SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks 2010; **00**:1–13
DOI: 10.1002/sec

RESEARCH ARTICLE

Location-Preserved Contention Based Routing in VANETs

Q. Yang[1]*, A. Lim[2], X.J. Ruan[3], X. Qin[2], and D. Kim[2]

[1]Dept. of Computer Science, Montana State University, Bozeman, MT 59717, USA
[2]Dept. of Computer Science and Software Engineering, Auburn University, Auburn, AL 36849, USA
[3]Dept. of Computer Science, West Chester University of Pennsylvania, West Chester, PA 19383, USA

ABSTRACT

Location privacy protection in VANETs considers preserving two types of information: the locations and identifications of users. However, existing solutions which either replace identifications by pseudonyms or hide locations in areas cannot be directly applied to geographic routing protocols since they degrade network performance. To address this issue, we proposed a location-preserved contention (LPC) based routing protocol, in which greedy forwarding is achieved using dummy DODs (distance to the destination) information instead of users' true locations. Unlike the contention based forwarding (CBF) protocol, the number of duplicated responses in LPC can be reduced by adjusting the parameter α, which is a timer scaling factor. To quantify the efficiency of location privacy protection, an entropy based analytical method is proposed. LPC is compared with existing routing and location privacy protection protocols in simulations. Results show that LPC provides 11.7% better network performance and a higher level of location privacy protection than the second best protocol. Copyright © 2010 John Wiley & Sons, Ltd.

KEYWORDS

location privacy, privacy protection, contention based routing, vehicular networks

***Correspondence**

Qing Yang, EPS 357, Montana State University, Bozeman, MT 59717, USA. E-mail: qing.yang@cs.montana.edu

Received . . .

1. INTRODUCTION

The proliferation of wireless technologies has spurred a new era of vehicular ad hoc networks (VANETs) which enables drivers to avoid accidents by exchanging safety-related information among one another. However, a driver has to share certain information such as location and identification, with its neighbors, to achieve efficient routing. Given such information being released, location privacy becomes a big concern in vehicular networks [1, 2, 3].

Location privacy in VANETs can be defined as follows: a certain user's current or historical locations should not be revealed to unauthorized people. This definition includes the protections of two types of information: identification and location of a user. By using randomly changed pseudonyms in [4], a user's identification can be preserved. On the other hand, [5, 6, 7] hide a user into an area to protect its location information. However, all the above-mentioned protocols sacrifice network performance to achieve location privacy protections [8]. A challenging issue is: how to protection location privacy without compromising network performance.

We propose a location-preserved contention (LPC) based routing protocol which protects user's location privacy by 1) completely hiding identification and 2) utilizing dummies to substitute true locations. To avoid periodically broadcasting the information of identification and location [9, 10, 11], we adopt the contention based forwarding scheme in which only nodes participating in routing share information between each others. For those nodes on routing paths, dummy distance to destination (DOD) and pseudonyms are used to protect every packet forwarder's location and receiver's identification, respectively. The dummy DOD should be carefully chosen so that a user's true location is preserved and the geographic routing goals are achieved. To measure how well the location privacy is protected in LPC, we proposed an entropy based analytical model. It gives the entropy (or hardness) required for an adversary to attack the location privacy of all users in networks.

The network and location privacy protection performance of LPC are compared with existing solutions, e.g. GPSR [9] and CBF [10]. Simulations results show that LPC provides a better network performance than other

routing protocols. Besides, we also discover LPC achieves a higher level of location privacy protection than others.

The rest of this paper is organized as follows. Section 2 discusses currently available geographic routing and location privacy protection protocols in VANETs. In Section 3, we describe the motivation and threat model of location privacy protection in VANETs. In Section 4 and 5, the detailed design of LPC and the entropy based analytical model are presented, respectively. Section 6 gives the simulation results and Section 7 concludes our work.

2. RELATED WORK

By exploiting location information, geographic routing provides [12, 13] superior scalability and robustness compared to traditional routing protocols in VANETs. However, location information shared among nodes will compromise user's location privacy. Although there are several location privacy protection protocols [6, 7, 5], they cannot be directly applied to existing routing protocols since they would degrade network performance. To the best of our knowledge, we are the first to apply location privacy protection to geographic routing protocols without sacrificing network performance.

2.1. Geographic Routing in VANET

Since GPS devices are becoming standard components in future vehicles, geographic routing has become more popular in VANETs such as [9, 10, 12, 14]. Most existing geographic routing protocols are based on Greedy Perimeter Stateless Routing (GPSR) [9]. In GPSR, a node selects from its neighbors the next hop which is closest to the destination. This process is called greedy forwarding which is the most commonly used geographic routing. To achieve routing goals, GPSR requires vehicles to periodically broadcast their locations and identifications to their one-hop neighbors. Therefore, the location privacy of vehicles are not protected in GPSR.

There is another type of geographic routing [10, 15, 16, 17] in which every next-hop node is elected through competitions. All these protocols are based on the contention based forwarding (CBF) [10]. In CBF, the current packet forwarder broadcasts a request-to-forward (RTF) packet with its location to its neighbors. After receiving the RTF, every neighbor computes its distance advance (i.e. the distance that it will advance) to the destination. Then, it sets a timer according to its distance advance. The node which is the closest to the destination has the smallest timer. So it will be the first to time out and be elected as the next hop by sending a clear-to-forward (CTF) packet with its identification. Finally, the data packet is forwarded to the next hop. In CBF-based routing protocols, the identification and location information of those involved in routing are not protected.

Some geographic routing protocols do not need node's exact location information instead of relative location information [18]. However, relative location computation is needed which takes time. Such overhead can not be ignored in VANETs as the relative position of vehicles need to be updated frequently.

2.2. Location Privacy Protection in VANETs

The location privacy issue in VANETs is brought out in [19, 20, 21, 22, 23, 24]. To address this issue, several work are proposed [25, 26, 27]. However, none of them provides a detailed analysis about how their protocols will affect network performance.

Location privacy protection can be achieved in two different ways: hiding the information of who sends the data or where the data come from. In the first case, pseudonyms of nodes are periodically changed, so an adversary cannot attack users' identities, and thus location privacy is protected [4, **?**, 23, 25, 28, 29, 30, 31]. However, changing identifiers has detrimental effects on routing efficiency and increases packet loss as shown in [8].

In the second case, a node's location is hidden in an area or a set of dummies. For instance, a node is considered residing in a rectangle or circle in which there are other $k-1$ nodes [5]. Similarly, dummy-based location privacy protection algorithms are proposed in [6, 7]. In [6], a node randomly generates several false position data with one that contains the true position. In [7], the authors hide a user's real location in a set of dummy positions which are deliberately generated according to either virtual grids or circles. Because an adversary cannot distinguish a node from other $k-1$ nodes [5] or from dummies [6, 7], location privacy is protected. However, none of these approaches can be applied directly to geographic routing protocols because network performance will be drastically decreased.

In summary, to achieve an efficient geographic routing in VANETs, both identification and location are important. The approaches of location privacy protection proposed in application layers cannot be directly applied in routing protocols. Thefore, it is non-trivial to explore the possibility of applying location privacy protection without sacrificing network performance. In LPC, we address the location privacy issue in VANETs by 1) replacing a user's location with a dummy DOD during 2) generating pseudonyms to preserve a user's identification.

3. MOTIVATION AND THREAT MODEL

Geographic routing in VANETs can facilitate the release of vehicle's location information. Because location and identity data are shared among neighbors in geographic routing protocols [9, 10, 11], attackers can easily eavesdrop and track the movements of vehicles. This causes the leakage of a driver's privacy, e.g. a patient at an AIDS

Security Comm. Networks 2010; **00**:1–13 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/sec
Prepared using secauth.cls

testing clinic might not want his or her movements (or even evidence of a visit) revealed to others.

3.1. Adversary

The objective of location privacy attack is to create a service (e.g. database or website) which allows querying a user's historical location information. For example, the service can answer "where was vehicle DS5478 on Mon. around 11:00 AM". To achieve this goal, an adversary need to constantly collect tuples containing time instances, identifications, and the locations of nodes.

In reality, the adversary possibly builds up such database by passively listening to packets transmitted in geographic routing protocols, e.g. GPSR [9] which release both identification and location information of nodes. The adversary may be external, which installs its own wireless receivers along roads and passively eavesdrops on communication messages. For instance, by exploiting already deployed 802.11 infrastructures, it is possible to build a global adversary which eavesdrops on the entire network [32, 33]. The attackers can also be internal, which utilizes devices that are legitimate members in VANETs. These malicious nodes passively collects data transmitted among neighboring nodes. To build the above-mentioned database, huge amount of malicious node are needed. Therefore, it is cheaper and easier to implement the external adversary (infrastructures). In this article, we assume a global external adversary in VANETs.

To avoid location information being recorded by the adversary node, the proposed protocol LPC does not release any true location and ID data. However, if the adversary node is powerful enough, it can still identify or localize a node. The problem is: it is too expensive to enable the global adversary to have such ability. In the next section, we will discuss about how the adversary node identifies or localizes nodes in networks.

3.2. Fingerprinting and Localization

In this paper, although LPC preserves identifications in the network layer by using pseudonyms, the adversary can detect IDs from other layers. For example, network nodes can be identified by their physical (PHY) fingerprints [34, 35, 36]. Similarly, the adversary may use localization algorithms [37, 38] to find user's location data. However, a MAC address can be easily changed or cloned by software. Therefore, it is difficult for the adversary to distinguish faked MAC address from others.

It may be easier for the adversary node to identify a node by discovering wireless device fingerprints. Fingerprinting technique is based on the assumption that subtle differences of manufacturing and hardware components can create unique signaling characteristics in wireless devices. In [34], the authors present a technique to identify wireless nodes based on the unique analog signal characteristics. Although the results of fingerprinting wireless devices are significant, [34] requires expensive hardware. Considering a nodes's clock skew, a technique is presented in [35] which uses slight drifts in a device clock to identify a wireless device via its unique clock skew. Similarly, other fingerprinting methods [36, 39] are proposed to identify nodes by analyzing the characteristic of wireless radio signals.

According to [37], there are numerous localization algorithms for wireless networks. In most IEEE 802.11 networks, received signal strength indicator (RSSI) can be used to estimate the location of a mobile station (STA). If antenna arrays are installed on the adversary node, localizing a mobile node can be solved by angle of arrival (AOA) method [40]. Based on time of arrival (TOA) and time difference of arrival (TDOA), many other localization algorithms are proposed [38, 41].

Using specially designed hardware, the location of a mobile station can be accurately measured [38]. In [38], silicon and firmware modifications are made, experiment results show the accuracy of localization is significantly improved. Localization can also be made by software such as [41]. In [41], packet sequences and a listening node are used to measure the distance between a STA and an AP.

Currently, it is possible for an adversary to identify and localize nodes through advance technologies. However, these new techniques either requires special hardware or sophisticated software to process huge amount of data. In other words, it is expensive (almost impossible) to apply them to all adversary infrastructures in the networks. In Section 5, we model the difficulty of applying identification and localization functions to the adversary as D^I and D^L, respectively.

4. LOCATION-PRESERVED CONTENTION (LPC) BASED ROUTING

There are four major differences between LPC and contention based forwarding (CBF), which makes LPC a unique routing protocol. First, the greedy forwarding in LPC is achieved by using a sender's dummy DOD instead of its true location. Second, buffers are used to avoid duplicated RTF, CTF and data packets. Third, the number of duplicated responses in LPC is reduced by adjusting the parameter α. Fourth, LPC provides better location privacy protection than CBF.

4.1. Exchange of Control Messages

Before a data packet is transmitted, the current packet forwarder sends a request-to-forward (RTF) message. In this message, the sender provides a dummy DOD (not real DOD). Then, an elected next hop sends a pseudonym (not tru e ID) in its clear-to-forward (CTF) message. Finally, the data packet is transmitted from the current packet forwarder to the next hop. This process is different from the active selection in CBF (so-called CBF-AS) because LPC neither releases the forwarder's location nor the next hop's identification.

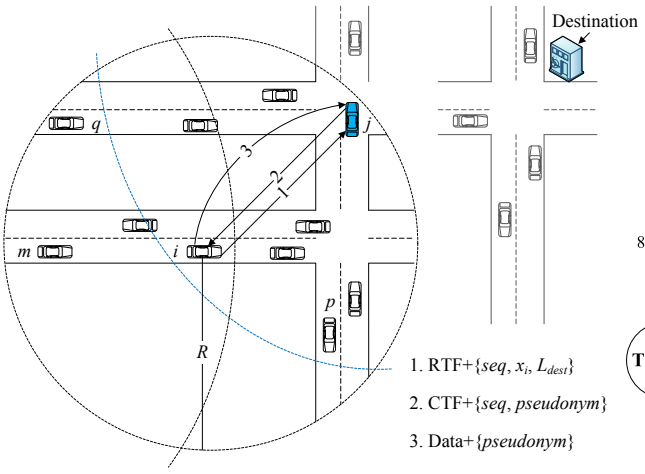

Figure 1. Dummy DOD based RTF/CTF exchange among vehicles

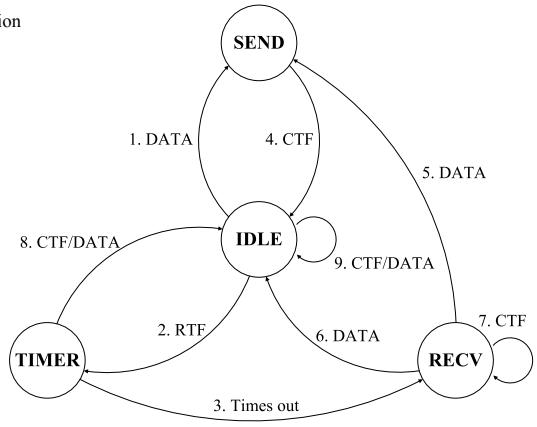

Figure 2. Finite state machine of LPC

As shown in Figure 1, suppose node i received (from node m) a packet with sequence number seq and destination L_{dest}. Before forwarding the packet, node i broadcasts a RTF message containing the tuple $< seq, x_i, L_{dest} >$ where x_i denotes node i's distance to destination (DOD). To illustrate the basic idea of greedy forwarding, we currently do not use the dummy DOD which will be introduced later.

When a neighbor (e.g. node k) receives the RTF, it checks if the packet was received before by comparing its seq. If there is a cache hit, node k drops the RTF because it is a duplicated request. Otherwise, node k saves this RTF and sets up a timer. The value of node k's timer is:

$$f(x_k) = T\left(1 - \frac{x_i - x_k}{R}\right) \quad (1)$$

where x_k is node k's DOD, T is the maximal one-hop forwarding delay, and R is the communication range. $x_i - x_k$ indicates the distance advance (to the destination) of node k compare to node i.

According to Equation 1, the value of the timer on each node is proportional to its DOD. Therefore, the timer on the node which is the closest to the destination will first time out. As shown in Figure 1, node j first times out and sends a CTF message including $< seq, pseudonym >$ where $pseudonym$ is an ID randomly chosen by j . When the CTF is received, node i immediately sends node j the data containing $< pseudonym >$. Then, node j checks the $pseudonym$ in the data. If it is the same as what it sent out previously, the packet is delivered; otherwise, the packet is dropped.

CTF and data can serve as suppression messages in LPC. For example, when the neighbors of node i and j receive the CTF from j, they immediately cancel their timers because a next hop (node j) has been elected. Since the CTF from j can only suppress its neighbors, duplicated CTF messages may be generated from node i's neighbors which are out of node j's communication range. That means duplicated CTFs may be received at node i. If node i sends the data before the second CTF is received, all its neighbors are suppressed by the data message. If another CTF is received before the data is sent, node i omits this CTF.

Although multiple CTF messages do not affect the next hop selection, they may cause network congestion and decrease network performance. The detailed analysis of duplicate responses in CBF and LPC will be discussed in Section 4.4. In Section 4.5, we will show how to reduce duplicate CTFs by setting an appropriate timer.

4.2. State Machine

The whole process of control message exchange can be better explained in a state machine transition diagram. As shown in Figure 2, when the system starts, all nodes are in the IDLE mode. Depending on what type of message is received, a node changes its mode as follows.

1. A node changes its mode from IDLE to SEND only if it intends to send new data to a destination. In this case, this node first sends a RTF message to its neighbors.

2. When the sender's neighbors receive this RTF, they enter the TIMER mode in which timers are set up according to their DODs.

3. The node which is closest to the destination first times out and sends a CTF message, i.e. it becomes the next hop. After sending the CTF, it goes into the RECV mode and wait for data packets.

4. When the sender receives the CTF, it unicasts the data to the next hop and returns to IDLE mode.

5. If the next hop is not the destination, it enters the SEND mode and keeps forwarding the data packet.

6. If the next hop is the destination, it delivers the message and comes to IDLE.

7. It is possible that the next hop receives a duplicated CTF from others, it simply drops the duplicated CTF.

8. Because other neighbors of the sender also receive the RTF message, they set up timers as shown in Step 2.

Security Comm. Networks 2010; **00**:1–13 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/sec
Prepared using secauth.cls

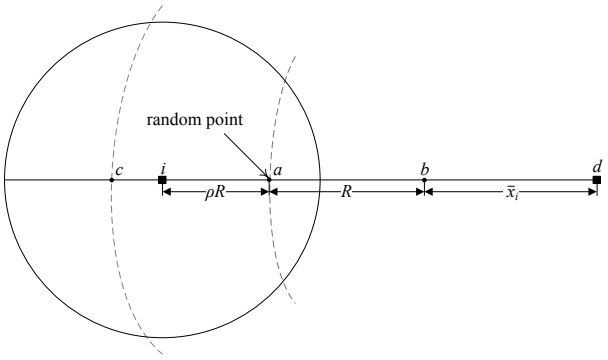

Figure 3. Dummy DOD selection on the current packet forwarder

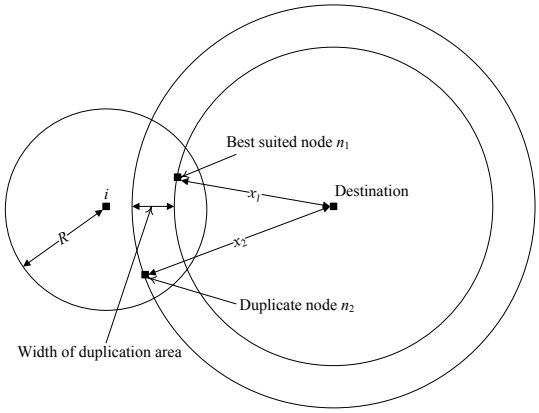

Figure 4. Duplicate responses and the width of duplication area

When they receive a CTF or data packet before their timers expire, they cancel the timers because a better next hop was selected. Therefore, they return to IDLE from TIMER mode.

9. When a node is in IDLE mode and receives a CTF or data packet, it will ignore those messages and remain in IDLE mode.

4.3. Dummy DOD Selection

The DOD x_i in a RTF message may release the current packet forwarder node i's real location. For example, the adversary can draw a circle centered at the destination with a radius of x_i. Node i must be on this circle. If node i has two other packets with different destinations, another two circles can be drawn. Then, the location of node i can be easily compute by triangulation algorithms.

To protect node i's location privacy, we use a dummy DOD $\bar{x}_i$ in RTFs, which hides node i's real location in an area. The selection of dummy DOD is shown in Figure 3. First, a point a is randomly chosen on the line between node i and the destination d if $ia < R$. Then, a point b which is R away from a can be found. Finally, the length of segment bd is considered the dummy DOD of node i. A dummy DOD is only used when node i is more than $2R$ away from the destination; otherwise, the dummy DOD will be zero, which means point b is located at the destination d. In summary, the dummy DOD $\bar{x}_i$ can be computed by the following equation:

$$\bar{x}_i = \begin{cases} x_i - (1+\rho)\cdot R, & x_i > 2R \\ 0, & x_i \leq 2R \end{cases} \qquad (2)$$

where x_i is node i's real DOD, ρ is a random number within $(0, 1)$. Since $\rho \in (0, 1)$, the difference between the real and dummy DODs is $(x_i - \bar{x}_i) \in (R, 2R)$. In other words, node i's location is hidden in the area enclosed by the two dashed lines as shown in Figure 3.

Because dummy DODs are used in RTFs, a neighbor receiving these RTFs will have problems in setting its timer. According to Equation 1, if we simply substitute x_i by $\bar{x}_i$, the computed runtime may be a negative value which is not acceptable. To avoid this issue, we change Equation 1 to:

$$f(x_k) = \begin{cases} T\left(\dfrac{x_k - \bar{x}_i}{2R}\right), & x_k - \bar{x}_i < 2R \\ \infty, & x_k - \bar{x}_i \geq 2R \end{cases} \qquad (3)$$

According to Equation 2, $x_k - \bar{x}_i$ is equal to $x_k - x_i + (1+\rho)\cdot R$. Since node i and k are neighbors, $x_k - x_i$ must be within $(-R, R)$, i.e. $x_k - \bar{x}_i \in (0, 3R)$. However, we are only interested in the nodes whose DODs satisfy the condition $x_k - \bar{x}_i \in (0, 2R)$. For those nodes having $x_k - \bar{x}_i \in (2R, 3R)$, their DODs must be larger than x_i because $x_k - x_i + (1+\rho)\cdot R \in (2R, 3R)$ yields $x_k - x_i \in (0, 2R)$. Since those nodes do not make any progress in forwarding packets to the destination, they ought to be omitted in the next-hop selection process. Therefore, those nodes will not have timers set up and participate in the routing procedure.

4.4. Analysis of Duplicate Messages

As we discussed in Section 4.2, a node sends a CTF to suppress the timers on others. However, it is possible that a neighbor times out right before it receives the CTF. In other words, if the difference between timers are too short, multiple CTF responses are generated. This problem occurs in both CBF-AS and LPC.

As shown in Figure 4, there is a best suited (next hop) node n_1 with a DOD x_1, and a duplicate node n_2 which is x_2 from the destination. With δ denoting the minimal time required for a successful suppression, the width of duplication are Δ is defined as

$$\Delta = \sup\{|x_2 - x_1| : |f(x_2) - f(x_1)| < \delta\} \qquad (4)$$

So, if x_1 is DOD of the best suited next hop for any particular packet, then duplicate CTF packets will be generated from all nodes j such that $x_j leq x_i + \Delta$; thus, the smaller is the Δ, the fewer is the number of duplicated

packets and hence the better is the network performance. According to the timer setting of CBF protocol, we have $\Delta = \frac{\delta R}{T}$.

4.5. Timer Setting Strategy

According to Equation 3, the timer on a node is a linearly function of its DOD. Therefore, the timer's difference is proportional to the DOD's difference. To achieve a larger timer's difference with the same DOD's difference, we need to change the timer setting function to:

$$f(x) = T\Big(\frac{x - \bar{x}_i}{2R}\Big)^{\alpha} \tag{5}$$

where x is the DOD of a receiver which receives a CTF containing the dummy DOD $\bar{x}_i$. $\alpha \in (0, 1)$ is a scaling factor which increases the value of $f(x)$ when it decreases. With the scaling factor α, the function $f(x)$ is still proportional (but not linearly) to the DOD x. For the same DOD's difference $\Delta = x_2 - x_1$, if x_1 and x_2 are closer to the 0 end, the timer's difference $f(x_2) - f(x_1)$ will be bigger than when they are nearby the $2R$ end.

Because duplicate messages are usually generated by the nodes with smaller DODs (closer to 0), Equation 5 can increase timer differences and reduce the number of duplicate messages. Another benefit of using Equation 5 is that the width of duplication area is not fixed but changeable by tuning the parameter α.

According to Equation 5, if the difference between any two DODs is very small, we have the following approximation:

$$f(x_2) - f(x_1) \approx (x_2 - x_1) \cdot \frac{d}{dx} f(x)\Big|_{x=x_1} \tag{6}$$

taking derivative of the above approximation yields:

$$f(x_2) - f(x_1) \approx (x_2 - x_1)\frac{T\alpha}{2R}\Big(\frac{x_1 - \bar{x}_i}{2R}\Big)^{\alpha-1} \tag{7}$$

giving

$$\Delta \approx \frac{2\delta R}{\alpha T}\Big(\frac{x_1 - \bar{x}_i}{2R}\Big)^{-\alpha+1} \tag{8}$$

The expression takes its largest value when $x_1 - \bar{x}_i$ is the largest. We have

$$x_1 - \bar{x}_i = x_1 - x_i + R + \rho R \in [\rho R, (1 + \rho)R] \tag{9}$$

So the maximal Δ is obtained when $x_1 - \bar{x}_i = \max\{(1 + \rho)R\} = 2R$, i.e. the upper bound of Δ is:

$$\frac{2}{\alpha} \cdot \frac{\delta R}{T} \tag{10}$$

From the above equation, we can see the duplication area in LPC is at most $\frac{2}{\alpha}$ times of that in CBF-AS. Because duplication often occur on nodes with $x_k - x_i \approx -R$, the actual value of Δ in LPC can be expressed as:

$$\frac{2}{\alpha}\Big(\frac{\rho}{2}\Big)^{-\alpha+1} \cdot \frac{\delta R}{T} \tag{11}$$

Assume $\alpha = 0.9, \rho = 0.1$, the duplication area in LPC is at most 1.64 times of that in CBF-AS. Since duplicate (CTF) messages are ignored by the current packet sender, they may not drastically affect the network performance. This speculation is also verified by simulation results.

5. ANALYSIS OF LOCATION PRIVACY PROTECTION

The efficiency of a location privacy protection system is usually evaluated by the hardness for an adversary to attack the proposed system. The hardness is defined as the entropy of an adversary correctly predicting all nodes' locations. For example, if an adversary can accurately predict each node's location at any time (e.g. in GPSR), the hardness of privacy attack is 0.

5.1. Entropy Based Analysis of Location Privacy Protection

To model the hardness for an adversary to attack the location privacy of the whole system, we need to compute the attack probability for every node in the networks. Not only is the attack probability itself important but also the distribution of the attack probability. For example, if the attack probability for every vehicle is the same, the entropy will be the largest and thus the highest level of protection is achieved. Although it is expensive for an adversary to apply advanced technologies (e.g. fingerprinting and localization algorithms), we need to consider this possibility as well.

As we described in previous sections, the location privacy of vehicles includes two types of information: node's identification and location. We claim that preservation of both identification and location is necessary because it is possible for the adversary to detect the node's ID or location through other techniques.

We first investigate how to compute the attack probability for a node (or vehicle) in a network. To compute this probability, we define two matrices $IM(s, p)$ and $LM(s, p)$ which record the number of times that node s appears at location p. Without location protection, e.g. in GPSR, there is only a 1 in each row and column. In other word, a node definitely appears at a certain location. If location protection is applied, a node could be at several locations and thus the adversary only predict the node's location probabilistically. Therefore, these matrices can be considered snapshots of the networks which record the possible locations of every node.

The matrix $LM(s, p)$ is used when identification detection technologies (e.g. fingerprinting) are available to the adversary. If localization algorithms are available to the adversary, $IM(s, p)$ will be used.

As shown in Figure 5, the dimension denoted as $I = I_0, I_1, \cdots, I_n$ records the IDs of all nodes in the network, and the other dimension $L = L_0, L_1, \cdots, L_n$ denotes the possible locations of all nodes. The initial value of every

Security Comm. Networks 2010; **00**:1–13 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/sec
Prepared using secauth.cls

$$\begin{array}{c|cccccc} & L_0 & L_1 & \cdots & L_i & \cdots & L_n \\ I_0 & 1 & 1 & \cdots & 1 & \cdots & 0 \\ I_1 & 1 & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ I_j & 1 & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ I_n & 1 & 1 & \cdots & 1 & \cdots & 0 \end{array}$$

Figure 5. Matrix recording possible identifications and locations of nodes

entry in $IM(s,p)$ and $LM(s,p)$ is 0. It increases by 1 if the adversary captures a message indicating that node I_i is (probably) located at L_j.

At a certain time instance t, the adversary may capture many packets from concurrent communications in the network. Let us look at one of them, the communication between node i and j as shown in Figure 1. Node i sends a RTF including its dummy DOD to its neighbors. After receiving this message, node j replies a CTF containing a pseudonym. Then, data is delivered from i to j with the pseudonym.

In the first step, node i hides its true location in an area using a dummy DOD. Since the dummy DOD is randomly chosen, there may be other nearby nodes providing the same DOD. Those nodes can be presented as:

$$\chi_i = \{k : x_k \in (x_i - R, x_i + R)\},\ (k \neq i) \qquad (12)$$

where x_k and x_i are the DODs of node k and i, respectively. Without other information, the adversary only knows that a node in χ_i sends the packet. So it updates $LM(s,p)$ to $LM(s,p)+1$ where $p = L_k (k \in \chi_i)$ because any node $k \in \chi_i$ can send the same dummy DOD. For example, suppose there are two node n_1 and n_2 in χ_i. The adversary only knows a message is sent from location L_1, L_2 or L_i, but has no clue who sent this packet. Therefore, the adversary node updates $LM(s,p) = LM(s,p)+1$ where $p = L_1, L_2, L_i$ and $s = I_0, I_1, \cdots, I_n$.

In the second step, node j sends a randomly chosen pseudonym in its CTF message. Because every node can send the same pseudonym, the adversary updates the matrix by changing $IM(s,p)$ to $IM(s,p)+1$ for all entries where $s = I_0, I_1, \cdots, I_n$ and $p = L_0, L_1, \cdots, L_n$.

In the third step, since no identification or location information is contained in data packets, the adversary learns nothing from this packet. Therefore, we simply omit this step in the analysis of location privacy protection.

If we look at the CBF-AS protocol, the current packet forwarder i sends a RTF along with its location. The next hop j sends a CTF with its ID. In the first step, we set $LM(s,p) = LM(s,p)+1$, where $p = L_i$, for every $s = I_1, I_2, \cdots, I_n$. In this case, the adversary node only needs to predict from which node this message is sent. In the second step, we set $IM(s,p) = IM(s,p)+1$, where $s = I_j$, for every $p = L_1, L_2, \cdots, L_n$. This is because the adversary node only needs to predict where node j is.

For the matrix M (either IM or LM), the value of $M(s,p)$ records the number of times that node s probably appears at location p. From this number, we can compute the joint probability of node s locating at p. In other words, the attacking probability of node i can be obtained by normalizing the matrix.

$$P(s = I_i, p = L_i) = \frac{M(I_i, L_i)}{\sum_{s,p} M(s,p)} \qquad (13)$$

This equation models the probability of the adversary attacking the location privacy of node i. For example, if the adversary receives a RTF from node i, the attack probability for node i will be $1/(|\chi_i|+1)$ where $|\chi_i|$ is the cardinality of χ_i. If the adversary receives a CTF from node j, since the pseudonym can be sent by other nodes, the attack probability for j will be $1/|\psi_j|$ where ψ_j is the set of nodes which can send the same pseudonym as node j does. Usually, the size of ψ_j is equal to $n+1$ the number of nodes in the network.

If the adversary node is able to detect a node's ID (through other advanced techniques), the conditional probability of the node i being located at L_i will be:

$$P(p = L_i | s = I_i) = \frac{LM(I_i, L_i)}{\sum_{p} M(I_i, p)} \qquad (14)$$

Therefore, the Shannon's entropy required by the adversary to correctly predict that node i is located at L_i will be:

$$H_i^L = \sum_{p} P(p|s = I_i) \cdot \log \frac{1}{P(p|s = I_i)} \qquad (15)$$

Similarly, if the adversary can detect a node's location through localization algorithms, the conditional probability that the node at location L_i must be I_i is:

$$P(s = I_i | p = L_i) = \frac{IM(I_i, L_i)}{\sum_{s} M(s, L_i)} \qquad (16)$$

So we can compute the entropy of predicting the node at location L_i must be I_i as:

$$H_i^I = \sum_{s} P(s|p = L_i) \cdot \log \frac{1}{P(s|p = L_i)} \qquad (17)$$

If the adversary can localize nodes in the network, the hardness of predicting IDs will be the cumulative entropy:

$$E^I = \sum H_i^I, i = 0, 1, \cdots, n \tag{18}$$

where we assume the network events (e.g. sending RTF message from different nodes) are independent to each other. If these events are dependent to each other, we can use the average entropy to model the hardness for the adversary to predict IDs. This average entropy can be computed as $\bar{E}^I = E^I/(n+1)$.

If the adversary is able to identify nodes, the uncertainty of predicting the locations of nodes can be modeled as the cumulative entropy:

$$E^L = \sum H_i^L, s = 0, 1, \cdots, n \tag{19}$$

where we also assume network events are independent to each other. Similarly, when these events are dependent on each other, the average entropy $\bar{E}^L = E^L/(n+1)$ is used.

As we described previously, it is expensive for the adversary to identify or localize nodes in the network. We consider the difficulty of enabling the identification and localization functions in the adversary to be D^I and D^L, respectively. Then, the hardness of the adversary attacking the whole system is:

$$H(M) = \min\left(D^I + E^L,\ D^L + E^I\right) \tag{20}$$

where E^I and E^L can be computed from IM and LM by Equation 18 and 19, respectively. If no identification or localization techniques are available to the adversary, we have $D^I = E^I$ and $D^L = E^L$, which are the most common case in reality.

The proposed analytical method relies on two matrices IM and LM which model the hardness of predicting identification and location, respectively. The update policy of these two matrices should be changed accordingly if new privacy protection scheme is applied. Because Equation 20 is still applicable, the proposed model can be used to evaluate the location privacy protection of other protocols as well.

5.2. Impact of Prior Knowledge

The proposed entropy based analysis assumes that the adversary relies purely on communication packets to attack the location privacy of vehicles in networks. However, the location information of the adversary may be helpful in localizing nodes. For instance, a captured RTF message must be sent from a vehicle within the adversary's communication range. Therefore, the area in which the vehicle is hidden is reduced, i.e. $|\chi|$ is smaller. Therefore, the attack probability of this vehicle's location privacy becomes larger.

We define the information that helps the adversary to identify or localize nodes as prior knowledge. The easiest way for the adversary to gain prior knowledge is analyzing the information provided in digital maps. For example, using the pre-determined published public transport timetable, the adversary can determine the route and the timing of a city bus. Combing the location information of the adversary, a bus's location can be restricted to a small area or a short road segment. Therefore, the attack probability of this bus is drastically increased.

On the other hand, the adversary may infer a vehicle's identification from existing location information. For instance, in a rural area where everyone may own a house with a parking space. Usually, the owner will park his/her car in the garage, so the adversary can identify this car by the location of this house which is provided in digital maps. Furthermore, the house owner's information may be found by searching the local phone book.

There may be other types of prior knowledge used by the adversary to attack the location privacy of vehicles. It is impossible to name all of them, and the list is increasing with the advance of new technologies. No matter what type of prior knowledge is used by the adversary, the proposed analytical model still works. The only change is that, with prior knowledge, fewer entries in the matrices need to be updated because the size of χ and ψ becomes smaller. The worst case occurs when $\chi = 1$ and $\psi = 1$ for all nodes in networks, i.e. no protection is provided for any vehicle. In that case, because the prior knowledge already releases the location and identification information of all vehicles, it is meaningless to investigate the location privacy any more. As far as we know, no such prior knowledge is available for the adversary in VANETs.

6. SIMULATIONS AND RESULTS

We implement the LPC protocol in ns-2.33 and compare its network performance with the original geographic routing protocols: GPSR and CBF-AS. To exam the effectiveness of the location privacy protection in LPC, we implement the identification protection scheme which periodically changes the pseudonyms of nodes. By adding this protecting mechanism, we generate another two protocols: CBF-AS-ID and GPSR-ID.

Since the mobilities of vehicles in networks are important in evaluating network performance and location privacy protections, we generate the movements of nodes using VanetMobiSim [42]. The road topology in simulations is imported from the TIGER (Topologically Integrated Geographic Encoding and Referencing) database, which is used by the United States Census Bureau to describe land attributes of U.S. The road map is a $2000m \times 2000m$ square area which centers at latitude: 42095405 and longitude: -72530684, and has 11 intersections and 25 road segments.

In simulations, different parameters are used to evaluate the data delivery ratio, end-to-end delay, network

Security Comm. Networks 2010; **00**:1–13 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/sec
Prepared using secauth.cls

throughput and location privacy protection of all routing protocols. Details of the simulation setup parameters are listed as follows: 1) the communication range is $250m$, 2) α and ρ are chosen from $[0.1, 0.9]$, and 3) the maximal one-hop delay T and suppression delay δ are $0.1s$ and $2\mu s$, respectively. Since we are interested in how different location privacy protection schemes affect network performance, the impact of different error models are not considered and evaluation in this paper. In each simulation, source and destination nodes are randomly selected from all nodes in networks. The simulation time is 2000 seconds and each scenario is repeated 20 times to achieve a result with a high level of confidence.

6.1. Networking Effectiveness of LPC

Because LPC hides a vehicle's true location in an area when it forwards packets, it is extremely important to know if the network performance is affected; if so, in which direction is the impact and by how much. In the simulations, we set $\alpha = 0.9$ and ρ as a random number from $(0, 1)$. Different data sending rates (0.1 to 1 pkt/s) are used and four protocols LPC, GPSR, CBF-AS, and CBF-AS-ID are compared. Because the degraded network performance of GPSR-ID has been verified in [8], we do not represent the similar results in this work.

Since we are interested in how location privacy protections affect network performance, the impact of channel fading and errors are not evaluation in the simulations. With lossy channels, more RTFs and CTFs are expected to be lost, which will cause longer networking delay and network congestion. In addition, the loss of CTF is more critical than RTF since CTF is considered a suppression signal for others to cancel their timers, which may cause more duplicate CTFs in LPC. Although duplicate CTFs could be handled well by the current data forwarder, more duplicate CTFs could potential cause network congestion, and thus affect the effective delivery of RTF and DATA messages.

6.1.1. Data Delivery Ratio

Data delivery ratio is defined as the number of packets received at the destination divided by the number of packets sent from the source. As shown in Figure 6, LPC delivers almost the same number of packets as CBF-AS does. In GPSR, every node selects the next hop based on the stored neighbors' information obtained from beacon messages. Such information may contain out-of-date neighbors, so packets may be sent to the next hop which moved out of the communication range. Although the CBF-AS-ID periodically changes nodes' pseudonyms, its data delivery ratio is similar to CBF-AS. Because the time difference between a packet being delivered and a CTF being sent is very small, there is no chance for the next hop to change its pseudonym. Therefore, no packet is dropped because nodes change their pseudonyms.

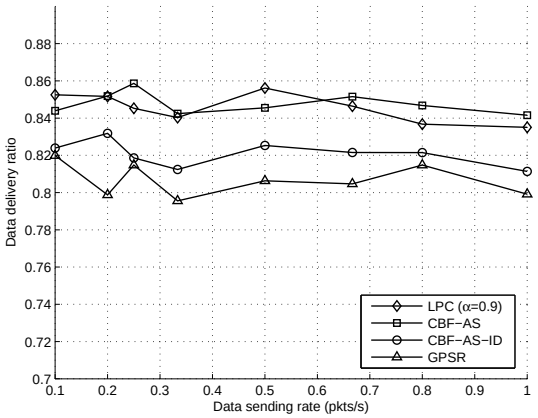

Figure 6. Data delivery ratio vs data sending rate

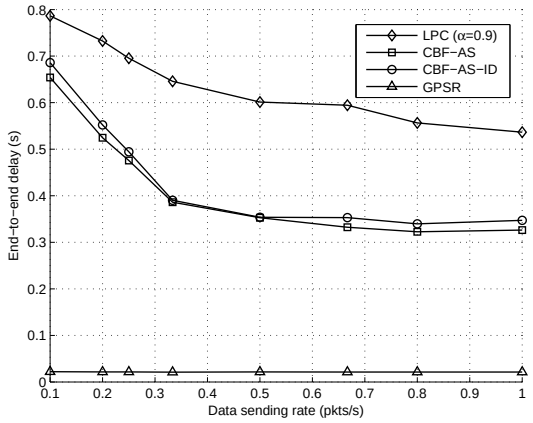

Figure 7. End-to-end delay vs data sending rate

6.1.2. End-to-end Delay

The end-to-end delay is defined as the average time for a packet to be transmitted from the source to the destination. As shown in Figure 7, GPSR gives the smallest end-to-end delay because there is no delay in selecting every next hop. However, in LPC, CBF-AS and CBF-AS-ID, after a node sends a RTF message, its neighbors compete with each other by setting different timers, which causes delays.

According to Equation 5, the timer on every elected next hop in LPC is larger than that in CBF-AS. Therefore, the end-to-end delay of LPC is larger than CBF-AS and CBF-AS-ID due to its longer single-hop delay. The end-to-end delay of LPC can be further reduced using a smaller maximal one-hop delay T with an appropriate α.

6.1.3. Network Throughput and Overhead

Network throughput is defined as the number of packets delivered to the destination every second. The network throughput of all protocols increase as the data sending rate

increases, and do not show differences between different approaches. Since network congestion is not significant in the simulations, the higher the delivery ratio, the larger the throughput.

Although the network throughputs of different protocols are similar, their networking overheads could be significantly different. Broadcasting beacon messages is considered the largest overhead of GPSR. In GPSR, each node periodically sends location and ID to its one-hop neighbors even though it is not, or will not be, participating any routing activities. In other words, the overhead of GPRS highly depends on the number of nodes in networks and the broadcasting interval of beacon messages. On the other hand, the overheads of CBF-AS and LPC are determined by the number of packets being forwarded within the network, i.e. higher the volume of network traffic, larger the overhead. Therefore, the overheads of GPSR, CBF-AS and LPC could be quite different depending on the number of nodes, beacon interval and network traffic. In terms of communication overhead, CBF-AS and LPC are similar although larger computation overhead is expected on the LPC front. In summary, LPC does not add significant extra communication overhead comparing to CBF-AS.

The computation overhead brought by LPC is bigger than its competitors. According to Equation 2, the LPC protocol needs to compute a dummy DOD for every packet being forwarded. To guarantee randomness in computed dummy DODs, a random number generator is also needed on each node. According to Equation 5, to set up a timer in LPC, exponentiation operations are required to compute the value of $T\left(\frac{x-\bar{x}_i}{2R}\right)^{\alpha}$, which needs more CPU cycles. However, because large computation and processing capacities are expected on regular vehicles, the extra computation overhead added will not degrade LPC's performance.

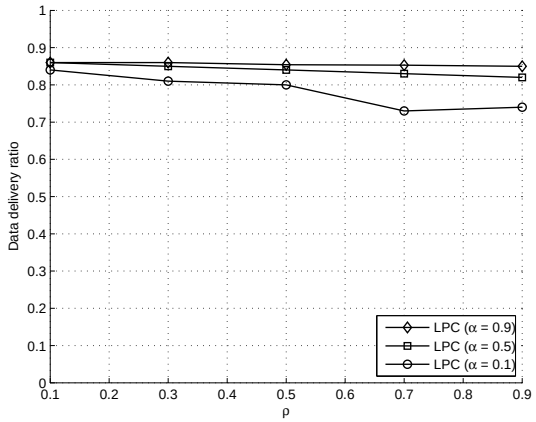

Figure 8. Data delivery ratio vs different settings of ρ and α

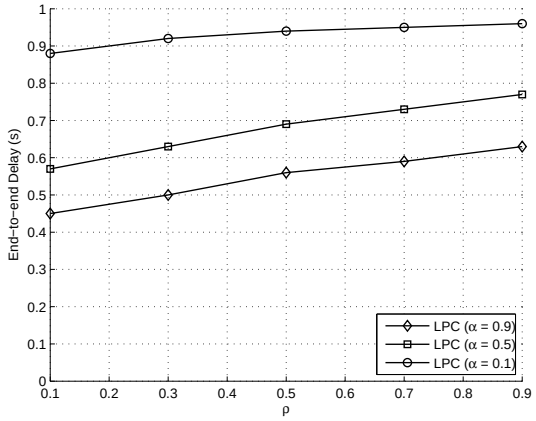

Figure 9. End-to-end delay vs different settings of ρ and α

6.2. Impact of Different Parameters

In LPC, there are two important parameters ρ and α which control the randomness of dummy DODs and the number of duplicate messages. By default, LPC uses random ρ to choose dummy DODs, but we use fixed ρ to investigate how it affects network performance. In addition, we need to investigate the impact of different α on the performance of LPC.

As shown in Figure 8 and 9, when the ρ increases from 0.1 to 0.9, the data delivery ratio of LPC decreases but the end-to-end delay increases. When ρ becomes larger, the timer on every next hop runs longer, and thus the end-to-end delay increases. Due to the delay at each hop, the network becomes more congested so the data delivery ratio drops. This phenomena can be seen in Figure 6 where the delivery ratio decreases when the data sending rate increases.

When α increases, the network performance becomes better in terms of higher delivery ratios and lower delays. Similar to the parameter ρ, when α becomes larger, the timer on every next hop becomes shorter, so the network performance is improved. The parameter α becomes more significant when the maximal one-hop delay T becomes smaller. For example, when $T = 0.05s$, we need to set $\alpha > 0.4$ to obtain an acceptable network performance. Due to the space limit, we do not show the relevant results.

6.3. Impact of Network Density

Besides the two scaling factors ρ and α, network density also plays an important rule in evaluating network performance. As shown in Figure 10 and 11, when the number of nodes increases, the data delivery ratio increases drastically. Because better next hops can be selected as more nodes are involved in next-hop elections, the delivery ratio of LPC increases. When network is sparse (e.g. 70 nodes), the performance of different strategies are quite different; however, they becomes similar when network becomes dense (e.g. 170 nodes). When network density is higher, it is likely that the chosen next-hop is closer

Security Comm. Networks 2010; **00**:1–13 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/sec
Prepared using secauth.cls

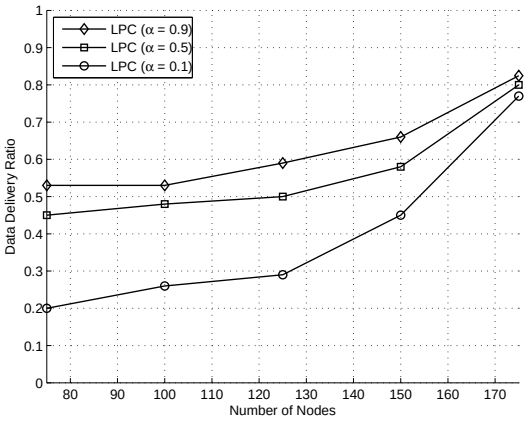

Figure 10. Data delivery ratio vs number of nodes

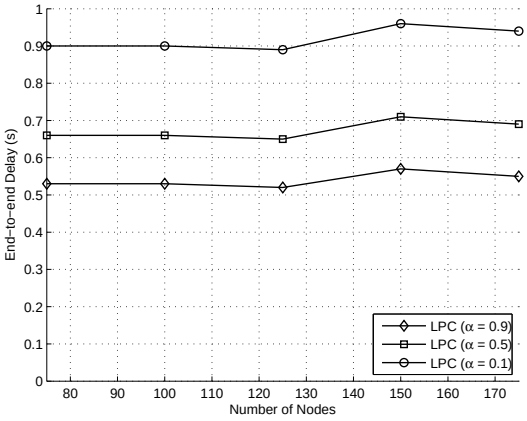

Figure 11. End-to-end delay vs number of nodes

to the destination, i.e. $\left(\frac{x-\bar{x}_i}{2R}\right)$ gets close to 0. In this case, the duplication areas of various α are almost the same. If $\left(\frac{x-\bar{x}_i}{2R}\right)$ approaches to 1, i.e. sparse networks, the duplication area of $\alpha = 0.1$ is much larger than $\alpha = 0.9$.

The end-to-end delay slightly increases when the network density increases. In ad hoc networks, end-to-end delay highly depends on the number of hops. Because the number of hops of LPC does not change much when the network becomes denser, the end-to-end delay just increases a little.

6.4. Location Privacy Protection

According to Equation 20, the hardness for an adversary to attack the location privacy of different protocols can be computed. As shown in Table I, assuming no prior knowledge, the attack hardness of GPSR, GPSR-ID, CBF-AS, CBF-AS-ID and LPC (with $1pkt/s$ data sending rate) are 0, 243, 305, 531, and 1320, respectively. We note that the hardness of LPC is about 2.5 times of the second best protocol CBF-AS-ID.

In GPSR, every node periodically broadcasts its location and ID information, so the adversary node eavesdropping these data can easily compute the matrices IM and LM. Suppose no identification and localization technologies are available on the adversary node, $M = IM = LM$ becomes an identity matrix in which all elements on the main diagonal are equal to 1 and all other elements are equal to 0. With this identity matrix, the adversary can certainly compute the location of any node at anytime. Therefore, the entropy of attacking the location privacy of vehicles in GPSR is 0.

In CBF-AS, each packet forwarder sends its location (no ID) in RTFs, and every self-elected next hop sends its ID (no location) in CTFs. If the adversary overhears a RTF message, it knows a node is communicating from a certain location (e.g. L_i), but does not know who is sending this message. Therefore, it adds n 1s into the ith column of matrix M. Similarly, when a CTF message (e.g. from node j) is overheard, the adversary add n 1s into the jth row of M. Therefore, some degree of location privacy protection is achieved in CBF-AS. Since the CBF-AS-ID periodically changes node pseudonyms, vehicle identification information are hidden. That is why CBF-AS-ID gives a higher level of location privacy protection than CBF-AS. Similarly, GPSR-ID achieves a better privacy protection than GPSR.

In LPC, dummy DODs and pseudonyms are used in RTFs and CTFs, respectively. Even though the adversary eavesdrops a RTF, it does not know where exactly it comes from and has no idea of who is actually sending it. Instead, it could only estimate that the message is sent from a certain area (due to the computation of dummy DOD), it will then update several columns in M which correspond to the locations with the area. While a CTF is overheard, the adversary needs to update all rows in M because it does not know whoever and wherever the message is being sent. Therefore, LPC provides the best location privacy protection.

7. CONCLUSIONS AND FUTURE WORK

The major contribution of this paper is a protocol that lets users benefit from location-based geographic routing while at the same time retain their location privacy. Although LPC is designed for VANETs, it can be easily generalized and used in any other types of mobile ad hoc networks. By changing the values of ρ and α, the level of location privacy protection and the number of duplicate responses can be easily tuned. This is a unique feature which makes LPC suitable for different network settings.

In future, we will replace the RTF/CTF by the RTS/CTS messages because the current active selections of next hops generate too much overhead. Some new information (e.g. packet sequence number, destination location, dummy

Table I. Hardness of attacking location privacy by the adversary

λ **(pkt/s)**	LPC	CBF-AS-ID	CBF-AS	GPSR-ID	GPSR
0.10	1320	531.02	305.32	243.14	0
0.20	1321	531.24	305.32	243.25	0
0.25	1322	531.62	305.32	243.25	0
0.33	1320	531.65	305.15	243.15	0
0.50	1320	531.55	305.21	243.15	0
0.67	1320	531.65	305.62	243.25	0
0.80	1320	531.64	305.62	243.22	0
1.00	1320	531.62	305.36	232.55	0

IDs) will be added to current RTS/CTS packets. The small modifications can be easily implemented in the IEEE 802.11 MAC protocol. Moreover, as geographic routing is widely used in pervasive computing, LPC can be also applied to the applications in pervasive computing to achieve better location privacy protections.

REFERENCES

1. Ying B, Makrakis D, Mouftah H. Dynamic mix-zone for location privacy in vehicular networks. *IEEE Communications Letters* 2013; **17**(8):1524–1527.
2. Lu R, Li X, Luan T, Liang X, Shen X. Pseudonym changing at social spots: An effective strategy for location privacy in VANETs. *IEEE Transactions on Vehicular Technology* 2012; **61**(1):86–96.
3. Tang L, Hong X, Vrbsky S. Using camouflaging mobility to protect privacy in mobile ad hoc networks. *Security and Communication Networks* 2009; **2**(6):580–594.
4. Pan Y, Li J. Cooperative pseudonym change scheme based on the number of neighbors in VANETs. *Journal of Network and Computer Applications* 2013; **36**(6):1599 – 1609.
5. Sweeney L. k-Anonymity: A model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* 2002; **10**(5):557–570.
6. Kido H, Yanagisawa Y, Satoh T. Protection of location privacy using dummies for location-based services. *Data Engineering Workshops, 2005. 21st International Conference on*, 2005; 1248–1248.
7. Lu H, Jensen CS, Yiu ML. PAD: Privacy-area aware, dummy-based location privacy in mobile services. *Proceedings of the 7th International ACM Workshop on Data Engineering for Wireless and Mobile Access*, 2008; 16–23.
8. Schoch E, Kargl F, Leinmuller T, Schlott S, Papadimitratos P. Impact of Pseudonym Changes on Geographic Routing in VANETs. *Third European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS 2006)*, 2006; 43–57.
9. Karp B, Kung HT. GPSR: Greedy perimeter stateless routing for wireless networks. *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*, 2000; 243–254.
10. Fler H, Widmer J, Ksemann M, Mauve M, Hartenstein H. Contention-based forwarding for mobile ad hoc networks. *Ad Hoc Networks* 2003; **1**(4):351 – 369.
11. Yang Q, Lim A, Li S, Fang J, Agrawal P. ACAR: Adaptive connectivity aware routing for vehicular ad hoc networks in city scenarios. *Mob. Netw. Appl.* Feb 2010; **15**(1):36–60.
12. Lochert C, Mauve M, Füssler H, Hartenstein H. Geographic routing in city scenarios. *SIGMOBILE Mob. Comput. Commun. Rev.* 2005; **9**(1):69–72.
13. Zhao J, Cao G. VADD: Vehicle-assisted data delivery in vehicular ad hoc networks. *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 2006; 1–12.
14. Naumov V, Gross T. Connectivity-aware routing (CAR) in vehicular ad-hoc networks. *Proceedings of the 26th IEEE International Conference on Computer Communications*, 2007; 1919–1927.
15. Li T, Li Y, Liao J. A contention-based routing protocol for vehicular ad hoc networks in city environments. *Distributed Computing Systems Workshops, 2009. ICDCS Workshops '09. 29th IEEE International Conference on*, 2009; 482–487.
16. Korkmaz G, Ekici E, Özgüner F, Özgüner U. Urban multi-hop broadcast protocol for inter-vehicle communication systems. *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, VANET '04, 2004; 76–85.
17. Chen D, Deng J, Varshney P. On the forwarding area of contention-based geographic forwarding for ad hoc and sensor networks. *Second Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2005; 130–141.
18. Rao A, Ratnasamy S, Papadimitriou C, Shenker S, Stoica I. Geographic routing without location

DOI: 10.1002/sec
Prepared using secauth.cls

information. *MobiCom '03*, ACM: New York, NY, USA, 2003; 96–108.
19. Emara K. Location privacy in vehicular networks. *IEEE 14th International Symposium and Workshops on World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013*, 2013; 1–2.
20. Alganas A, Lin X, Grami A. Evse: An efficient vehicle social evaluation scheme with location privacy preservation for vehicular communications. *Communications (ICC), 2011 IEEE International Conference on*, 2011; 1–5.
21. Raya M, Hubaux JP. The security of vehicular ad hoc networks. *Proceedings of the 3rd ACM workshop on Security of ad hoc and sensor networks*, SASN '05, 2005; 11–21.
22. Popa RA, Balakrishnan H, Blumberg A. VPriv: Protecting Privacy in Location-Based Vehicular Services. *18th USENIX Security Symposium*, Montreal, Canada, 2009.
23. Beresford A, Stajano F. Location privacy in pervasive computing. *Pervasive Computing, IEEE* Jan-Mar 2003; **2**(1):46–55.
24. Song JH, Wong VW, Leung VC. Wireless location privacy protection in vehicular ad-hoc networks. 2010; 160–171.
25. Sampigethaya K, Li M, Huang L, Poovendran R. AMOEBA: Robust location privacy scheme for VANET. *Selected Areas in Communications, IEEE Journal on* Oct 2007; **25**(8):1569–1589.
26. Dötzer F. Privacy issues in vehicular ad hoc networks. *Privacy Enhancing Technologies*, 2005; 197–209.
27. Lu R, Lin X, Zhu H, Ho PH, Shen X. ECPP: Efficient conditional privacy preservation protocol for secure vehicular communications. *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, 2008; 1229–1237.
28. Sampigethaya K, Huang L, Li M, Poovendran R, Matsuura K, Sezaki K. CARAVAN: Providing location privacy for VANET. *Proceedings of Embedded Security in Cars (ESCAR)*, 2005.
29. Gerlach M, Guttler F. Privacy in VANETs using changing pseudonyms - ideal and real. *IEEE 65th Vehicular Technology Conference, 2007.*, 2007; 2521–2525.
30. Pan Y, Li J, Feng L, Xu B. An analytical model for random changing pseudonyms scheme in VANETs. *International Conference on Network Computing and Information Security (NCIS), 2011*, vol. 2, 2011; 141–145.
31. Wasef A, Shen XS. REP: Location privacy for VANETs using random encryption periods. *Mobile Networks and Applications* Feb 2010; **15**(1):172–185.
32. Mehta K, Liu D, Wright M. Location privacy in sensor networks against a global eavesdropper. *Network Protocols, 2007. ICNP 2007. IEEE International Conference on*, 2007; 314–323.
33. Freudiger J, Raya M, Flegyhzi M, Papadimitratos P, Hubaux JP. Mix-Zones for Location Privacy in Vehicular Networks. *ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS)*, 2007.
34. Kohno T, Broido A, Claffy K. Remote physical device fingerprinting. *IEEE Transactions on Dependable and Secure Computing* 2005; **2**(2):93–108.
35. Gerdes RM, Daniels TE, Mina M, Russell SF. Device identification via analog signal fingerprinting: A matched filter approach. *Proceedings of the 2006 Network and Distributed System Security Symposium (NDSS '06)*, 2006.
36. Brik V, Banerjee S, Gruteser M, Oh S. Wireless device identification with radiometric signatures. *Proceedings of the 14th ACM international conference on Mobile computing and networking*, MobiCom '08, 2008; 116–127.
37. Gezici S. A survey on wireless position estimation. *Wireless Personal Communications* 2008; **44**:263–282.
38. Golden S, Bateman S. Sensor measurements for Wi-Fi location with emphasis on time-of-arrival ranging. *IEEE Transactions on Mobile Computing* 2007; **6**(10):1185–1198.
39. Desmond LCC, Yuan CC, Pheng TC, Lee RS. Identifying unique devices through wireless fingerprinting. *Proceedings of the first ACM conference on Wireless Network Security*, WiSec '08, ACM: New York, NY, USA, 2008; 46–55.
40. Niculescu D, Nath B. Ad hoc positioning system (APS) using AOA. *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications*, 2003; 1734–1743.
41. Hoene C, Willmann J. Four-way TOA and software-based trilateration of IEEE 802.11 devices. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2008; 1–6.
42. Härri J, Filali F, Bonnet C, Fiore M. VanetMobiSim: Generating realistic mobility patterns for VANETs. *VANET '06: Proceedings of the 3rd international workshop on Vehicular ad hoc networks*, ACM Press, 2006; 96–97.