n} P(ra)^n . (1 - P(ra))^{N-n} \qquad (5)$$

Plotting (3) & (5) in Fig. 7 it is seen that while the probability for the distance based method is found to increase as the number of nodes in the network increases the probability plot for the prediction method tends to increasingly plateau.

V. Simulation Results

Simulation of the proposed model was done making use of OPNET modeler 8.0. The number of nodes in the network was varied to be 6,12,18,24,30&42 for different simulation runs. For each network configuration twenty runs with different random number seeds were completed. The movement of the cluster head and the mobile nodes was assumed to be an unrestricted random walk. There was no restriction of movement in velocity, direction or step size. The simulation environment allowed users to move in any arbitrary direction (between 0 & 2π) and vary their speed in random intervals. We assume that each node moves constantly without pausing in between steps. For each of the network configuration we simulated the Flooding algorithm, LAR scheme 2 (LAR-distance) and our prediction based LAR. For each simulation run, a source and a destination are chosen randomly and the source initiates a routing request to a different destination every five seconds. Each node makes multiple moves during the simulation time period with the direction of motion chosen randomly. All the nodes have the same transmission range and to maintain uniformity of results the same mobility image is retained for a specific topology across all three protocols.

The comparative results of the number of routing packets generated in the network by each of the routing strategies are presented in Fig. 8. The simple strategy of specifying the entire network space to be the request zone as in flooding algorithms results in a large number of routing packets being generated as all nodes of the network that are within transmission range are involved in the route discovery process. As the overhead of route discovery increases exponentially with the number of nodes in the network flooding algorithms are not best suited for ad hoc networks as the amount of available bandwidth is low. Flooding algorithms are not scalable. Also most nodes are working under strict energy constraints and reception, processing and transmission of packets all result in energy loss. The LAR scheme 2 does offer some improvements on the flooding algorithm. The number of packets generated is considerably lower for each of the network topologies. For the 42-node network the LAR scheme 2 has roughly 50% less route discovery overhead than the flooding algorithm. As the number of forwarding nodes is limited with respect to their distance from the destination node there is a reduction in the overhead. However because of the scheme reducing to the flooding algorithm in the event of a route reply not being received within a specified time period it suffers from scalability issues similar to the flooding algorithm. It can be seen that the number of packets generated though lower still increases exponentially with the number of nodes in the network. It is quite significant that if the number of nodes were very large then due to the latency of route discovery the overhead generated would be comparable to the flooding algorithm.



No.of Nodes	6	12	18	24	30	36	42
LAR/ Flooding	2160	5504	6189	8192	9720	12328	13564
LAR/ Distance	328	870	1360	2478	3868	4769	6047
LAR/ Prediction	113	223	364	510	684	725	735

Fig. 8. Simulation results of LAR methods w.r.t number of packets generated during the route discovery phase. The prediction based LAR is found to perform best amongst the three algorithms. The overhead of route discovery is minimal compared to the flooding and LAR scheme 2 algorithms. This is due to the request zone being minimal and the incremental increase of the request zone rather than a one step reduction to the flooding algorithm. As the network size increases the routing overhead is found to increase gradually moving into a plateau. It can be seen from the plot that the overhead for 30-node, 36-node and 42-node networks are 684,725 & 735 routing packets respectively. Thus making the prediction based LAR scalable. The reason for this efficiency can be explained from the cluster structure that is employed. As the forwarding nodes are to be found in the optimal circle the number of nodes found in the circle would not vary much as the number of nodes in the network increases. This can be seen by the results of the mathematical model in section IV.

VI. CONCLUSION

Maximum knowledge of the user's movement minimizes the request zone. We have introduced in this paper Prediction based Location Aided Routing (P-LAR) that makes use of mobility prediction to reduce further the search space required for route discovery in mobile ad hoc networks. The mobility prediction technique employed works on a sectorized cluster structure and is able to predict the next position of the user from a network level. From our simulation results we can see that the proposed method reduces the routing overhead quite considerably in comparison to the flooding and LAR scheme 2 algorithms. The algorithm is found to be scalable, as the routing overhead tends to plateau after a gradual increase. We have also presented a method of route reconstruction that makes use of location information to sets up alternate routes proactively.

REFERENCES

[1] Y. B. Ko and N. H. Vaidya, "Location Aided Routing (LAR) in Mobile Ad Hoc networks," Proceedings of ACM/IEEE MOBICOM '98, Oct.1998, pp. 66-75.

[2] X. Hong, K. Xu and M. Gerla, "Scalable Routing Protocols for Mobile Ad Hoc Networks," IEEE Network, July/August 2002, pp. 11-21.

[3] E. M. Royer, C-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," IEEE Personal Communications, April 1999, pp. 46-55.

[4] J. C. Navas and T. Imielinski, "Geographic Addressing and Routing," Proc. 3rd ACM/IEEE Intn'l. Conf. Mobile Comp. Net. , Sept. 1997.

[5] S. Basagni et al., " A Distance Routing Effect Algorithm for Mobility (DREAM)," ACM/IEEE Int'l. Conf. Mobile Comp. Net. 1998, pp. 76-84.

[6] B. Karp and H. T. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," Proceedings of MobiCom2000, pp. 243-54.

[7] Kui Wu and Janelle Harms, ``Location Trace Aided Routing in Mobile Ad Hoc Networks," IEEE International Conference on Computer Communications and Networks ICCCN 2000, Las Vegas, Nevada USA, October 2000.

[8] Syski R., "Random Processes A First Look" 2nd edition, Revised and Expanded, Marcel Dekker, Inc., New York, 1989.

[9] R. Chellappa, A. Jennings, N. Shenoy, "A Comparative Study of Mobility Prediction in Fixed Wireless Networks and Mobile Ad hoc Networks", to appear in proceedings of IEEE Int'l Conference on Communications ICC 2003, Anchorage, Alaska USA, May 2003.