DOI 10.1007/s11036-009-0167-4

Wireless Location Privacy Protection in Vehicular Ad-Hoc Networks

Joo-Han Song · Vincent W.S. Wong · Victor C.M. Leung

Published online: 21 May 2009

Abstract Advances in mobile networks and positioning technologies have made location information a valuable asset in vehicular ad-hoc networks (VANETs). However, the availability of such information must be weighted against the potential for abuse. In this paper, we investigate the problem of alleviating unauthorized tracking of target vehicles by adversaries in VANETs. We propose a vehicle density-based location privacy (DLP) scheme which can provide location privacy by utilizing the neighboring vehicle density as a threshold to change the pseudonyms. We derive the delay distribution and the average total delay of a vehicle within a density zone. Given the delay information, an adversary may still be available to track the target vehicle by some selection rules. We investigate the effectiveness of DLP based on extensive simulation study. Simulation results show that the probability of successful location tracking of a target vehicle by an adversary is inversely proportional to both the traffic arrival rate and the variance of vehicles' speed. Our proposed DLP scheme also has a better performance than both Mix-Zone scheme and AMOEBA with random silent period.

Keywords vehicular ad-hoc network · location privacy

J.-H. Song · V. W. S. Wong (✉) · V. C. M. Leung
Department of Electrical and Computer Engineering,
The University of British Columbia, Vancouver,
British Columbia, Canada
e-mail: vincentw@ece.ubc.ca

J.-H. Song
e-mail: joohans@ece.ubc.ca

V. C. M. Leung
e-mail: vleung@ece.ubc.ca

1 Introduction

Due to the technological advancement and widespread deployment, nowadays there are various ways to access the Internet from a mobile device. For example, one can either use the IEEE 802.11 [1] compliant WiFi networks or by subscribing to the third generation (3G) wireless cellular services. Recently, a new type of wireless ad-hoc networks, also known as the vehicular ad hoc network (VANET), has gained increasing attention. VANETs aim to provide communications among vehicles (i.e., vehicle-to-vehicle (V2V)) and between vehicle and fixed roadside equipment (i.e., vehicle-to-roadside (V2R)). Significant progress has been made in Intelligent Transportation Systems (ITS) to create a safe and efficient driving environment. The DSRC (Dedicated Short Range Communications) [2] is a short to medium range wireless technology for V2R and V2V communications. VANET is an important component in ITS, and is expected to play a crucial role in various applications such as collision warning, road sign alarms, driver assistance, and infotainment.

In safety enhancing applications, each vehicle needs to periodically broadcast an authenticated safety message, which includes its verifiable identity, its current location, speed, and acceleration. Although these safety messages can help to prevent accidents, they may also be used by the adversaries for unauthorized location tracking of vehicles. By using an external IEEE 802.11 compliant wireless network, an attacker can eavesdrop on all the broadcast messages and determine the locations visited by the vehicles (or users) over a period of time. The location history information (or mobility traces of the target vehicles) can be associated with places of interest of vehicle users. This information can be exploited for advertisement or surveillance. Moveover, it can be misused for crimes, such as abductions or automobile thefts. Protecting the location privacy of vehicles is important because the lack of privacy may hinder the wide acceptance of VANET technology.

In general, location privacy protection schemes for mobile networks can be classified as policy-based [3] and anonymity-based [4]. In policy-based schemes, vehicles specify their location privacy preferences as policies and trust that the third party location-based service (LBS) providers adhere to these policies. In the anonymity-based approaches, the location tracking of a target vehicle can be mitigated by using a randomly chosen and changing identifier, called the *pseudonym*. Pseudonyms are generated in a predefined way such that the adversaries cannot link a new pseudonym to previous ones of the same vehicle. Pseudonyms can either be a set of public keys, network layer addresses, or link layer addresses [5]. The change of pseudonym denotes that the vehicle either changes its public key or addresses on the different layers (i.e., network and link). This approach regards anonymity as being untraceable between two successive locations of the target. Since pseudonyms cannot be linked to each other, they can provide a certain degree of privacy. In general, frequently changing pseudonyms are accepted as a solution to protecting the privacy of VANET [6]. However, in most communication protocols, it can be complicated to re-establish the session frequently as both communication parties should be re-authenticated.

Since adversaries can predict the movement of the vehicles based on the position and speed information from all the broadcast messages of a target vehicle, they can use this prediction to link different pseudonyms of the target vehicle with a high probability. In this paper, we propose a vehicle density-based location privacy (DLP) scheme which can provide location privacy by utilizing the neighboring vehicle density as a threshold to change the pseudonyms. The level of location privacy protection of changing pseudonyms depends on the road traffic condition. The neighboring vehicle density can be used to assess the road traffic condition. In particular, a high neighboring vehicle density usually indicates that there is traffic congestion. Each vehicle can easily obtain the number of neighboring vehicles through the exchange of the periodic link-layer HELLO messages in VANETS. In our proposed scheme, by monitoring the neighboring vehicle density, each vehicle updates its pseudonym only when there are at least $k_i - 1$ distinct neighboring vehicles. To the best of our knowledge, it is the first approach that considers the neighboring vehicle density as the triggering factor for updating the pseudonym. The goal of our proposed scheme is to minimize the probability of successful location tracking of a target vehicle by an adversary. The contributions of this paper are as follows:

1. We propose the vehicle density-based location privacy (DLP) scheme, which can mitigate the location tracking of vehicles by changing pseudonyms based on a threshold in neighboring vehicle count within a density zone.
2. We derive the delay distribution and the expected total delay of a vehicle within the density zone. Given the delay information, an adversary may still be available to track the target vehicle based on selection rules.
3. Simulation results show that the probability of successful location tracking by an adversary is inversely proportional to the intensity of the traffic

Table 1 Index of key notations

Notation	Description
k_i	Total number of neighboring vehicles in a density zone i
w	A vehicle in the VANET
m_i	Number of ports in density zone i
n_i	Number of intersections in density zone i
d_i	Distance from a port i to the intersection
S_i	Random variable for vehicle speed at the road segment i
$f_{S_i}(s)$	Probability distribution of random variable S_i
μ_i	Mean speed of vehicle at the road segment i
σ_i^2	Variance of vehicle speed at the road segment i
A_i	Random variable of the inter-arrival time of vehicles at port i
λ_i	Average arrival rate of vehicles at port i
ξ	Maximum number of vehicles per lane per second
α_{ij}	Probability of entering port i and exiting port j
T_i	Random variable of duration for traveling on the road segment i
p_i	Average signal delay at the road segment i
c	Average time (s) to display all traffic signal indications
g	Average green signal time (s)
x_i	The degree of saturation on the road segment i
l_i	Number of lanes at the road segment i
T_{ij}	Random variable of total delay for a vehicle to travel from port i to port j
Δt	Beacon interval
$t_{ij}(w)$	Time for vehicle w exiting port j after entering port i
$\Psi_{w,i}$	The ith pseudonym of vehicle w
I_w	Instantaneous neighboring node density of vehicle w (or neighboring vehicle count)
V_w	Cartesian coordinates of the velocity vector of a given vehicle w
U_n	Unit vector at direction n

and the variance of the vehicles' speed. Our proposed DLP scheme outperforms both AMOEBA (with random silent period) [7] and Mix-Zone [8] schemes in reducing the probability of successful tracking by an adversary.

This paper is organized as follows. The related work is summarized in Section 2. The system model is described in Section 3. Our proposed DLP scheme is described in Section 4. Performance evaluations and comparisons are presented in Section 5. Conclusions are given in Section 6. A list of the notations that we used in this paper is shown in Table 1.

2 Related work

In this section, we summarize the recent work on location privacy enhancement schemes. In [7], a protocol called AMOEBA is proposed which can mitigate the unauthorized location tracking of vehicles by using the concept of group navigation for V2R communications and by introducing the random silent period between update of pseudonyms [9] for V2V communications. Since the mobility of vehicles is spatially restricted and dependent, vehicles in geographical proximity can navigate as a group. These vehicles have the same average speed and moving direction over a period of time. Since each group can be represented by a single vehicle, which is a group leader, the location of other vehicle is protected from disclosing. When other vehicles are joining the network, AMOEBA can prevent an adversary to link a new pseudonym to the previous one by enforcing a target vehicle to remain silent for a random period.

In [8], the road network is divided into the observed zones and the unobserved zones from the viewpoint of the adversaries. The observed zones are those areas where the adversaries can track the locations of the target vehicles. The unobserved zones (also called the *mix zones*) are some predetermined locations (e.g., road intersections) where the vehicles vary their directions, speeds, and their pseudonyms. The adversaries would have difficulty in linking the vehicles that emerge from the mix zone to those that entered it earlier. Since the locations of mix zones are predetermined, the adversaries may still attempt to eavesdrop on transmissions originating from the mix-zone area.

The concept of location k-anonymity [10] is proposed for protecting the location information through spatial and temporal cloaking. In spatial cloaking, a vehicle broadcast its coarse-grained spatial range information when the number of vehicles within its range is greater than a certain threshold. In other words, the real location of vehicle hides somewhere within a threshold. In temporal cloaking, the beacon message will not be broadcast by the vehicle until a certain number of other vehicles have visited the same location. It refers to replacing a time point associated with a target vehicle with a time interval.

In [11], a protocol is proposed to create the CMIX (Cryptographic MIX-zone) at road intersection. Each vehicle obtains a public/private key pair from certificate authority (CA) via the roadside equipment, and utilizes these keys to encrypt all messages while they are within the mix zone. By key forwarding mechanism, vehicles, which are approaching to the mix-zone, can obtain the symmetric key directly from roadside equipment.

Since vehicles use multiple addresses simultaneously on the different layers, changing the pseudonym only at one layer implies the risk that an adversary can link two consecutive pseudonym by the unchanged address at another protocol layer. In [12], the multi-layer addressing scheme is proposed to support pseudonymity enabling user privacy at the different layers. It is a study of practicability in pseudonymity deployment and implementation. In [13], the effects of pseudonym changes on the performance of network layer are investigated in VANET environments. Simulation results show that pseudonym changes can affect routing procedure and result in packet losses. By introducing a callback mechanism from the link layer, the network layer can better cope with the pseudonym changes.

3 System model

We define the k_i-density zone of vehicle w as the area where at least, on average, $k_i - 1$ distinct neighboring vehicles always exist around w. The density zone i consists of m_i ports and n_i intersections. All vehicles can enter and exit the density zone only via these ports. An intersection is a road junction where two or more roads either meet or cross. Figure 1 shows an example of three density zones with different values of m_i and n_i. To guarantee a secure key management, public key infrastructure (PKI) with certificate authority (i.e., trusted third party) is assumed.

We study the privacy protection of the vehicle operation under a global passive adversary (GPA). GPA aims to locate and track the target vehicles within a region-of-interest by eavesdropping on their authenticated safety broadcast messages with verifiable identity and location information. GPA leverages the deployed infrastructure (e.g., WiFi network) and

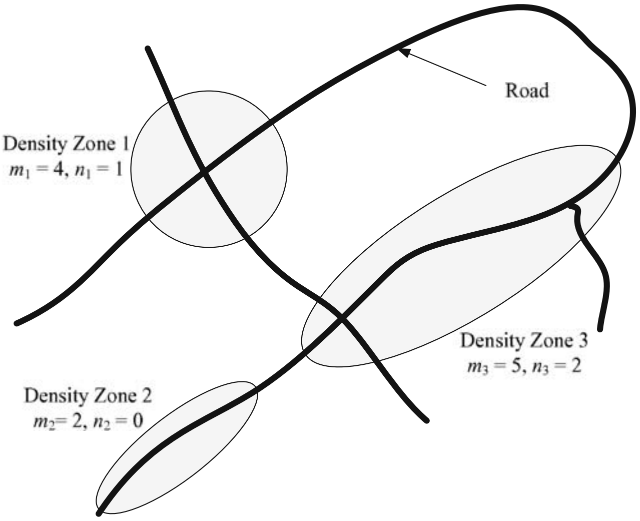

Fig. 1 An example of a VANET with three density zones. The *solid black lines* denote the road. The *shaded areas* correspond to the density zones

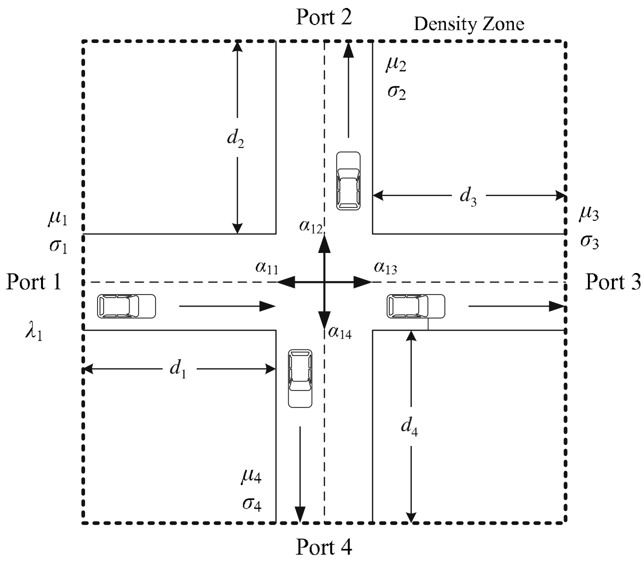

Fig. 2 Topology of a road intersection

utilizes the adversarial roadside equipment deployed to track the movement of the target vehicles within the region-of-interest. Compared to the GPA, the restricted passive attacker (RPA) proposed in [7] is aware of the number of vehicles entering and exiting a density zone, but not of the traveling time and trajectories of the vehicles. In this paper, we only consider the privacy protection against GPA.

Although the GPA cannot distinguish the target vehicle from other vehicles within the density zone due to the change of pseudonym, it can still eavesdrop on all the broadcast messages within the density zone. By installing radio receivers at opportune locations, the GPA can observe entering and exiting events of vehicles where an event is a pair consisting of a port number and a time stamp. In addition, a GPA can either measure (via extensive real measurements [8]) or estimate the probability distribution of the delay of vehicles within the density zone. Given the delay distribution, a GPA can attempt to link an entering vehicle and an exiting vehicle with certain success probability. Thus, GPA should have a knowledge of the delay characteristics of the intersection and on the trajectory of the vehicles in the density zone. In the following subsections, we describe how the GPA obtain the delay distributions via estimation.

3.1 Road traffic model

The topology of a road intersection is shown in Fig. 2. After entering the density zone via port i, each vehicle travels at a distance d_i with constant speed S_i which is chosen independently from a normal distribution $f_{S_i}(s_i)$ with mean μ_i and variance σ_i^2 as:

$$f_{S_i}(s_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(s_i-\mu_i)^2}{2\sigma_i^2}}, \qquad s_i > 0. \tag{1}$$

From an empirical study on the real freeway traffic (i.e., 5-lane highway at California for 24 h) [14], we can assume that the inter-arrival time A_i of vehicles to port i has an exponential distribution $f_{A_i}(a_i)$ with parameter λ_i:

$$f_{A_i}(a_i) = \lambda_i e^{-\lambda_i a_i}, \qquad a_i > 0, \tag{2}$$

where the average arrival rate λ_i (vehicles/s) can be estimated via traffic flow measurement. Thus, vehicles arrive at the port i according to a Poisson process with rate λ_i. At the intersection, each vehicle chooses the output port j with probability α_{ij} where

$$\sum_{j=1,\ j\neq i}^{m} \alpha_{ij} = 1, \tag{3}$$

where m is the number of ports. In this paper, we assume that vehicles cannot enter and exit the density zone via the same port (i.e., $\alpha_{ii} = 0$).

3.2 Delay model in a density zone

In this section, we determine the probability density function (pdf) of the total delay of a vehicle from entering port i to exiting port j. From Fig. 3, when the vehicle is on the road segment i, it moves at a constant

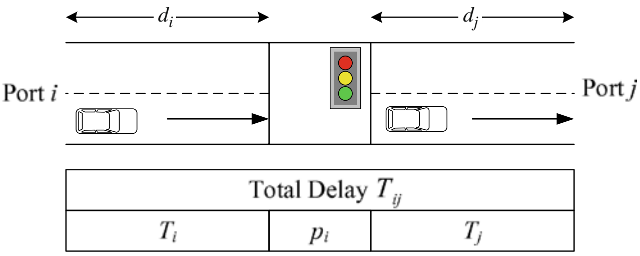

Fig. 3 Total delay in a density zone

speed S_i chosen independently from Eq. 1. Given the distance of the i^{th} road segment d_i, the duration for traveling on this road segment $T_i = d_i/S_i$. Thus,

$$\begin{aligned} P(T_i \le t_i) &= P(S_i \ge d_i/t_i) \\ &= 1 - P(S_i \le d_i/t_i). \end{aligned} \tag{4}$$

By taking the derivative with respect to t_i, we have

$$f_{T_i}(t_i) = \frac{d_i}{t_i^2} f_{S_i}\left(\frac{d_i}{t_i}\right). \tag{5}$$

By substituting Eq. 1 into Eq. 5, we obtain

$$f_{T_i}(t_i) = \frac{d_i}{\sqrt{2\pi}\sigma_i {t_i}^2} e^{\frac{-(d_i/t_i-\mu_i)^2}{2\sigma_i^2}}, \qquad t_i > 0. \tag{6}$$

The delay characteristics of a density zone depend on the road intersection. The *average signal delay* p_i is the time it takes for a vehicle from port i waiting at the intersection for the traffic light to turn green. We choose the widely used approximated average signal delay formula [15, 16] from a combination of theoretical and numerical simulation approaches. This model is reasonably simple and can cover a variety of real world conditions:

$$\begin{aligned} p_i = {} & \frac{c(1-g/c)^2}{2\,(1-(g/c)x_i)} + \frac{{x_i}^2}{2\lambda_i(1-x_i)} \\ & -0.65\left(\frac{c}{{\lambda_i}^2}\right)^{1/3} {x_i}^{2+5(g/c)}, \end{aligned} \tag{7}$$

where c denotes the average time (s) to display all traffic signal indications (i.e., green, red, and yellow) at an intersection, g denotes the average green signal time (s), x_i degree of saturation on road segment i, and λ_i denotes the arrival rate (vehicles/s) at port i. Given the estimated values of c, g, x_i, and λ_i, the average signal delay p_i can be determined and be considered as a constant in the subsequent derivation. The degree of saturation x_i is calculated based on the following equation:

$$\begin{aligned} x_i &= \frac{\text{Demand}}{\text{Capacity}} \\ &= \frac{\lambda_i}{l_i\,\xi} \times \frac{c}{g} \end{aligned} \tag{8}$$

where l_i denotes the number of lanes at road segment i, and ξ is the the maximum number of vehicles per lane per second [17].

As shown in Fig. 3, the total delay that a vehicle experienced in the density zone from entering port i to exit port j, denoted by T_{ij}, is

$$T_{ij} = T_i + p_i + T_j. \tag{9}$$

The cumulative distribution function (cdf) of T_{ij} is

$$\begin{aligned} F_{T_{ij}}(t) &= P(T_{ij} \le t) \\ &= P(T_i + T_j \le t - p_i). \end{aligned} \tag{10}$$

Therefore, it can be obtained as follows [18, Ch. 3]:

$$F_{T_{ij}}(t) = \int_0^{t-p_i} \int_0^{t-p_i-t_i} f_{T_{ij}}(t_i, t_j) dt_i dt_j, \; t > p_i. \tag{11}$$

Since the random variables T_i and T_j are independent and identically distributed (i.i.d.), we have $f_{T_{ij}}(t_i, t_j) = f_{T_i}(t_i) f_{T_j}(t_j)$. Thus,

$$\begin{aligned} F_{T_{ij}}(t) &= \int_0^{t-p_i} \left(\int_0^{t-p_i-t_i} f_{T_j}(t_j) dt_j \right) f_{T_i}(t_i) dt_i \\ &= \int_0^{t-p_i} F_{T_j}(t - p_i - t_i) f_{T_i}(t_i) dt_i, \quad t > p_i. \end{aligned} \tag{12}$$

Equation 12 allows us to determine the cdf of T_{ij}, given the average signal delay p_i and the pdf of random variable T_i. From Eqs. 6 and 12, the pdf of T_{ij} (i.e., $f_{T_{ij}}(t)$) can be determined numerically. From Eq. 12, the average delay for a vehicle to travel from port i to port j is

$$E[T_{ij}] = E[T_i] + p_i + E[T_j]. \tag{13}$$

3.3 Tracking of target vehicle

An adversary can utilize the location information of the target vehicles to infer details about the traveling pattern of the individuals. By installing radio receivers at some sections of the roads, an adversary can eavesdrop on the beacon messages sent by the vehicles with VANET capabilities.

We assume that the adversary has information of the system model of the density zones. It can also observe the entering and exit events corresponding to vehicles entering and exiting the density zone, respectively. An *entering event* consists of the port where the vehicle entered the density zone, and the time when it happened. Similarly, an *exit event* consists of the port where the vehicle left the density zone, and the time when it happened.

The objective of an adversary is to relate exit events to entering events. Specifically, in our model, the adversary selects a target vehicle w and tracks its movement until it enters the density zone. Within the density zone, the target vehicle may change its pseudonym if it satisfies the criteria in the DLP scheme (as explained in the next section). The adversary then needs to determine which vehicle coming out from one of the exiting ports corresponds to the target vehicle.

In the following, we denote the port at which the target vehicle w entered the density zone by i. Without loss of generality, at time 0, the adversary observes that the target vehicle w entered the density zone via port i. The adversary observes the exit events at all ports $j \in \mathcal{J}$ for a time T_{max} such that the probability that the target vehicle w leaves the density zone before T_{max} approaches to 1. Since each vehicle travels a road segment with a constant speed, which is chosen independently from a normal distribution, about 99.7% of vehicles are within three standard deviations 3σ away from the mean μ in terms of speed. To be more precise, the area under the bell curve between $\mu - 3\sigma$ and ∞ in terms of the normal distribution function is about 0.9985. Therefore, in Fig. 3, the value of T_{max} can be defined as

$$T_{max} = \left(\frac{d_i}{\mu_i - 3\sigma_i}\right) + p_i + \left(\frac{d_j}{\mu_j - 3\sigma_j}\right). \tag{14}$$

The adversary records the time $t_{ij}(w_r)$ for each vehicle w_r which exits port j before T_{max}. Let $\{w_1, \ldots, w_R\} \in \mathcal{W}$ be the set of vehicles observed during the time interval $(0, T_{max})$. We propose two selection rules for an adversary to choose a vehicle $w' \in \mathcal{W}$ to be the target vehicle w.

Rule 1 The adversary chooses a vehicle which minimizes the time difference between the average delay $E[T_{ij}]$ to the exit time t_{ij} of all candidate vehicles:

$$\left(w', j'\right) = \arg \min_{\{w_1,\ldots,w_R\} \in \mathcal{W},\ j \in \mathcal{J}} \alpha_{ij} \left(t_{ij}(w_r) - E[T_{ij}]\right)^2 \tag{15}$$

where the multiplication of α_{ij} gives a different weight value depending on the direction of vehicle at the intersection.

Rule 2 The adversary chooses a vehicle which can maximize the cdf of T_{ij} between $t_{ij} + \delta$ and $t_{ij} - \delta$:

$$\left(w', j'\right) = \arg \max_{\{w_1,\ldots,w_R\} \in \mathcal{W},\ j \in \mathcal{J}} \alpha_{ij}(\mathcal{P}_{ij}(w_r) - \mathcal{Q}_{ij}(w_r)), \tag{16}$$

where $\mathcal{P}_{ij}(w_r) = F_{T_{ij}}(t_{ij}(w_r) + \delta)$, and $\mathcal{Q}_{ij}(w_r) = F_{T_{ij}}(t_{ij}(w_r) - \delta)$.

The range value δ is a configurable parameter. The adversary is successful in tracking the target vehicle if the selected vehicle w' is indeed the target vehicle w.

4 Density-based location privacy (DLP)

In DLP scheme, each vehicle can provide connectivity information by broadcasting local beacon messages periodically. A beacon message is a short packet with the current pseudonym and location information of the vehicle. At every beacon interval Δt, the vehicle checks whether it has sent a broadcast within the last Δt (e.g., the default beacon interval in 802.11-based networks is 100 ms [1] and the default HELLO message interval in AODV [19] is 1000 ms). If it has not, it will broadcast a beacon message with time-to-live (TTL) value equals to 1.

Each vehicle w determines its *instantaneous neighboring node density* I_w, (or *neighboring vehicle count*) by listening for beacon messages from its set of neighbors. If a vehicle w has not received a beacon message from another neighbor for Δt, the vehicle w assumes that the link to its neighbor is lost. Therefore, the vehicle w decreases its neighboring vehicle count by 1. On the other hand, whenever vehicle w receives a beacon message from a new neighbor (i.e., with a new pseudonym), w will increase its neighboring vehicle count by 1. Due to the change of pseudonym of neighboring vehicles, the neighboring vehicle count can be greater than the real density for Δt at maximum. By using neighboring node density, each vehicle w can update the average neighboring node density at every beacon interval Δt. DLP uses the arithmetic average neighboring node density, which is the sum of I_w's divided by the number of observations. DLP can also use an exponential weighted moving average (EWMA), which applies weighting factors. The weighting for a previous value decreases exponentially, giving much more importance to recent observations.

If the deployed network layer protocol does not support the exchange of beacon messages, each vehicle can maintain accurate information about its continued connectivity to its neighboring vehicles by using either link layer or other network layer mechanisms. Any adequate link layer notification, such as those provided by IEEE 802.11, can be used to determine connectivity. For example, the absence of a link layer feedback or failure to receive a clear-to-send (CTS) after sending a request-to-send (RTS) may indicate the loss of the link to its neighboring vehicle. If link layer notification is not available, passive acknowledgment can be used in network layer when the neighboring vehicle is expected to forward the packet, by listening to the channel for a transmission attempt made by the neighboring vehicle. If transmission is not detected within a predefined timeout value, an Internet control message protocol (ICMP) [20] echo request message can also be sent to the target neighboring vehicle. If a link to the neighboring vehicle cannot be detected by any of the above methods, the vehicle assumes that the link is lost.

We assume each vehicle w has been preloaded with Z different pseudonyms $\{\psi_{w,1}, \psi_{w,2}, \ldots, \psi_{w,Z}\}$, where Z is a large number. Pseudonyms can either be a set of public keys, network layer or link layer addresses. Many protocols need to commit to a sequence of random values. For this purpose, we use a one-way hash function [21] to generate a one-way chain. One-way chains are a widely-used cryptographic primitive [22]. Each vehicle chooses a random initial value $\psi_{w,Z}$ and generates a one-way chain by repeatedly computing a one-way hash function H on this starting value:

$$\begin{aligned}
\psi_{w,1} &= H(\psi_{w,2}, PSK) \\
&= H^2(\psi_{w,3}, PSK) \\
&\ \ \vdots \\
&= H^{S-2}(\psi_{w,Z-1}, PSK) \\
&= H^{S-1}(\psi_{w,Z}, PSK) \qquad (17)
\end{aligned}$$

where PSK is a pre-shared key with the trusted third party.

Before joining the VANET, each vehicle registers with the trusted third party. We can verify any element of the chain through $\psi_{w,1}$. For example, to verify that element $\psi_{w,k}$ is indeed the element with index k of the hash chain, we check that $H^{k-1}(\psi_{v,k}, PSK) = \psi_{w,1}$. Due to the lack of PSK, attackers cannot link any two pseudonyms from the same vehicle. Moreover, this pseudonym chain can also be used for non-repudiation purpose. The third party authority can provide proof of the integrity and origin of pseudonyms of each vehicle in case of a dispute.

Depending on the frequency of pseudonym update, the size of chain can be large. For example, when the pseudonym update interval is 1 s, each vehicle needs to store 3600 pseudonyms for an hour driving. It may cause an issue of storing lots of pseudonyms. Since hash chain can be used to compute any other element on demand simply by storing $\psi_{w,Z}$, it is efficient in terms of memory usage. The number of hash function evaluations can vary depending on its implementation. For example, the sender can compute approximately 10^6 MD5 hash function evaluations per second (i.e., software-based implementation) [23]. Moreover, the speed of MD5 hash function evaluation can be accelerated with the help of simple stand-alone hardware [24].

A change of pseudonym is triggered by vehicles only when the average neighboring vehicle count is more than or equal to $k_i - 1$. In other words, DLP prevents a privacy breach by ensuring that each vehicle triggers a pseudonym change only when there are at least $k_i - 1$ neighboring vehicles on average.

4.1 Extension of DLP with vehicle-heading (VH)

DLP can be extended by including any information that is used to characterize the situation of either density zone or other neighboring vehicles, such as the moving direction and relative speed. In order to calculate both the speed and direction of neighboring vehicles, distance during two consecutive beacon messages is measured from the same neighboring vehicle. Since every vehicle broadcasts its own location information in periodic beacon messages, all vehicles within the transmission range are able to maintain a status of a neighboring vehicle including its moving direction and relative speed. In our proposed DLP with Vehicle-Heading (VH) scheme, the moving direction of neighboring vehicle is considered in updating neighboring vehicle density.

In DLP with VH, vehicles are grouped into four different groups based on their velocity vectors [25]. In a Cartesian space, each group is characterized by one of the unit vectors, $U_1 = (1, 0)$, $U_2 = (0, 1)$, $U_3 = (-1, 0)$, $U_4 = (0, -1)$, as shown in Fig. 4. Vehicles are assumed to be equipped with a global positioning system (GPS) to detect their geographical location. Location detection is performed every Δt interval. Let $V_w = (v_x, v_y)$ denote the Cartesian coordinates of the velocity vector of a given vehicle w. Using the velocity vector and unit vectors, the group of vehicle w can be decided as follows: Vehicle w belongs to group n, if the inner product of its velocity vector V_w and the unit vector U_n takes the positive value. Vehicle w with negative inner product with the unit vector U_n is heading to the opposite direction, and vehicle w with zero inner product with the unit vector U_n is moving to the orthogonal direction to the vehicles in group n.

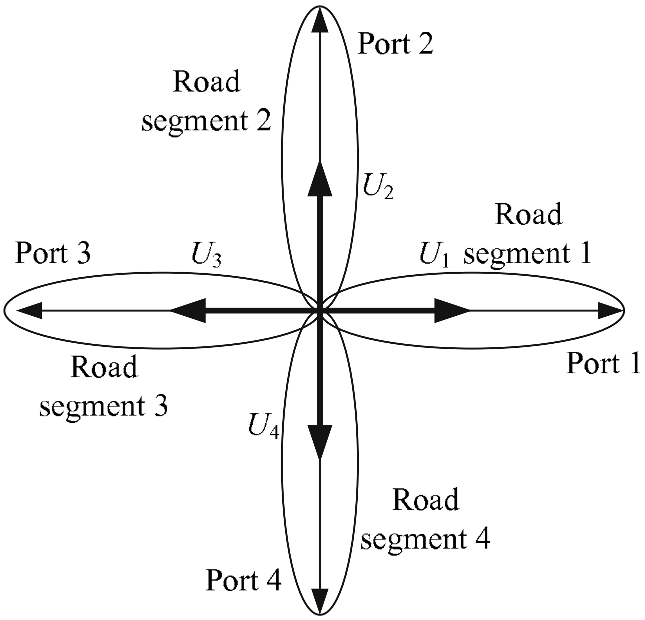

Fig. 4 Velocity vector and vehicle group

From Eq. 3, vehicles are not allowed to enter and exit the density zone via the same port. A vehicle w, entering the density zone via port 1, should not update its neighboring vehicle count when it receives beacon messages from neighboring vehicles heading to the intersection from port 2, 3, and 4. We call these neighboring vehicles as invalid vehicles. On the other hand, if a vehicle receives a new beacon message from neighboring vehicles heading to the intersection from port 1, it should update its neighboring vehicle count by one. Table 2 provides cases for invalid vehicles at the intersection.

Table 2 Cases for invalid vehicles

Entering port of a vehicle w	Invalid vehicles	Inner product
1	Group 1 at road segment 1	< 0
	Group 2 at road segment 4	$= 0$
	Group 4 at road segment 2	$= 0$
2	Group 1 at road segment 3	$= 0$
	Group 2 at road segment 2	< 0
	Group 3 at road segment 1	$= 0$
3	Group 2 at road segment 4	$= 0$
	Group 3 at road segment 3	< 0
	Group 4 at road segment 2	$= 0$
4	Group 1 at road segment 3	$= 0$
	Group 3 at road segment 1	$= 0$
	Group 4 at road segment 4	< 0

4.2 Selection of performance metrics

There are several metrics to quantify the level of privacy achieved in VANETs. For example, the level of privacy is often measured by using the *anonymity set size* (i.e., the entropy [26] of the distribution of vehicles) of the target network. However, it is an appropriate privacy metric only if each vehicle is *equally likely* to be a tracking target. As shown in Section 5, it is not the case in our simulation model. The *successful tracking time* by an adversary can also be used to quantify the level of privacy [6]. An adversary can link two consecutive pseudonyms when it detects that only one vehicle within the zone has changed the pseudonym. In our proposed scheme, each vehicle changes its pseudonym in the density zone only when there are at least $k_i - 1$ neighboring vehicles. Therefore, the successful tracking time is not a proper performance metric when k_i is greater than 1.

The performance metric in our model is the *probability of successful tracking of a target vehicle by an adversary* when making its decision as described in Section 3.3. If the success probability is large, the density zone and changing pseudonyms are ineffective. On the other hand, if the success probability is small, then tracking is difficult and the system ensures location privacy. The probability of successful tracking cannot be determined analytically due to the complexity of our model. Therefore, we ran simulations to determine its empirical value in realistic situations. The simulation setting and parameters as well as the simulation results are presented in the next section.

5 Performance evaluation

In this section, we evaluate the achievable location privacy under various traffic conditions. We first evaluate the performance of our proposed DLP scheme. We then compare the performance between our proposed DLP scheme with Mix-Zone [8] and AMOEBA with random silent period [7] schemes. Table 3 provides a summary of the simulation parameters. The ns-2 [27] simulator is used for the implementation of our proposed scheme. Using SUMO [28], all the necessary files for the network topology, traffic signal logic, and mobility models for the corresponding density zones. We extended several modules in SUMO so that it can support both the normal distribution of vehicle speed and the Poisson arrival of vehicles. Using TraNS [29], SUMO car movement file is converted to the ns-2 mobility file. The ns-2 source code is also modified to count the number of neighbors with varying beacon

Table 3 Simulation parameters

Parameter	Value
Arrival rate λ_1 (arrivals/s)	0.005 ∼ 0.180
Average speed $\mu_1, \mu_2, \mu_3, \mu_4$ (m/s)	14
Variance $\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$	1, 3, 5, 7, 9
Distance d_1, d_2, d_3, d_4 (m)	500
Probability $\alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14}$	0, 0.2, 0.6, 0.2
Acceleration factor (m^2/s)	2
Cycle length c (s)	60
Average green signal time g (s)	30
Degree of saturations x_i	0.02 ∼ 0.72
Number of simulations	100
Range value δ (s)	3

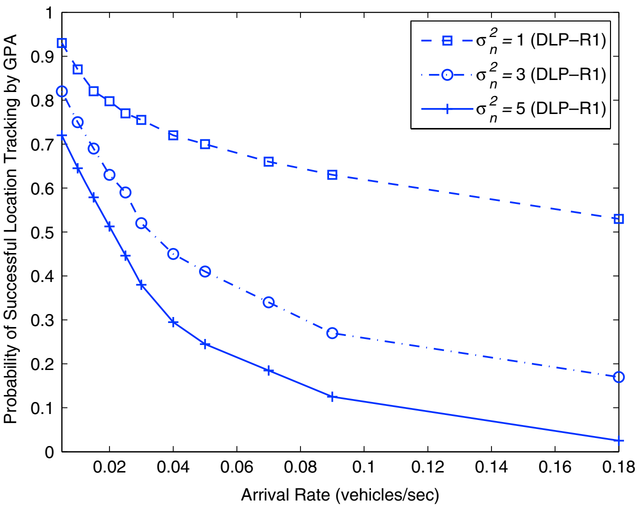

Fig. 5 Probability of successful location tracking by an adversary under different arrival rate and variance for DLP-R1

intervals. The number of vehicles is 100. Each simulation run takes 30,000 simulated seconds. The average speed of vehicles is 14 m/s (i.e., 50.4 km/h). In SUMO, a *deceleration factor* (2 m^2/s) is set when the vehicle is approaching the red traffic light at the intersection. An *acceleration factor* (2 m^2/s) is set when the traffic light changes from red to green and the vehicle needs to accelerate to reach the chosen speed. For medium access control, the IEEE 802.11 distributed coordination function is used. The nominal data rate is 2 Mb/s and a transmission radio range is 250 m. The propagation model combines both free space propagation model and two-ray ground reflection model. We performed multiple independent simulation runs to obtain an estimation of the probability of successful tracking of a target vehicle by a global passive adversary (GPA).

5.1 Performance of DLP

In the following, we outline the topology and mobility model for the performance evaluation of our proposed DLP scheme.

- *Network Topology*: A density zone is composed of one intersection and 4 road segments in each direction (see Fig. 2). Each segment has one lane that prevent following vehicles from passing preceding ones.
- *Mobility Pattern*: All vehicles within the density zone are assumed to travel at a constant speed given by Eq. 1 at each segment. At the intersection, all vehicles experience the delay based on the signal logic of intersection.

Figure 5 shows the success probabilities of the location tracking when both the arrival rate and variance σ_i^2 of vehicles' speed vary. Each vehicle uses the DLP scheme. An adversary applies Rule 1 and Eq. 15 to detect a target vehicle. In other words, for each exiting vehicle, the adversary chooses a vehicle which can minimize the time difference between the average delay to the exit time of all candidate vehicles. We call this scenario as DLP-R1. Each curve matches a different value of σ_i^2. Results indicate that the success probability of the adversary decreases as the variance of vehicles' speed increases. The main reason is that the total delay is inversely proportional to the speed of vehicles. Therefore, as the variance of vehicles' speed increases, the variance of total delay decreases. This makes it difficult to find a target vehicle with the highest probability as the variance of vehicles' speed increases.

Figure 6 also shows the success probabilities of the location tracking when both the arrival rate and variance σ_i^2 of vehicles' speed vary. Each vehicle still uses

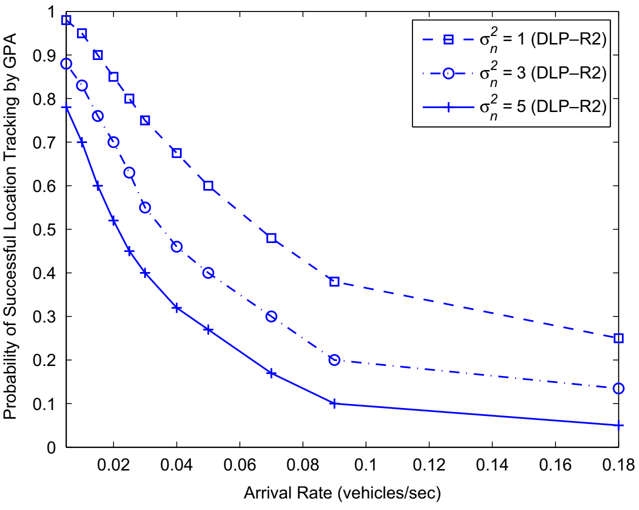

Fig. 6 Probability of successful location tracking by an adversary under different arrival rate and variance for DLP-R2

the DLP scheme. This time, the adversary applies Rule 2 and Eq. 16 to detect a target vehicle. For each exiting vehicle, the adversary chooses a vehicle which can maximize the cdf of the total traveling time in a given interval. We denote this scenario as DLP-R2. Each curve corresponds to the different value of σ_i^2. Results in Figs. 5 and 6 show that DLP-R1 scheme outperforms DLP-R2 scheme as the arrival rate increases. It shows that the accuracy of derived cdf decreases as the arrival rate increases.

5.2 Performance comparison with other schemes

The topology for the performance comparison among AMOEBA, Mix-Zone, and DLP is shown in Fig. 7. There are three density zones with different values of AR, which is the arrival rate of vehicles in density zone. Each vehicle is allowed to change its own pseudonym only one time during the whole travel. For example, if a vehicle w changes its pseudonym in the density zone #1, it cannot change its pseudonym in the other two density zones. In AMOEBA, each vehicle can change its pseudonym only when there are new neighboring vehicles joining the density zone via the entrance ramp. After a silent period chosen randomly between 0.1 to 3 s (recommended values in [7]), each vehicle can update its own pseudonym. In Mix-Zone scheme, each vehicle changes its pseudonym in any density zone with the same probability of 1/3. In our proposed DLP scheme, each vehicle changes its pseudonym in the density zone only when there are at least $k_i - 1$ neighboring vehicles on average. The value of k_i is set to 10 in the simulation.

Figures 8 and 9 show the success probabilities of location tracking by an adversary between AMOEBA, Mix-Zone, and DLP in multiple density zones. In Fig. 8, each vehicle either uses the DLP, AMOEBA, or Mix-Zone schemes. An adversary uses Rule 1 and Eq. 15 to detect a target vehicle. We call the respective scenario as DLP-R1, AMOEBA with Rule 1 (AMOEBA-R1), and Mix-Zone with Rule 1 (Mix-Zone-R1). On the other hand, in Fig. 9, an adversary uses Rule 2 and Eq. 16 to detect a target vehicle. We call the corresponding scenarios as DLP-R2, AMOEBA with Rule 2 (AMOEBA-R2), and Mix-Zone with Rule 2 (Mix-Zone-R2), respectively. Since DLP can choose the density zone where the average number of neighboring vehicles (or the average neighboring vehicle density) is

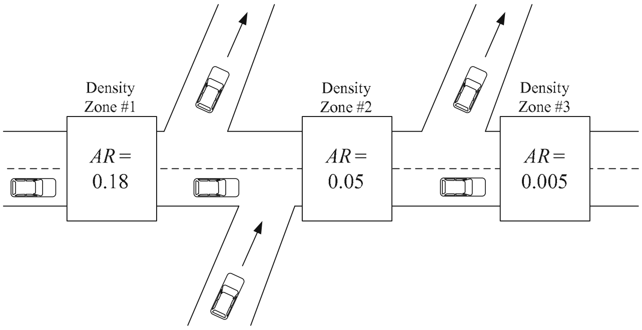

Fig. 7 Topology for performance comparison between AMOEBA [7], Mix-Zone [8], and our proposed DLP scheme. AR denotes the arrival rate of vehicles in a density zone

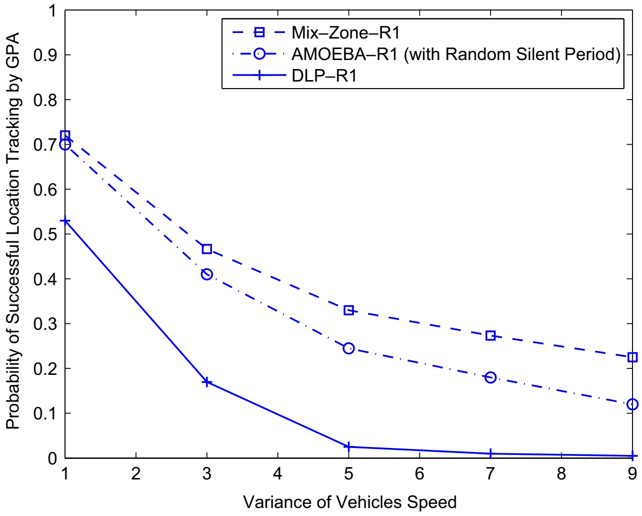

Fig. 8 Probability of successful location tracking by an adversary under different arrival rates when it uses Rule 1

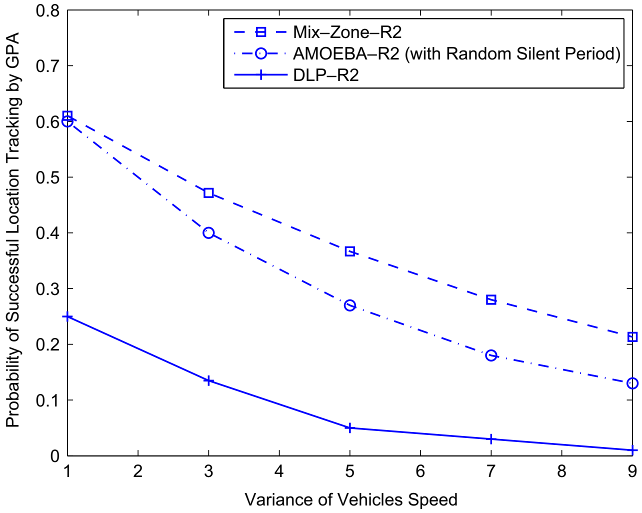

Fig. 9 Probability of successful location tracking by an adversary under different arrival rates when it uses Rule 2

greater than or equal to k_i, the probability of successful location tracking of a target vehicle by an adversary is lower than those of both AMOEBA and Mix-Zone schemes. Although AMOEBA can provide unlinkability between the new and old pseudonyms by using a random silent period, it cannot always find the density zone with k_i neighboring vehicles on average.

6 Conclusion

In this paper, we studied the effectiveness of changing pseudonyms to provide location privacy in VANETs. The approach of changing pseudonyms to make location tracking more difficult was proposed in prior work, but its effectiveness has not been investigated in either an analytical or numerical manner. In order to tackle this issue, we derived a delay model of vehicles in the density zone. We assumed that the adversary has sufficient knowledge (i.e., the delay distribution of the vehicles) in density zone. Based on this information, an adversary may try to select a vehicle which exits the density zone to the target vehicle that entered it earlier. We proposed the vehicle density-based location privacy (DLP) scheme, which can mitigate the location tracking of vehicles by changing pseudonyms based on a threshold in neighboring vehicle count within a density zone. We performed extensive simulations to study the probability of successful tracking of a target vehicle by an adversary under different scenarios. Simulation results showed that our proposed DLP scheme has a better performance than both Mix-Zone and AMOEBA with random silent period in terms of a lower probability of successful tracking by an adversary.

In this paper, we assumed that the frequency of the update of pseudonyms has no effect to the privacy. However, in general, frequent updates of pseudonym may give an advantage to the privacy. On the other hand, the higher the frequency, the larger the cost that the pseudonym updates induce on the system in terms of the design of communication protocols between layers. Future work will investigate the optimal frequency of the pseudonym updates, and enhance the analytical studies with different intersection delay models and service time distributions.

Acknowledgements This research is funded in part by AUTO21, a member of the Network of Centres of Excellence of Canada program, and by Nokia Products Limited, Canada.

References

1. IEEE 802.11 WG (1999) Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications
2. ASTM (2003) Standard specification for telecommunications and information exchange between roadside and vehicle systems—5 GHz band dedicated short range communications (DSRC) medium access control (MAC) and physical layer (PHY) specifications. ASTM E2213-03
3. Myles G, Friday A, Davies N (2003) Preserving privacy in environments with location-based applications. IEEE Pervasive Computing 2(1):56–64
4. Gruteser M, Grunwald D (2003) Anonymous usage of location-based services through spatial and temporal cloaking. In: Proc of ACM int'l conf mobile systems, applications, and services (MobiSys), San Francisco
5. Lei M, Hong X, Vrbsky SV (2007) Protecting location privacy with dynamic MAC address exchanging in wireless networks. In: Proc of IEEE globecom, Washington, DC
6. Garlach M, Guttler F (2007) Privacy in VANETs using changing pseudonyms—Ideal and real. In: Proc of IEEE VTC-Spring, Dublin
7. Sampigethaya K, Li M, Huang L, Poovendran R (2007) AMOEBA: robust location privacy scheme for VANET. IEEE J Select Areas Commun 25(8):1569–1589
8. Buttyan L, Holczer T, Vajda I (2007) On the effectiveness of changing pseudonyms to provide location privacy in VANETs. In: Proc of European workshop on security and privacy in ad hoc and sensor networks (ESAS), Cambridge
9. Huang L, Matsuura K, Yamane H, Sezaki K (2005) Enhancing wireless location privacy using silent period. In: Proc of IEEE WCNC, New Orleans
10. Gedik B, Liu L (2008) Protecting location privacy with personalized *k*-anonymity: architecture and algorithm. IEEE Trans Mobile Comput 7(1):1–18
11. Freudiger J, Raya M, Felegyhazi M, Papadimitratos P, Hubaux J-P (2007) Mix-zones for location privacy in vehicular networks. In: Proc of int'l workshop on wireless networking for intelligent transportation systems (WiN-ITS), Vancouver
12. Fonseca E, Festag A, Baldessari R, Aguiar R (2007) Support of anonymity in VANETs—Putting pseudonymity into practice. In: Proc of IEEE WCNC, Hong Kong
13. Schoch E, Kargl F, Leinmuller T, Schlott S, Papadimitratos P (2007) Impact of pseudonym changes on geographic routing in VANETs. Lect Notes Comput Sci (LNCS) 4357:43–57
14. Wisitpongphan N, Bai F, Mudalige P, Sadekar V, Tonguz O (2007) Routing in sparse vehicular ad hoc wireless networks. IEEE J Select Areas Commun 25(8):1538–1556
15. Wolshon B, Taylor WC (1999) Analysis of intersection delay under real-time adaptive signal control. Transp Res Part C: Emerg Technol 7(1):53–72
16. Wunderlich R, Liu C, Elhanany I, Urbanik T (2008) A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection. IEEE Trans Intell Transp Syst 9(3):536–547
17. Lo HK (2006) A reliability framework for traffic signal control. IEEE Trans Intell Transp Syst 7(2):250–260
18. Ross S (2006) Introduction to probability models, 9th edn. Academic, London
19. Perkins C, Belding-Royer E, Das S (2003) Ad hoc on-demand distance vector (AODV) routing. http://www.ietf.org/rfc/rfc3561.txt

20. Deering S (1991) ICMP router discovery messages. http://www.ietf.org/rfc/rfc1256.txt
21. Rivest R (1992) The MD5 message-digest algorithm. http://www.ietf.org/rfc/rfc1321.txt
22. Perrig A, Canetti R, Song D, Tygar D, Briscoe B (2005) Timed efficient stream loss-tolerant authentication (TESLA): multicast source authentication transform introduction. http://www.ietf.org/rfc/rfc4082.txt
23. Perrig A (2001) The BiBa one-time signature and broadcast authentication protocol. In: Proc of ACM conference on computer and communications security
24. Helion (2009) High performance solutions in silicon—MD5 core. http://www.heliontech.com/md5.htm
25. Taleb T, Sakhaee E, Jamalipour A, Hashimoto K, Kato N, Nemoto Y (2007) A stable routing protocol to support ITS services in VANET networks. IEEE Trans Veh Technol 56(6):3337–3347
26. Serjantov A, Danezis G (2002) Towards an information theoretic metric for anonymity. In: Proc of the workshop on privacy enhancing technoligies (LNCS), San Francisco
27. NS-2 simulator (2009) NS-2 simulator homepage. http://www.isi.edu/nsnam/ns/
28. Simulation of urban mobility (SUMO) (2009) Simulation of urban mobility (SUMO) homepage. http://sumo.sourceforge.net
29. EPFL (2008) Traffic and network simulation environment VANETs (TraNS). http://trans.epfl.ch