Location Management in Heterogeneous VANETs: A Mobility Aware Server Selection Method

(Invited Paper)

Seyedali Hosseininezhad and Victor C.M. Leung

Electrical and Computer Engineering Department
University of British Columbia, Vancouver BC V6T1Z4, Canada
{seyedali,vleung}@ece.ubc.ca

Abstract. Heterogeneous wireless networks are capable of providing customers with better services while service providers can offer more applications to more customers with lower costs. Many services require the support of location management functions, which still need further innovations to become viable in multi-hop vehicular ad-hoc networks (VANETs). High mobility in vehicular networks causes conventional location management systems to overuse the bandwidth and induce extra handovers to clients trying to synchronize themselves to location servers. We provide a routing algorithm for transactions between location servers and mobile nodes. We assume location servers are vehicles equipped with at least one long range and one short range radio interfaces, whereas regular nodes (clients) only need a short range radio interface. The primary goal of our design is to minimize handovers between location servers while limiting the delays of location updates. Taking advantage of vehicle mobility, we propose a mobility-aware server selection scheme and show that it can reduce the number of handovers and yet avoid large delays during location updates. We model the proposed scheme in NS-2 and apply vehicular mobility patterns generated with SUMO for urban and highway scenarios for performance evaluations. We show that proposed scheme significantly lowers the costs of signaling and rate of server handovers by increasing the connection lifetime between clients and servers.

Keywords: Heterogeneous Networks, Mobility Aware Routing, Location Management, VANET.

1 Introduction

Vehicular ad-hoc networks (VANETs) are emerging as one of the most important practical applications of mobile ad-hoc networks (MANETs). As the demand for pervasive computing is increasing, location management becomes one of the most important modules in vehicular networking. Multimedia streams, news broadcasting, entertainment and other applications which require Internet connectivity, peer to peer applications, local advertisements, vehicle pooling and local

J. Zheng, D. Simplot-Ryl, V.C.M. Leung (Eds.): ADHOCNETS 2010, LNICST 49, pp. 64–81, 2010.

cab services are some examples of the broad range of feasible applications when vehicles are equipped with positioning and communication equipments [1, 2].

1.1 Location Management in MANETs

Several protocols have been designed to handle mobility of nodes [3–10]. Location servers are responsible for handling geographical location information of nodes in the vicinity and provide them to others when needed.Different categories of location management have been classified: Flooding based location management is considered as the most straight forward method for passing location information. Due to high redundancy overhead, researchers have strived to decrease unnecessary packet relays. The hypothesis of methods like DREAM [4] is that relative locations of closer nodes are changing faster compared to nodes far away from each other. Therefore location updates are being sent to close location servers more frequently than others. Quorum based location management is another category which is based on assuring a rendezvous between queries and updates. A localized quorum based location service is proposed in [5]. In this method location of every node is dispersed horizontally and vertically. As the authors have stated, this method is proper for networks without a significant relative motion. VANETs with high relative speeds are improper environment for this class of location management systems.

The GLS method [6] for distributed location service management divides the area into different degrees of grids in a way that in every grid around the node there is a fixed number of servers that collect location information about that node. As grids grow larger, the probability of a server being chosen for other nodes decreases. This method is not very flexible for highly variant environments like VANETs. Hierarchical methods [9, 10] for server allocation are highly scalable because the rates of location updates are reduced for servers in higher levels. However in this work we would prefer not to consider these methods because of the following reasons.

We consider private and public transit vehicles in this work. Private vehicles have intermittent availability and may not be trusted by other vehicles; therefore they are not candidates for selection as location servers. We propose to host location servers in public transit vehicles. Because not every node in a VANET can be selected as location server, and the density of location server candidates will not be very high in most road traffic scenarios, application of hierarchical methods is not justified. To deal with the sparsity of potential server nodes, we shall also propose to utilize long-range radio communications to interconnect these nodes.

In high mobility networks such as VANETs, keeping track of location information would in general be a huge overhead on the network if location information is saved in all location servers. Therefore having records of every node in network while nodes are rapidly changing their locations can be performed better if we are able to save these information in a specific set of nodes. Moreover, to reduce the effect of mobility on precision, we are focusing on hop by hop packet relaying rather that finding a deterministic route from clients to servers. We will look at location management as a service to vehicles which is offered by some specific nodes in the area. Therefore we need a service discovery mechanism to find service providers. The main contribution of this paper is the proposal of an efficient routing mechanism for this category of networks with high mobility profiles.

In [11] a new metric has been introduced for routing assessment: Longevity of a route. Specially important in MANETs, changes in the path can cause extra signaling and delay overhead during route reconstruction. Therefore they have proposed association between two nodes to measure route longevity. It is assumed that a link is reliable if association between two ends is higher than a certain threshold. Moreover, a routing algorithm is proposed in [12], in which route selection is based on a hybrid criterion of route lifetime and path length. Route lifetime is measured by a definition of link affinity which is calculated based on received signal strength. Since in practical propagation channels the signal strength is not constant over time, RABR can make wrong decisions especially in an urban environment. However, it is possible to utilize the concept of link lifetime as a decision factor in routing but it needs a new measurement tailored for variable conditions of vehicular networks.

Many different wireless access technologies can be employed in VANETs. Short-range technologies include wireless local area networks (WLANs) and its variant called Dedicated Short-range Radio Communications (DSRC) targeting specifically vehicular communications. Long-range technologies include cellular networks and wireless metropolitan area networks such as WiMAX. VANETs employing short-range radio access face problems in area coverage and fast hand-offs between nodes. Because of high mobility speed, rate of hand-offs in the network becomes a bottleneck in location registration and updates. Heterogeneity can come to the rescue for services like location management. Using long-range wireless access as a higher layer of communications, we can interconnect location servers together as a logical mesh network. We can assume that a connected graph of location servers can exchange signaling messages through this logical mesh network. We shall base our work on utilizing available long-range wireless connections to facilitate location management in VANETs.

1.2 Service Discovery Inspired by Field Theory

Lenders et. al. in [13] defined an approach for efficient and robust service discovery. This concept is similar to anycast routing, which is supported in IPV6 [14]. In anycast routing, an address is associated with more than one interfaces that belong to distinct nodes that are similar in nature. As it is preferable for clients to get service from the nearest among several potential servers, use of anycasting would allow the desired server to be reached easily.

From electromagnetic field theory, the point potential of a spot is related to its distance to the maxima potential charge. In wireless networks, the most commonly used definition for distance is based on hop count; nonetheless geographical distance is also applicable. In [13], hop count is considered as the distance between nodes:

$$\varphi(n) = \sum_{i=1}^{N} \frac{Q_i}{dist(n, n_i)} \tag{1}$$

where Q_i is the potential assigned to server i and $\varphi(n)$ is the total received potential by node n from all servers. The amount of potential assigned to each server could be a factor of their capacity or quality of service (QoS) metrics.

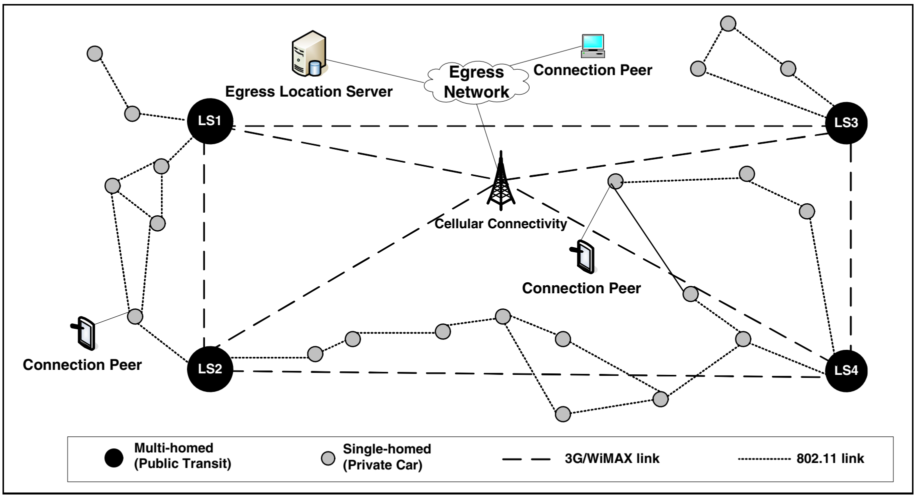

Fig. 1. Location Management over Heterogeneous Architecture

2 Location Management over Heterogeneous Networks – The Architecture

Fig. 1 depicts a heterogeneous network architecture with partial Internet connectivity. In this system, heterogeneous nodes are connected to each other and edge gateways using their long-range wireless access capability.

The requirements and assumptions in aforementioned architecture are:

1. All vehicles are considered to have a mechanism to extract their own geographical location, e.g. using a onboard global positioning system (GPS) receiver.
2. All nodes are equipped with at least one short-range radio (e.g. 802.11a,b,g,p).
3. Some special nodes are equipped with all valid short range communication interfaces and one long range communication interface (e.g. WiMAX) and function as location servers. These nodes are interconnected to each other in a logical mesh network to exchange their location records, and to stationary gateways for Internet access.

4. In consideration of valuable licensed-spectrum used for long-range wireless access, the use of long-range radio should be minimized. Therefore it is our goal to reduce the numbers of queries between servers and server handovers for vehicles.
5. Location queries and updates should not be propagated more than a certain number of hops.
6. Server advertisements should not be rebroadcasted more than a certain limit.

In one scenario of this architecture, a public transportation system provides a wireless Internet relay service inside an urban area. Public transit vehicles are equipped with multiple radios and they are tasked to provide connectivity and related services to other vehicles. They utilize their long-range radios to relay local data network traffic to stationary gateways and to provide a location management service to vehicles in their vicinity by exchanging location information with other location servers.

To advertise location service of servers and receive updates and queries from vehicles, we propose a service discovery mechanism to find routes to location servers in the area with the best matching mobility pattern. We will evaluate the effect of this service discovery method with different scenarios of urban and highway mobility.

3 Mobility Aware Service Selection and Packet Relay

Vehicular mobility patterns (urban or highway) generally follow roadways with probabilistically change of directions at intersections. We assume that every vehicle responsibly sends its location information to a location server. This location information can be used by other vehicles or service providers to present location based services. When a vehicle and its location server move away from each other and the distance grows more than a certain hop distance, path delay and high link breakage probability make their interactions ineffective. Therefore the client has to hand off from the old server to a better server in terms of delay, robustness and lifetime. Every hand-off between two location servers is comprised of several 'server to server' and 'server to client' signalling interactions. However, server to server interaction are more expensive because they use licensed spectrum to communicate.

Based on expectations and assumptions in the architecture of Fig. 1, if we want to use the approach explained as field theory before, clients should send their location management packets toward other relays or servers in vicinity who have the highest potential. Signalling for a location update comprises of a primary phase of registration between client and server. After the registration, client is able to synchronize the server by sending periodic or event triggered updates. If a client is unable to send updates to designated server, a new registration with an available server is required. Based on the our desired architecture and location management procedure, the field theoretic method reviewed in Section 1.2 has the following deficiencies that should be addressed:

1. The measure for distance between nodes is unrealistic, since mobility pattern is not considered in server selection. In our case the relative speed between a client and its server defines the connectivity lifetime and we prefer to choose a server that has a higher connectivity lifetime as long as the path delay is less than a certain limit.
2. The server selection is stateless. Service discovery would lead to a set of choices for each relay to forward the packet. However there is no guarantee that a packet will be relayed to the same server which the former packet is sent to. It is desirable for a client to send location updates to a server that it has already registered in. It means that if a client selects a server with highest potential as its location server, all the relay nodes should be notified to relay the packet from that client toward the same server. Consequently a server hand-over does not happen unless the delay threshold is exceeded or disconnection occurs.

By modifying the service discovery method proposed in [13], we are going to define a location management method that minimizes hand-overs, which is applicable for geographical and topological location managements.

3.1 Reliability vs. Distance

Hop distance is a simple and effective criterion for route selection, but in cases with high mobility this measure is very unstable. To avoid this problem we propose to use link stability and usability (also known as reliability) as the route selection criteria instead. Let's denote the set of links in the chosen path between s and d as $P(s,d)$. We want to account for reliability of each link $l \in P(s,d)$ and choose a path with highest aggregated reliability. Reliability of a link is directly related to the estimated link lifetime. However calculation of reliability includes error and an unmeasured factor of future alternative connections. For instance, a weak link could be replaced in future by a new relay node which is not present at the moment. Due to this factor, it is not rational to underrate a path by considering the reliability of the weakest link in path as decision factor. In the other hand, we cannot rely on arithmetic average because strong links in the path would cause overestimation of path reliability.

Since we want the factor to tend toward the most realistic reliability value of the path (to mitigate the impact of links with excellent reliability in calculation of total path reliability), we desire to have the smallest average value as the measuring factor. Instead of using arithmetic mean, we would use harmonic mean to calculate the reliability factor. Harmonic means tend toward the smallest values compared to geometric and arithmetic means. Hence, with r_l being reliability of link l and the number of hops $|P(s,d)|= n$ we have:

$$\frac{1}{\frac{1}{n}\sum_{l \in P(s,d))} \frac{1}{r_l}} \leq \sqrt[n]{\prod_{l \in P(s,d)} r_l} \leq \frac{1}{n} \sum_{l \in P(s,d)} r_l$$

To apply the value of reliability in routing decision, we define the distance factor in (1) as:

$$D(s,d) = \sum_{l \in P(s,d)} \frac{1}{reliability(l)} \tag{2}$$

In every calculation period, each node will predict the locations of current neighbors and based on estimated pathloss exponent of environment and foreseen mobility patterns, calculate the link lifetime. Path prediction and lifetime estimation are two major ongoing research topics and they are explained more in 3.2. Intuitively, a longer link lifetime leads to a more reliable path between nodes and location servers. Notwithstanding, due to high error rates in prediction mechanisms [15], we cannot rely solely on measures of one link. Therefore we will define *reliability factor* for all (*node,server*) pairs to show how reliable the *node* could be to relay packets toward the *server*. Based on the assumption that a higher node density can make the route more reliable, we define the *reliability factor* as the probability of a packet being successfully relayed by a *node* to another which is closer to *server*.

3.2 Reliability Measurement

We define link reliability for two neighbors as the estimated expected remaining time of connectivity between the pair. To calculate the link reliability factor we assume that each node will listen to data packets and beacons sent by its neighbors. Using sampled signal characteristics and location information, receiver predicts how link condition will change in future. Therefore, to extract reliability factors, nodes need to have knowledge about future variations in link connectivity. In [16] a method for estimating link residual time and link stability has been proposed. In this method, after denoising and classification of the radio signal strength indications (RSSI) from neighbors, future lifetime is estimated. In [15], Euclidean distance information is utilized to estimate future trajectory. It seems that by using relative mobility between nodes and digital map information together, future estimations can become more precise [17]. A method to calculate the probability of turns in road intersections is proposed in [18], based on the theory that turning options that lead to more destinations in shorter times are more popular than those which lead to local areas or take more time to reach destinations. In [19], mobility behavior of nodes is used to classify their transportation mode. Moreover, using particle filter method they estimate parameters in a Bayesian network for path selection decisions. By learning these parameters, they try to estimate future velocities and turning selections.

To calculate the link reliability in a path toward a server, every node will calculate the cumulative probability of connectivity to next hop for each server. The next hop is defined as any node in communication range that has a loop-free path to the server. In practice every node can put a short hash value of its unique address in forwarded advertisement to avoid considering routes which are originated from itself. Hashing can reduce the length of a string to a few bits and yet avoid duplicate indexes with a high probability.

We assume that a link connectivity prediction method can provide a process consisting of connectivity probabilities during a prediction period. Suppose for a specific environment and mobility pattern, a prediction method is able to predict future motions and channel states for n time units. Moreover, we can extract average percentage of error in lifetime prediction (which can be empirically found) as $E(\hat{e})$. This error extraction could be enforced to system as a-priori known measure or it could be re-adjusted based on observations from past (By comparing prediction and real condition after it happened). Therefore $P_{l(i)}(t)$ is the probability of link $l(i)$ being connected from t_0 (now) until $t_0 + t$ and is equal to:

$$P_{l(i)}(t) = (1 - E(\hat{e}))P(\hat{L}_i, t) \tag{3}$$

where $P(\hat{L}_i, t)$ represents the link condition (alive/dead) which is calculated using the desired mobility prediction method. As mentioned before, each mobility prediction and link classification method has a distinct estimation capability in different environments. Therefore values of $E(\hat{e})$ and $P(\hat{L}_i, t)$ are highly dependent on the method of prediction being used. We will evaluate our method in Section 4 based on a simple linear prediction, but as long as a prediction method yields a prediction of link lifetime and an estimate of the error, it can be integrated in our approach.

Having the estimate of link lifetime for all links in the path, the probability of having an undisrupted path from node k toward the server S for the next t time units (complement of the probability of no link being capable of relaying packets from k to S) is:

$$C_S^k(t) = 1 - \prod_{l(i) \in H^k(S)} \big(1 - P_{l(i)}(t)\big) \tag{4}$$

where $H^k(S)$ is set of links between k and its neighbors that have a loop free path to server S. We use the cumulative distribution of $C^k(t)$ (for $t = 0 \cdots t_{max}$ with t_{max} equal to maximum duration of predictability) as a factor which shows how reliable the node k is to pass the packets toward server S:

$$reliability(k) = cdf_{t0}^{t_{max}}(C_S^k) = \sum_{t=0}^{t_{max}} C_S^k(t_0 + t) \tag{5}$$

We need to extract the reliability of a node for all servers being discovered. To avoid extra calculations, we define a maximum hop threshold for acceptable potentials received from neighbors. Intuitively it is obvious that information regarding far away servers are not of any interest because of extra relay overhead and delay.

Finally we define the distance between client c and server S as:

$$D(c, S) = \sum_{k \in P(c,S)} \frac{1}{cdf_{t0}^{t_{max}}(C_S^k)} \tag{6}$$

Notice that as hop number increases, the distance is affected and the chance of being chosen as best path decreases. This definition of distance would result in such way that nomadic mobility patterns lead to higher potentials and connection between vehicles with opposite directions and/or sparse connectivity conditions cause less potential dispersion.

3.3 Potential Assignment for Path Construction and Server Selection

Using (2) as the distance measure for (1), every node can receive a potential from servers based on the relative mobility and link condition of all nodes in the path from that server to the node. As described in Algorithm 1 every node advertises all valid server information received from neighbors to adjacent nodes. After receiving these advertisements, the node sets the current potential received from a server to the highest received value. These values are valid up to a certain time after last advertisement. Whenever the node wants to send or relay a packet toward a server, it will choose the server with the highest potential. This policy leads to selection of a server which is having the best known mobility correlation with the transmitter.

Algorithm 1. Potential Assignment

Input $servers\_advertisements[\,]$
for each $servers\_info$ **in** $servers\_advertisements$
  $L \leftarrow servers\_info.source$
  predict_link_condition(L)
  $servers[L] \leftarrow servers\_info$
  for each S in $servers\_info$
    $rel\_factor \leftarrow \text{reliability}(S)$
    $D(S.id) = \frac{S.pot.original}{\frac{S.pot.original}{S.pot.received} + \frac{1}{rel\_factor}}$
    $pot[S.id] \leftarrow \max\left(pot[S.id], \frac{S.pot.original}{D(S.id)}\right)$
    $next\_hop[S.id] \leftarrow \arg\max_{l \in neighbors} l.servers[S.id].potential$
  end for
end for

3.4 Location Update

To choose a new server for location updates, each node will select the server with the highest accessible potential. After choosing the best server, location updates are sent toward it using the neighbor who has the largest potential from that particular server. Using this approach instead of [13] we can make sure that location update packets will not face misroute to other servers which are not moving in favorable directions. Decision to hand over to another server is performed by a client when the hop distance to the current server has exceeded a certain threshold. Since it is assumed that location update messages are not in a high priority class and their packet size is reasonably small, packet relay in short range wireless would still be more favorable compared to expensive long range network. Anyhow, decision for when to hand off is still open to users and they can choose between prompt updates and lower cost. We will discuss about this trade-off in the next section. Packet relay to a chosen server is also done very easily by comparing received potential of the server from current alive links and therefore source routing is not necessary.

3.5 Location Query

To find the location of another user, a requester would send a location query to the best available server at the moment (the one with the highest available potential). The packet relay mechanism would be similar to that for location updates. After receiving the query, the server looks up in its local database to see if it has up-to-date location entry. Otherwise it will send a query to its neighbors using long range wireless. Since we assume that long range media would lead to a connected graph topology, queries will have answer from one of the servers and this answer will reach the original server.

4 Performance Evaluations

To evaluate performance of our proposed framework, we have modeled the system using the NS-2 network simulator [20]. We have added a new service discovery agent over the currently implemented network stack and added our logic as an application agent. Using application agent, we can use any routing algorithm for packet routing mechanism. We have tried our protocol on several test scenarios. These scenarios are based on realistic vehicular traffic generated by SUMO network mobility generator [21]. This microscopic vehicle traffic generator is able to create mobility patterns based on defined traffic flows. The trace generated by SUMO is a mobility log for vehicles moving based on road and traffic regulations. We can import different maps to SUMO to generate different test cases. We have imported several maps with different key features. The first imported map is a 10 Km long highway with 2 lanes in each direction. Two kinds of vehicles have been considered to commute on the road: private vehicles with short-range radios and public transit vehicles equipped with long-range and short-range radios. The two categories of vehicles have different characteristics in speed limit, acceleration and deceleration. The second scenario is realistic urban area extracted from actual street maps. These maps are extracted from free maps available in OpenStreetMap [22]. After adding traffic lights to map, we have used SUMO to generate traffic information for 10000 seconds. The procedure of map extraction and simulation has been shown in Fig. 2. After generation of mobility traces, they are fed into NS-2 as mobility scenario and simulation is performed by NS-2. Since we need prediction in our method and it is not performed in NS-2, we do the simulation twice; The first run is done to extract exact location of every

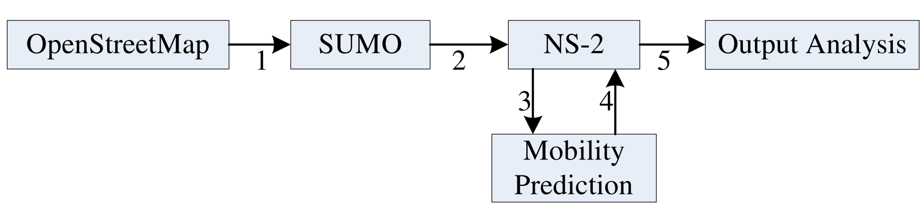

Fig. 2. Simulation Procedure

vehicle during the simulation. Then we use this data in the next run as a precise prediction of future mobility patterns in network. To make the prediction more realistic, we add noise to location information. Since prediction precision is strongly dependent on prediction mechanism, we use one of the simplest predictions: In every second, each node predicts $x(t+1) = v + x(t)$, where $x(t)$ is the location of node in time t $(0 \leq t \leq t_{max})$. t_{max} is the maximum time that prediction can be reasonably valid. We will set t_{max} to a value which:

$$\begin{gathered} E\big(Pr(C(t))Pr(\hat{C(t)}) + Pr(\overline{C(t)})Pr(\overline{\hat{C(t)}})\big) > threshold \\ 1 < t < t_{max} \end{gathered} \tag{7}$$

This is the sum of expected probability for having a true guess, whether positive or negative, on having a connection. This probability should be more than a certain threshold. To find this value we run the simulation and calculate the predicted and actual locations in future. We consider a link as active if its RSSI is more than a threshold. Since measurement of RSSI is impacted by environmental clutters, it is impractical to deterministically define the link connectivity threshold. So we use the propagation model in [23]:

$$P_r(d) = P_t - PL(d) = P_t - (\overline{PL}(d) + X_\delta) \tag{8}$$

$\overline{PL}(d)$ is the log-distance path loss from transmitter to receiver and X_δ is a zero-mean Gaussian distributed random shadowing effect with standard deviation δ. Values of path loss exponent and δ are usually extracted from empirical data. We have borrowed these values from the experiment done by Otto et. al. in [24]. Finally, the probability of RSSI being more than γ (dBm) in distance d(m) is:

$$Pr[P_r(d) > \gamma(dBm)] = Q(\frac{\gamma - \overline{P_r(d)}}{\sigma}) \tag{9}$$

Fig. 3 shows the estimated error in aforementioned prediction method. Results show that in highway scenario prediction is performing close to reality and connection condition after 40 seconds is predicted correctly with a 70% probability. However, in suburban scenarios and downtown areas, nondeterministic stops and turning probabilities causes prediction error to grow. For downtown scenario we find that predictions are 50% successful only for 20 seconds ahead. In suburban areas with less stops and turns compared to downtown, it is up to 35 seconds. We apply these errors in calculating path reliability factor for each scenario.

To avoid excessive delay caused by late hand-offs we have to set a threshold for maximum allowable hop distance between nodes and servers. The trade-off is between location update cost (which is related to amount of relayed data and type of media used for it) and end-to-end delay.

$$X = \begin{cases} d(i,S)\,(f_u.LU + f_q(LQ + LR)) & d(i,S) < thr \\ \\ K * (N-1) * SYN + & \\ d(i,S_{new})(f_u.LU + f_q(LQ+LR)) & otherwise \end{cases} \tag{10}$$

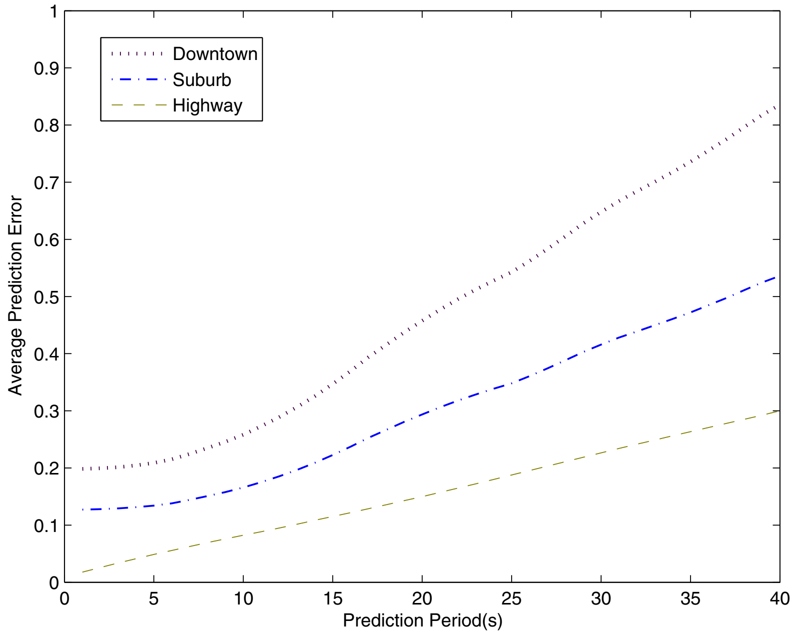

Fig. 3. Linear prediction average error for three scenarios

Table 1. Variables definition in (10)

$d(i, S)$	Hop distance between i and Server S_i
f_u	Frequency of location update messages
f_q	Frequency of location query messages
LU	Size of location update message
LQ	Size of location query message
LR	Size of location reply message
K	Usage cost/Kb for long range network
N	Number of location servers
SYN	Size of synchronization message

To calculate the proper threshold, we set our objective to minimize $Cost * Delay$ for location update packets. Using (10) as the cost function and by knowing $d(i, S')$, the distance between a node and its second best server with eligible hop distance, every node can calculate the threshold as follows:

$$thr = d(i, S') + \sqrt{\frac{K * (N-1) * SYN}{f_u.LU + f_q(LQ + LR)}.d(i, S')} \quad (11)$$

Here we assumed that delay is only based on hop distance and did not consider the delay caused by collisions.

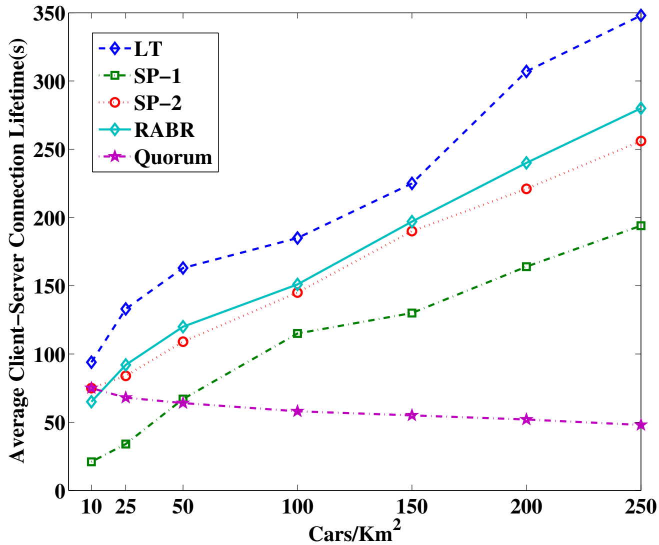

Fig. 4. Average Client-Server Connection Lifetime (Downtown)

After finding the estimation errors, we run the algorithm based on these estimation properties. We try to establish connection between nodes with heterogeneous connectivity (as servers) and nodes with single wireless network.

We compare client-server path length and traffic cost compared to three other methods: [13] (shortest path is the metric for route selection), [12] (affinity based on signal-to-noise ratio) and [5] (quorum based method with column based advertisement and row based query). Fig. 4 shows the average lifetime of connections between mobile nodes and location servers in downtown mobility pattern. Hereafter we will refer to our method as Life Time based method (LT). SP-1 represents shortest path anycasting based on [13]. In SP-2 we use the same method as SP-1 but whenever a server is selected for a node as a location server, it will remain chosen as long as their distance is less than maximum hop distance. Results show that using lifetime as the distance metric has led to significant connection lifetime improvement specially for higher densities.

In affinity based method, SNR is considered as the measuring factor for decision making. Therefore for downtown areas with highly volatile SNR conditions RABR can not perform much better than SP-2. Since in quorum based method, chosen servers are changed rapidly after any change in topology, connection lifetimes are not comparable to other methods.

Fig. 5 shows the same measures as Fig. 4 but for highway scenarios. Results show 57% overal improvement in connection lifetime compared to SP-2. Especially in lower vehicle densities, our proposed method achieves more improvements compared to the shortest path method because of the steady mobility of vehicles which leads to higher lifetime if paths are selected from vehicles moving along the same direction. RABR performs better in highway scenarios duo

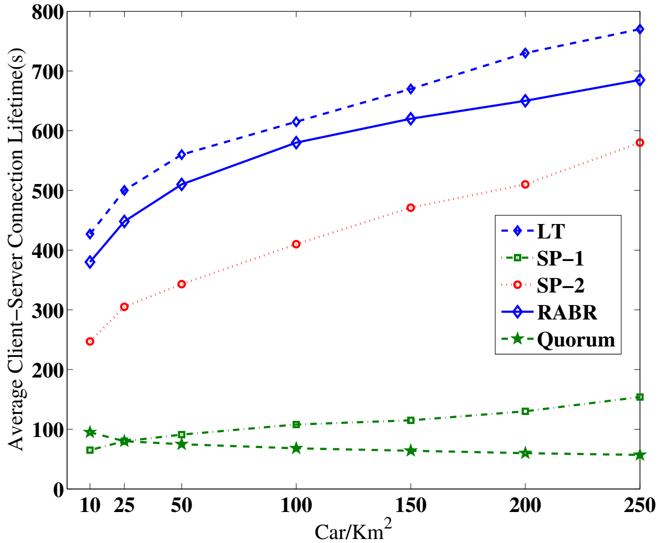

Fig. 5. Average Client-Server Connection Lifetime (Highway)

to less perturbations in SNR. However, quorum based method is following the same behavior in downtown. Fig. 6 compares the overhead caused by location update packets. Quorum based method uses several location servers, hence location updates become costly. Moreover, as mobility and interactions between nodes increase, the overhead of quorum based method increases drastically. We have assumed that the quorum based method uses WLAN and WiMax based on availability with no preferences. To compare the proportion of WLAN usage, we assumed that parameter K in (10) is equal to 100 (every transmission on WiMax is 100 time more expensive than WLAN). We can see that usage of WiMax in low traffic densities is significantly low and as mobility patterns grow more dynamic, the difference between LT and other three methods become noticeable.

Fig. 7 depicts the normalized total cost of location management for different values of K. Since the cost of RABR is very close to SP and the cost of the quorum based method is significantly higher than other methods, we have only compared the cost of SP vs. LT. As one can observe, for $K \geq 100$ our method outperforms the shortest path method. K can be interpreted as a priority or preference parameter and could be tuned based on the trade-off between delay and cost efficiency. If providers prefer faster and more precise location updates, they can decrease K. In contrast, for better efficiencies (e.g. for less location sensitive applications) higher values will lead to better spectrum conservation.

In Fig. 8 we have shown the average delay experienced by signaling packets (Location Update, Location Query and Responses) for downtown scenario. It is important to mention that the quorum based method uses more than one

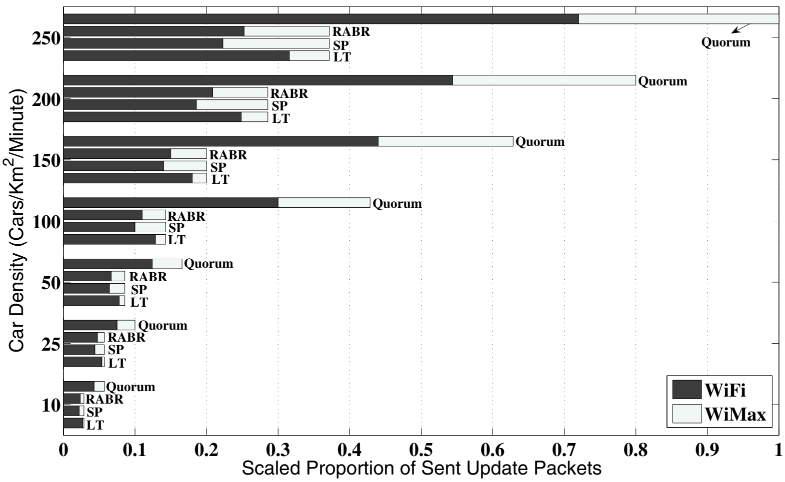

Fig. 6. Normalized overhead of update packets being sent over WLAN and WiMax (K=100)

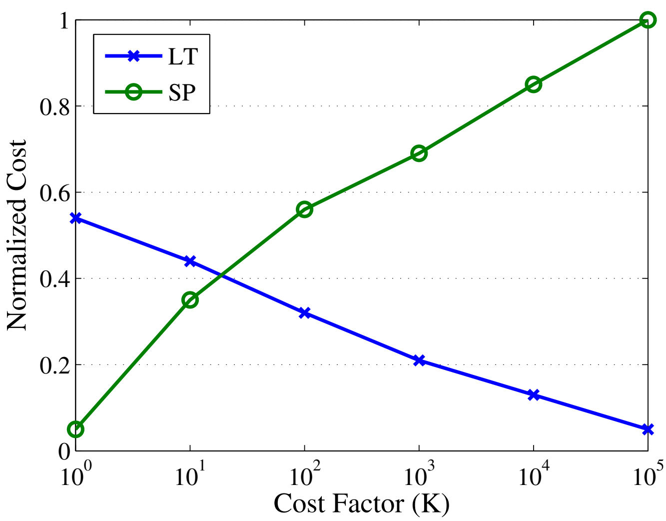

Fig. 7. Signaling cost based on cost factor (K) in (10)

location management server and always every immediate neighbor in the same row of the node is acting as a location server. As a result signaling delays for location updates and responses are very low. However, when it comes to location query, signaling delay is relatively higher than updates. Our method is always performing better than RABR in terms of delays but compared to SP, it suffers

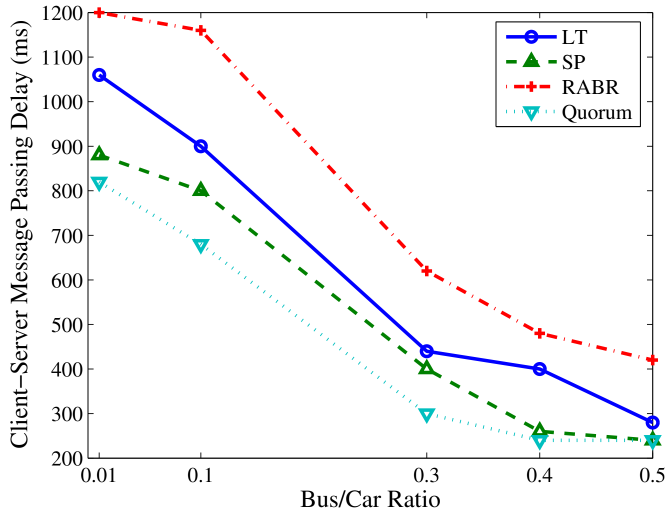

Fig. 8. Average Client-Server Message Passing Delay in Downtown (Only Query/Response messages are considered for Quorum Based Method)

a 20% increase in delay. As the number of location servers (in this case described as buses) increases, the overall signaling delay for all methods decreases. We can see that if half of the vehicles in our system could act as location servers, delay would have become as low as SP and quorum based methods.

5 Conclusion

Location management is a critical part for vehicular ad-hoc networks. In this paper we have assumed that some of the mobile nodes in vehicular network are equipped with heterogeneous wireless connectivity. These vehicles are able to act as location servers for other vehicles and cooperate using their long-range radio. We have proposed a new server selection and packet relay mechanism that minimizes the rate of server handovers by relaying location update packets toward the server that has the lowest possibility of disconnection. This is done by proposing a new definition of distance. The proposed method has been evaluated by extensive computer simulations. The results show significant improvements in client-server connection lifetimes. Higher connection lifetimes lower the costs of handovers, which require the use of long-range communications to update the record for the client at all the servers. We have provided a tuning factor which can be used for decision making based on tolerable delay and cost. The comparison has been made against three methods: associativity based routing, shortest path selection and quorum based location management. Results show that in scenarios with high mobility our method achieves the lowest costs and acceptable delays compared to other three methods.

References

1. C2c-cc manifesto. Technical report, Car 2 Car Communication Consortium (2007)
2. Willke, T., Tientrakool, P., Maxemchuk, N.: A survey of inter-vehicle communication protocols and their applications. IEEE Communications Surveys and Tutorials 11, 3–20 (2009)
3. Friedman, R., Kliot, G.: Location services in wireless ad hoc and hybrid networks: A survey. Technical report, Department of Computer Science,Technion (2006)
4. Basagni, S., Chlamtac, I., Syrotiuk, V., Woodward, B.: A distance routing effect algorithm for mobility (DREAM). In: Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 76–84 (1998)
5. Liu, D., Stojmenovic, I., Jia, X.: A scalable quorum based location service in ad hoc and sensor networks. In: 2006 IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS), pp. 489–492 (2006)
6. Li, J., Jannotti, J., De Couto, D.S., Karger, D.R., Morris, R.: Scalable location service for geographic ad hoc routing. In: Proceedings of the Annual International Conference on Mobile Computing and Networking (MOBICOM), pp. 120–130 (2000)
7. Saleet, H., Langar, R., Basir, O., Boutaba, R.: Proposal and analysis of region-based location service management protocol for vanets. In: IEEE Global Telecommunications Conference (GLOBECOM), pp. 491–496 (2008)
8. Dikaiakos, M.D., Florides, A., Nadeem, T., Iftode, L.: Location-aware services over vehicular ad-hoc networks using car-to-car communication. IEEE Journal on Selected Areas in Communications 25, 1590–1602 (2007)
9. Kieß, W., Füßler, H., Widmer, J., Mauve, M.: Hierarchical location service for mobile ad-hoc networks. ACM SIGMOBILE Mobile Computing and Communications Review (MC2R) 8, 47–58 (2004)
10. Ahmed, S., Karmakar, G., Kamruzzaman, J.: Hierarchical adaptive location service protocol for mobile ad hoc network. In: Proceedings of the 2009 IEEE Conference on Wireless Communications & Networking Conference, The Institute of Electrical and Electronics Engineers Inc., pp. 2932–2937 (2009)
11. Toh, C.: Associativity-based routing for ad hoc mobile networks. Wireless Personal Communications 4, 103–139 (1997)
12. Agarwal, S., Ahuja, A., Singh, J., Shorey, R.: Route-lifetime assessment based routing (RABR) protocol for mobile ad-hoc networks. In: IEEE International Conference on Communications. Citeseer, vol. 3, pp. 1697–1701 (2000)
13. Lenders, V., May, M., Plattner, B.: Service discovery in mobile ad hoc networks: A field theoretic approach. Pervasive and Mobile Computing 1, 343–370 (2005)
14. Hinden, R., Deering, S.: IP Version 6 Addressing Architecture, RFC 2373 (1998), `http://ietfreport.isoc.org/rfc/rfc2373.txt`
15. Haas, Z.J., Hua, E.Y.: Residual link lifetime prediction with limited information input in mobile ad hoc networks. In: Proceedings of IEEE INFOCOM, pp. 26–30 (2008)
16. Sofra, N., Leung, K.K.: Link classification and residual time estimation through adaptive modeling for vanets. In: IEEE VTC, pp. 1 –5 (2009)
17. Menouarand, H., Lenardi, M., Filali, F.: Movement prediction-based routing (mopr) concept for position-based routing in vehicular networks. In: IEEE Vehicular Technology Conference, pp. 2101–2105 (2007)

18. Krumm, J.: Where will they turn: predicting turn proportions at intersections. (Personal and Ubiquitous Computing, Springer)
19. Patterson, D., Liao, L., Fox, D., Kautz, H.: Inferring high-level behavior from low-level sensors. LNCS, pp. 73–89 (2003)
20. The network simulator ns-2, http://www.isi.edu/nsnam/ns/
21. Simulation of urban mobility, http://sourceforge.net/apps/mediawiki/sumo/
22. Open street map, http://www.openstreetmap.org/
23. Rappaport, T.: Wireless communications: principles and practice. Prentice Hall PTR, Upper Saddle River (2001)
24. Otto, J.S., Bustamante, F.E., Berry, R.A.: Down the block and around the corner: The impact of radio propagation on inter-vehicle wireless communication. In: Proceedings - International Conference on Distributed Computing Systems, Montreal, QC, Canada, pp. 605–614 (2009)