Privanet: An Efficient Pseudonym Changing and Management Framework for Vehicular Ad-Hoc Networks

Abdelwahab Boualouache, Sidi Mohammed Senouci, Samira Moussaoui,

PRIVANET: An Efficient Pseudonym Changing and Management Framework for Vehicular Ad-Hoc Networks

Abdelwahab Boualouache , *Member, IEEE*, Sidi-Mohammed Senouci , *Member, IEEE*, and Samira Moussaoui

***Abstract*— Protecting the location privacy is one of the main challenges in vehicular ad-hoc networks (VANETs). Although, standardization bodies, such as IEEE and ETSI, have adopted a pseudonym-based scheme as a solution for this problem, an efficient pseudonym changing and management is still an open issue. In this paper, we propose PRIVANET, a complete and efficient pseudonym changing and management framework. The PRIVANET has a hierarchical structure and considers the vehicular geographic area as a grid. Each cell of this grid contains one or many logical zones, called vehicular location privacy zones (VLPZs). These zones can easily be deployed over the widespread roadside infrastructures (RIs), such as gas stations, to provide a secure changing and management of pseudonyms. The proposed framework consists of different building blocks: 1) an effective VLPZ-based pseudonym changing strategy; 2) a reputation-based mechanism to motivate selfish vehicles to enter VLPZs; 3) an adapted user-centric privacy model; 4) a secure hybrid mechanism for the distribution of pseudonyms sets and CRLs; 5) a method to generate the IP and MAC addresses from the pseudonym; 6) a stochastic model to estimate the number of VLPZs required at a given cell; and 7) a mathematical model for an optimal placement of the VLPZs over RIs to reduce the transportation cost of vehicles in terms of time. An extensive simulation study using a realistic map and with real traffic mobility measurements is carried out to evaluate and validate the performance of the PRIVANET. The simulation results demonstrate the effectiveness of the proposed framework.**

***Index Terms*— Vehicular ad-hoc networks (VANETs), security, location privacy, pseudonym changing.**

I. Introduction

AS PART of the intelligent transportation system (ITS), Vehicular Ad-Hoc Networks (VANETs) have witnessed a huge interest both in academia and industry. This technology is mainly intended to make future transportation systems safer and comfortable via two main categories of applications [1]: safety-related applications such as emergency reporting and collision warning, and non-safety-related applications such as Internet access and location-based services.

A. Boualouache and S. Moussaoui are with the RIIMA Laboratory, Department of Computer Science, University of Sciences and Technology Houari Boumediene (USTHB), Algiers 16025, Algeria (e-mail: webwahab@gmail.com).

S.-M. Senouci is with DRIVE EA1859, Université Bourgogne Franche-Comté, F58000 Nevers, France.

A beaconing service is often required by safety-related applications. Every vehicle regularly broadcasts a safety message, which includes its status information such as time, position, and speed [2]. Safety messages are authenticated and sent with a high frequency in clear text. They significantly help in road safety by providing a vicinity view to vehicles. However, these messages can also be exploited by passive adversaries to perform location tracking attack. Indeed, the broadcast messages can easily be collected by eavesdropping wireless communications and linked after that according to vehicles' identifiers. Thereby, all locations visited by vehicles can be known by adversaries, which threatens the location privacy of VANETs' users since there is usually a strong relationship between users and their vehicles [3].

Pseudonym changing is a common approach to overcome this problem [4]. Instead of using one static identifier, each vehicle is equipped with a set of fictive identifiers, called pseudonyms. Current security standards such as IEEE 1609.2 [5] and ETSI 102941-v1.1.1 [6], consider pseudonyms as public keys certified by a Certificate Authority (CA) and stored in vehicles' On-Board Units (OBUs). In order to alleviate the tracking of their positions, vehicles should frequently change their pseudonyms. Indeed, a higher frequency of pseudonym changing is an important parameter to provide a good degree of privacy protection. However, the frequency value should be reasonable to avoid the impact communication performances [7]. In addition, to avoid ease linkability of pseudonyms, all the communication identifiers such as the MAC and IP addresses should be changed with the pseudonym at the same time [8].

Although both of academia and industry have agreed to apply a pseudonym changing approach in the future deployment of VANETs, different challenges are still to be addressed [9]. For example (i) many works have demonstrated that vehicle's pseudonyms can easily be linked without using a pseudonym changing strategy [10]. Although, different pseudonym changing strategies have been proposed (e.g. [11]–[14]), an effective pseudonym changing strategy is not achieved yet [15], (ii) the level of location privacy could significantly be affected if rational vehicles refuse to change their pseudonyms with other vehicles [16], and last but not least, (iii) high deployment costs are generated if existing solutions to distribute certificates (pseudonyms) revocation lists (CRLs) and pseudonyms sets are applied. Indeed, these

solutions require a total coverage by Roadside Units (RSUs), which cannot always be achieved [17].

In this paper, we extend our work published in [18] to propose PRIVANET, a complete and efficient pseudonym changing and management framework. This framework has a hierarchical structure and mainly based on logical zones, called Vehicular Location Privacy Zones (VLPZs). These zones aim at managing and changing pseudonyms and can easily be deployed over the widespread Roadside Infrastructures (RIs) such as gas stations, electric vehicles charging stations, and toll booths or created as new roadside infrastructures for the future deployment VANETs. VLPZs are equipped with a reputation-based mechanism for motivating vehicles to enter to them. Different building blocks are proposed by PRIVANET that address each issue of the pseudonym changing approach. For example, the framework includes building blocks for an efficient changing of pseudonyms, an optimal distribution of pseudonyms sets, and a synchronized changing of all identifiers of the communication stack layer. In the other hand, a stochastic model to estimate the number of VLPZs required at a given cell as a function of traffic density and the demand for vehicles to enter VLPZs is proposed. In addition, the problem of the optimal placement of VLPZs over RIs is mathematically modeled. The modularity, the extensibility, the flexibility and the ease of deployment are the main features of this framework, which makes it well adapted for the future deployment of VANETs.

The main contributions of this paper can then be summarized as follows:

- We propose a complete and efficient pseudonym changing and management framework for VANETs. It has a hierarchical structure to facilitate the control and management of the system. The building blocks of our framework address the key issues of the pseudonym changing approach.
- We propose a stochastic model to estimate the number of VLPZs required at a given cell as a function of time. This model is mainly based on traffic density and the demand for vehicles to enter the VLPZs.
- We mathematically model the problem of the optimal placement of VLPZs over widespread RIs to reduce the transportation cost of vehicles in terms of time. We then define an objective function to select the best RIs to host VLPZs and propose a simple solution to solve this problem.
- We evaluate and validate the proposed framework through an extensive simulation study using a realistic map and with real traffic mobility measurements. The simulations are performed based on a reliable vehicular network simulation framework composed of Veins [19], OMNet++ and SUMO [20].

The remainder of this paper is organized as follows. Some related work is described in Section II. The proposed pseudonym changing and managing framework is presented in Section III. Section IV describes the reputation-based mechanism that is used to motivate vehicles to enter VLPZs. An optimal deployment of VLPZs is described in Section V. The results of the performance evaluation are presented in Section VI. Finally, the conclusion is given in Section VIII.

II. Related Work

In this section, we present an overview of recent state-of-art solutions for pseudonym changing and management in VANETs. At the end of this section, a comparison is made to highlight the features of our proposed framework comparing to relevant presented solutions.

Pseudonym linking attack makes a simple changing of pseudonym ineffective to protect the location privacy of VANETs' users [10]. Two kinds of this attack are considered [21]: (i) *syntactic linking* that can be performed if only one vehicle changes its pseudonym among the group of vehicles running on the road, and (ii) *semantic linking* that can be performed despite all vehicles change their pseudonym simultaneously, where adversaries use advanced tracking algorithms to predict the future positions of vehicles based on their current positions, and thereby link their pseudonyms. Different pseudonym changing strategies have been proposed to overcome this attack. These strategies were classified according to the used protection mechanism into three categories: (i) a synchronizing mechanism-based strategies (e.g. [11], [12]), (ii) encryption-based strategies (e.g. [13]), and (iii) radio silence based strategies (e.g. [14]). The strategies of the last category are the most effective as they allow protection again the two kinds of pseudonym linking attack. However, radio silence may have negative impacts on road safety [22]. To address this last issue, we proposed in [23], the VLPZ model and the VLPZ-based pseudonym changing strategy. This strategy provides protection against pseudonym linking attack while preserving road safety. Recently in [21], we proposed a comprehensive survey and classification of pseudonym changing strategies. We also compare and discuss them according to important parameters.

Furthermore, the cooperation between vehicles is important to ensure a successful pseudonym changing. Indeed, rational vehicles may decide to do not change their pseudonyms when they asked for it. This is due to the costs that could be generated from pseudonym changing [7]. To address this issue, a game theoretic based approach is first proposed in [16]. The authors suggested a user-centric privacy model, where vehicles take maximizing their payoffs as a decisive parameter to change their pseudonyms. However, vehicles, in this solution, only cooperate when their location privacy protection level is below the required level, otherwise they will not. Recently, in [13] a novel solution called MPSVLP is proposed. This solution consists of a strategy of pseudonym changing and a mechanism to motivate vehicles to cooperate in this strategy. The strategy consists in creating dynamic mixed zones called, DMPLs. A request is sent to a strategy controller server (CS) each time a vehicle wants to create its own DMPL. The length of vehicle's DMPL area is calculated by CS, which after that sends a command to all vehicles found this area requesting them to cooperate in creating this DMPL. Relying on a encryption-based strategy is, however, a major drawback of this solution as it does not provide any protection again the semantic linking of pseudonyms [21].

TABLE I
A COMPARISON OF PRIVANET WITH THE STATE OF ART OF PSEUDONYM-BASED SOLUTIONS

Solution	PC Strategy	Pseudonym Management	Generating Net Identifiers	Selfishness Protection	Privacy Model
TPSRP [26]	X	X	X		
MPSVLP [13]	X			X	
PMS [24]		X	X		
PRIVANET	**X**	**X**	**X**	**X**	**X**

Pseudonym management is another important issue in the pseudonym changing approach. Indeed, the authors in PMS [24] highlighted the fact that more focus is placed in the literature on pseudonym changing strategy than pseudonym management and implementation issues. For this reason, they proposed a solution that completely relies on RSUs to distribute pseudonym sets and exchanging them to increase the anonymity. This solution could however generates high deployments costs as it is totally based on road infrastructures. In addition, as no strategy proposed is clearly proposed, pseudonyms linking could easily be performed. In [25], we proposed HPDM a hybrid method for the distribution of the pseudonyms sets. This method uses both RSUs and vehicles to distribute pseudonyms sets. The aim is to reduce the deployment costs generated to perform this operation.

The authors in TPSRP [26] viewed pseudonyms as costly resources that need to be considered as services. Therefore, instead of pro-actively issuing pseudonyms, they proposed to provide pseudonyms only to vehicles requesting them. This solution seems interesting as it significantly helps to reduce the number of needed pseudonyms. However, this solution is not aligned with requirements of VANETs as location privacy protection is required and protection mechanisms are basically based on the cooperation of a large number of vehicles [27].

Our main goal in this paper is thus to propose a complete framework that efficiently considers the key issues of the pseudonym changing approach (PCA). Indeed, to the best of our knowledge, there is no solution considers all these issues together in the literature. In Table I, we compare between state-of-art solutions for pseudonym-based system in VANETs. This comparison highlights which of the key issues of PCA are addressed by each solution such as the pseudonym changing strategy (PC Strategy) and the non-cooperative behavior (Selfishness). We can see that, in contrast to state-of-art solutions, all the key issues are addressed by PRIVANET.

III. AN EFFICIENT PSEUDONYM CHANGING AND MANAGEMENT FRAMEWORK FOR VEHICULAR AD-HOC NETWORKS

In this section, we present the design of the proposed framework. Figure 1 illustrates a global view of this framework.

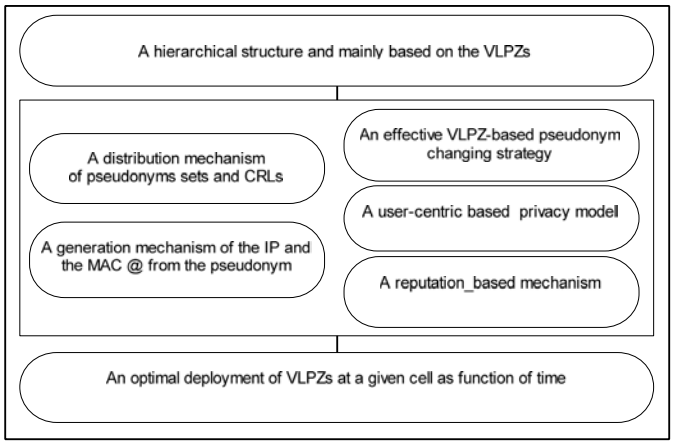


Fig. 1. A global view of the PRIVANET.

It has a hierarchical structure and mainly based on logical zones, called Vehicular Location Privacy Zones (VLPZs) deployed over grid cells of the vehicular geographic area. This framework consists of different building blocks: (i) a VLPZ-based pseudonym changing strategy for an effective changing of a pseudonym, (ii) a secure hybrid mechanism for the distribution of pseudonyms sets and CRLs, (iii) a method to generate the IP and the MAC addresses from the pseudonym, (iv) an adapted user-centric privacy model, (v) a reputation-based mechanism to motivate selfish vehicles to enter VLPZs, and finally (vi) an optimal deployment of VLPZs over RIs of a given cell as a function of time. The first four building blocks [18] are described in this section, while the fifth and sixth ones are presented in sections IV and V respectively.

This section is thus structured as follows. We first describe the considered system and adversary models. We then present the model VLPZ, and the VLPZ-based pseudonym changing strategy. After that, we present the VLPZ-based pseudonyms sets and revocation lists distribution. Finally, we describe the used privacy evolution model and the mechanism of generating the communication layers identifiers from the pseudonym.

A. System and Adversary Models

To facilitate the management of pseudonyms and CRLs, we consider that the vehicular geographic area is partitioned as a grid [18]. The cells of the grid have a same predefined size. Each cell may comprise the entire downtown area of a small town or few city blocks. The VANET system is composed of vehicles and RSUs. Each vehicle has an OBU device that is equipped with a wireless technology based on the IEEE 802.11p/WAVE standard. The OBU allows the vehicle not only to communicate with other vehicles but also with RSUs. Each vehicle is also equipped with a map and a GPS receiver that allows obtaining the position and the current time. Each vehicle periodically broadcasts a safety message every t milliseconds, where each message includes information about the vehicle such as its position and its speed. We also assume the existence of a central trusted authority (TA) that provides public and private keys to vehicles and RSUs. Before joining the VANET, each vehicle registers with the TA with its vehicle identifier, denoted by ID_v. During the registration,

each vehicle V_i is equipped with a public and a private keys and sets of pseudonyms. Each set contains n pseudonyms $K_{i,j}$, where $j \in 1, \ldots, n$. For each pseudonym $K_{i,j}$ of vehicle V_i, the TA provides a certificate $Cert_{i,j}(K_{i,j})$. The private key $K_{i,j}^{-1}$ corresponding to the pseudonym $K_{i,j}$ is used by the vehicle V_i to digitally sign messages. The pseudonym is attached to each message to enable other vehicles and RSUs to verify the sender's authenticity. Each vehicle changes its pseudonym each δ minutes. Each cell contains one regional trusted authority (TA_R), and one or more Vehicular Privacy Zones (VLPZs). TA_Rs act as intermediates between the TA and the VLPZs. They aim at managing the pseudonyms sets and the CRLs distribution and control the location privacy protection level provided the VLPZs within the cells. Indeed, all the TA_Rs are connected to the TA, and each TA_R is contacted to the VLPZs within its cell via secure communication links.

Furthermore, we are interested to study the location privacy protection against a strong passive adversary model. This adversary is passive ie. it can only eavesdrop communications and composed of an external global adversary and few internal local attackers. The global one has a complete coverage on the system but it is not an authenticated member (external). However, the local attackers are internal i.e. authenticated members but with a limited coverage on the system (local).

This adversary model aims to track the target vehicle by eavesdropping all communications of any vehicle within a region of interest. The adversary model is well aware of the system model and the proposed framework design. However, it has no control over VLPZs. In addition, this adversary is not able to perform tracking using cameras, because the cost of the global eavesdropping with cameras is much higher than the radio-based eavesdropping. Therefore, camera-based global eavesdropping is beyond the scope of this paper.

B. VLPZ Model

We define the Vehicular Location Privacy Zone (VLPZ) as a logical zone managed by trusted regional authorities like municipalities or directly by the country transportation department. Each cell of the grid can contain one or more VLPZs. The VLPZ aims not only to increase the location privacy protection level of vehicles within the cell by providing an effective pseudonym changing [23], but also to distribute pseudonyms sets and CRLs to them. The design of VLPZ is seemingly similar to widespread Roadside Infrastructures (RIs) like gas stations. As illustrated in Figure 2, a basic VLPZ consists of one entry point called ***the router***, one exit point called ***the aggregator*** and a limited number of lanes l where $l > 1$. Each VLPZ is equipped with an RSU denoted by RSU_{vz} and used to: (i) periodically announce the existence of a VLPZ, (ii) stimulate vehicles passing through the VLPZ to enter, (iii) request pseudonyms sets from the TA_R and distribute them to the vehicles inside the VLPZ according to their requests, (iv) request the CRLs from the TA_R and distribute them to the vehicles inside the VPLZ, and finally (v) get information from vehicles, which helps the VLPZ to take certain decisions. The VLPZ can easily be deployed over RIs such as gas stations and toll booths. However, due to the increasing interest of users to protect their location privacy, we do not rule out that the VLPZ can be created as an independent roadside infrastructure in the future VANETs deployment. Ideally, VLPZs are installed in the two directions of two-way streets and their emplacements of are shown on the map.

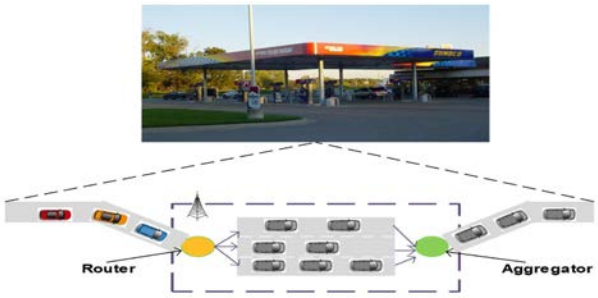


Fig. 2. The VLPZ basic model.

C. VLPZ-Based Pseudonym Changing Strategy

The strategy of pseudonym changing is executed as follows [23]. The RSU_{vz} periodically broadcasts notifications for informing the vehicles about the existence of a VLPZ. If a vehicle wants to access to the VLPZ, it sends a request to the RSU_{vz}. As Figure 2 shows vehicles arrive at a VLPZ, one after another, on one-lane. They keep broadcasting safety messages until they enter the VLPZ. When a vehicle reaches ***the router***, it stops broadcasting safety messages and heads for an assigned VLPZ's lane. The assigned lane is randomly and privately selected by ***the router***. The vehicle can then reside inside a VLPZ for a random period of time. This period mainly depends on the service time. For example, if we assume that a VLPZ is deployed in a gas station, the service time is the time taken by the driver to fill the fuel tank of its vehicle. A vehicle must change its pseudonym before it exits the VLPZ and all vehicles exit a VLPZ through the ***the aggregator***. However, the exit order is different from the entering order since the residency periods of vehicles are random. We also assume that ***the aggregator*** can select a certain order in random and private way. As discussed in [23], this strategy provides protection not only against both of the syntactic and the semantic linking of pseudonyms, but also against the FIFO attacks. In addition, differently from the strategies that rely on the radio silence technique, road safety is preserved in this strategy.

D. Pseudonyms Sets and Revocation Lists Distribution

As we already highlighted in the related work section, existing state-of-art solutions totally rely on available RSUs to distribute pseudonyms sets and pseudonyms revocation lists (CRLs) [24], which may generate high deployment costs. For this reason, in our framework, we propose that these operations are performed both inside VLPZs using RSU_{vz}s and by vehicles through V2V communications. Our solution allows a quick and wide distribution of pseudonyms sets and CRLs, while keeping low deployment costs. Indeed, vehicles can exchange small CRLs updates as described in [28] and involve

to distribute pseudonyms sets using the method proposed in [25] for example.

E. Privacy Level Evolution

The location privacy level of a vehicle changes over the time. It can be decreased due to the pseudonyms linking attack and increased each time that a vehicle accesses to a VLPZ. To capture the evolution that occurs to the location privacy level of a vehicle over time, we use the user-centric location privacy model introduced by [16]. The location privacy level of a vehicle i is modeled using a location privacy loss function $\beta_i(t, T_i^{vz}) : (\mathbb{R}^+, \mathbb{R}^+) \rightarrow \mathbb{R}^+$ where t is the current time and $T_i^{vz} \leq t$ is the time of the last pseudonym change of vehicle i inside a VLPZ. The privacy loss is set to 0, each time that i changes its pseudonym inside a VLPZ and increases with time according to a sensitivity parameter, $0 < \lambda_i < 1$ until it reaches a maximum value $A_i(T_i^{vz})$, which is the location privacy protection level achieved at the last pseudonym change of vehicle i inside a VLPZ. The privacy loss function is defined as follows:

$$\beta_i(t, T_i^{vz}) = \begin{cases} \lambda_i(t - T_i^{vz}) & \text{for } T_i^{vz} \leq t < T_i^{max} \\ \\ A_i(T_i^{vz}) & \text{for } t \geq T_i^{max} \end{cases}$$

where $T_i^{max} = \frac{A_i(T_i^{vz})}{\lambda_i} + T_i^{vz}$ is the time when the function reaches the maximal privacy loss. The location privacy level of vehicle i at time t is:

$$A_i(t) = A_i(T^{vz}) - \beta(t, T^{vz}), t \geq T^{vz}$$

F. Generating the Communication Layers Identifiers

Changing all communication identifiers with the pseudonym is required to avoid the ease linking of pseudonyms [8]. In our framework, the Cryptographically Generated Address (CGA) protocol [29] is used to build IP addresses from pseudonyms. Indeed, as described in [30], CGA uses a random 128-bit number and pseudonym (public key) to create the interface identifier, which is concatenated after that with a subnet prefix to build an IPv6 address. The concept of CGA can also be used to build a MAC address [24]. Indeed, the MAC can be generated by calculating a hash value of a set of concatenated values. The values that we propose in our framework are: a random 128-bit number, an interface identifier, collision count, a pseudonym, and extension fields. The collision count and the extension fields are described in [30].

IV. Motivating Vehicles to Enter to VLPZs

Two important parameters impact the level of location privacy provided a VLPZ: (i) Its capacity (K): that is the maximum number of vehicles that the VLPZ can contain. This parameter is static and can be set by the system designer, and (ii) Its occupancy ($|\text{AS}|_t$), which represents the number of vehicles that have accessed the VLPZ at the same time i.e. the number of vehicles inside a VLPZ at a given time. In contrast to the capacity, the occupancy is a dynamic parameter that mainly depends on road traffic density and the number of vehicles requesting access to the VLPZ. Thus, the occupancy could dramatically be decreased if vehicles are rational. Indeed, these vehicles always tend to protect their location privacy with minimum possible cost. In our framework, the cost is expressed as the time that a vehicle took to move to a VLPZ and quantified by the lost of pseudonyms along this time. To this end, a non-cooperation from these vehicles is expected if they reached their required privacy level (A_d). In other words, they request to access a VLPZ only if their location privacy level goes under A_d.

To overcome this problem, we propose a reputation-based mechanism for VLPZs to increase their occupancy. This mechanism consists in broadcasting invitations of access to vehicles through RSU_{vz}s. The reputation value of each vehicle is calculated according to its response to the received invitation. Indeed, the reputation value of a vehicle will be increased if a VLPZ received a positive response from this vehicle. However, if the vehicle refuses the invitation, its reputation value will be decreased. The reputation value is calculated on the basis of the occupancy of the VLPZ at t_j. t_j represents the time when the vehicle quits a VLPZ, if the vehicle accepts the j^{th} invitation from a VLPZ, otherwise, t_j is the time of refusing the invitation. The following formula gives the reputation value ($\mathbb{R}_i^j$) of a given vehicle i after receiving the j^{th} invitation.

$$\mathbb{R}_i^j = \begin{cases} \mathbb{R}_i^{j-1} + |AS|_{t_j} & \text{if v cooperates} \\ \\ \mathbb{R}_i^{j-1} - |AS|_{t_j} & \text{if v defeats and } \mathbb{R}_i^{j-1} \geqslant |AS|_{t_j} \\ \\ 0 & \text{if v defeats and } \mathbb{R}_i^{j-1} < |AS|_{t_j} \end{cases}$$

where $\mathbb{R}_i^{j-1}$ is the old reputation value of the vehicle i. The reputation value of i increases as much as it cooperates. The accumulated value of reputation is thus used each time that the vehicle i needs an access to a VLPZ.

V. An Optimal Deployment of VLPZs

As we previously pointed out, the design the VLPZ is similar to widespread RIs like gas stations. These RIs are not mainly intended to protect the location privacy of VANETs' users, but it will be interesting to exploit them for this purpose as well. However, the deployment of VLPZs over the widespread RIs should not disturb the purpose to which these RIs are constructed for. For this reason, an optimal placement of VLPZs over existing RIs is required, in such way that the VLPZ is activated at a given RI only when it is needed. In this section, we propose an optimal deployment of VLPZs on the RIs that exist at a given cell as a function of time. We first propose a model to estimate the required VLPZs at a given cell as a function of traffic density and the demand for vehicles enter the VLPZs. We then define an objective function to select the best roadside infrastructures to host VLPZs to reduce the transportation cost of vehicles in terms of time.

A. Required VLPZs as a Function of Time

In this subsection, we provide a stochastic model to estimate the number of vehicles that are looking to enter to a VLPZ

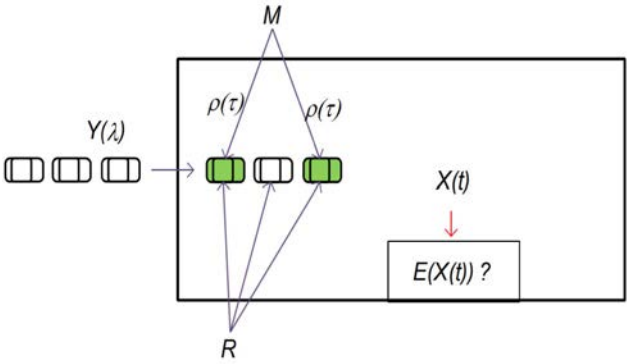


Fig. 3. Cell stochastic model.

as a function of time at a given cell. The aim is to be able to predict the optimal number of VLPZs required at a given cell as a function of time. In the following, we assume a given cell that contains a single VLPZ. We define the stochastic model, illustrated in Figure 3, as follows. Let the vehicle arrival V_A at the cell follows a Poisson Process Y and the inter-arrival time t_A for V_A has an exponential distribution with the mean $1/\lambda$. During its residency in the cell, a vehicle may enter to a VLPZ if its location privacy protection level is below than the required level or to respond to an invitation from a VLPZ. For this reason, we define another random process M, independently of the arrival stochastic process Y. As vehicles enter the cell, M marks some of them with a positive probability $\rho(t)$. $\rho(t)$ models the willingness of a vehicle to enter the VLPZ during its residency period in the cell. Thus, vehicles that are not marked by M, will not enter the VLPZ during their residency in the cell. In addition, $\rho(t)$ is a time-depended probability because it depends on the traffic density in the cell by time. Indeed, the more traffic density is, the more vehicles will enter the VLPZ. We also assume a random process R that specifies the time that will be spent by each vehicle inside the cell. This residency time can then be modeled by a random variable (G) with a distribution function F_G. The vehicles' residency time is independent of the arrival process Y and the marking process M.

At $t = 0$, we assume there is no vehicle inside the VLPZ. Let $\{X(t) \mid t \geq 0\}$ be the counting process that counts the number of vehicles that have the intention to enter the VLPZ during their residency in the cell as a function of time t. We are then interested in $E[X(t)]$: the expected number of marked vehicles resided in the cell at time t. According the law of total probability, we have:

$$Pr[X(t) = k] = \sum_{n \geq 0} Pr[X(t) = k \mid Y(t) = n] * Pr[Y(t) = n]. \tag{1}$$

where $\Pr[X(t) = k \mid Y(t) = n]$ is the conditional probability that $X(t) = k$ given n vehicles enter the cell in the period $[0,t]$. Knowing that the vehicles' arrival follows a Poisson process we have then:

$$Pr[X(t) = k] = \sum_{n \geq 0} \binom{n}{k} [\gamma(t)]^k [1 - \gamma(t)]^{n-k} \frac{[\lambda t]^n}{n!} \exp^{-\lambda t}. \tag{2}$$

$\gamma(t)$ is the probability that a given arrived vehicle marked with a positive probability by M and it stills in the cell at time t. Given that $n \geq 0$ vehicles have arrived in $[0,t]$, where their arrival times are uniformly distributed in $[0,t]$, $\gamma(t)$ is given as follows [26]:

$$\gamma(t) = \frac{1}{t} \int_0^t \rho(t-u)[1 - F_G(u)]\, \mathrm{d}u. \tag{3}$$

where u is the uniform random variable on $[0,t]$. [26] demonstrated that:

$$Pr[X(t) = k] = \frac{[\Lambda(t)]^k}{k!} \exp^{-} \Lambda(t) \tag{4}$$

where $\{X(t) \mid t \geq 0\}$ is a Poisson process with parameter:

$$\Lambda(t) = \lambda \int_0^t \rho(t-u)[1 - F_G(u)]\mathrm{d}u. \tag{5}$$

and with an expectation $E[X(t)]$:

$$E[X(t)] = \Lambda(t) = \lambda \int_0^t \rho(t-u)[1 - F_G(u)]\mathrm{d}u. \tag{6}$$

After estimating the number of vehicles that intend to enter a VLPZ during their residency period, we can then predict the number of VLPZs required in the cell by time, which is denoted by $N_{vlpz}(t)$ and given by the following formula:

$$N_{vlpz}(t) = \frac{E(X(t))}{K_{opt}}. \tag{7}$$

where K_{opt} is the optimal capacity of the VLPZ, which is fixed by the system designer. We denote N_{max} the maximum number of required VLPZs at a given cell.

B. An Optimal Placement of VLPZs

In this subsection, we try to answer to the following question. Given m RIs at a given cell, with $m >= N_{max}$, what are the best RIs that should host VLPZs taking reducing the transportation cost of vehicles as an objective function?

To answer to this question, we formulate the problem as follows. Let $i = \{1, \ldots, n\}$ the set of existing vehicles at a given cell at time t. Let $j = \{1, \ldots, m\}$ be the set of the RIs candidates to host VLPZs. Let c_{ij} the transportation cost that is spent by a vehicle i to move to a RI_j. Let y_j a binary decision variable, which indicates that the infrastructure is selected to host a VLPZ at time t. x_{ij} a binary variable, that indicates that the vehicle i is assigned to the RI_j.

In order to select the best RIs to host the required VLPZs, we should minimize the following objective function F that aims to minimize the transportation cost of vehicles to move to the VLPZs.

$$F = min \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} \tag{8}$$

The transportation cost c_{ij} can be expressed as the time taken by a vehicle i to reach a candidate RI j and quantified by the loss of pseudonyms during this time, which can be calculated using the following formula:

$$cij = \frac{d_{ij}}{v} * \eta \tag{9}$$

- d_{ij}: the distance between a vehicle i and a candidate roadside infrastructure j.
- v: the average speed of vehicles (m/s).
- η: the frequency of changing of pseudonym (pseudo/s).

We assume that v and η are fixed values. Thus, the objective function F can be rewritten as a function of d_{ij} as follows:

$$F = min \sum_{i=1}^{n} \sum_{j=1}^{m} d_{ij} x_{ij} \tag{10}$$

The solution's feasibility is dependent on different constraints, which are represented by the following equations: (11) ensures that each vehicle i is only assigned to one RI; (12) ensures that the number selected RIs is equal to the number of VLPZs required at time t ($N_{vlpz}(t)$). (13) ensures that the number of vehicles that are assigned to each RI does not exceed the number of the capacity of the RI (K_{opt}); and finally, (14) and (15) are the integrity constraints.

$$\begin{cases} \sum_{j=1}^{m} x_{ij} = 1 & (11) \\ \sum_{j=1}^{m} y_j = N_{vlpz}(t) & (12) \\ \sum_{j=1}^{n} x_{ij} <= K_{opt} & (13) \\ x_{ij} \in \{0,1\} & (14) \\ y_j \in \{0,1\} & (15) \end{cases}$$

VI. Performance Evaluation

In this section, we evaluate the performance of the proposed framework. This evaluation consists of three parts. We first evaluate the privacy protection provided by the framework. We then perform an analytical evaluation of the expected number of VLPZs required at a given cell. Finally, we run a simple algorithm to select the best RIs to host VLPZs.

A. Privacy Protection

A set of simulations are realized to evaluate the privacy protection degree provided by the proposed framework. Veins is the simulation framework that is used to perform these simulations. The main foundations of Veins are OMNet++ and SUMO [20]. These simulation tools are bi-directionally coupled and communicate through a TCP socket during the simulation runtime. The strong feature of Veins is that it is fully based full 802.11p and IEEE 1609.4 DSRC/WAVE network layers. The parameters considered in our simulation are summarized in Table II

The considered scenario, models the traffic of the city of Manhattan New York, USA. We focused on a region of interest (ROI) of dimensions 2km x 2km. The vehicles were generated using SUMO to take trips of 1 hour duration over the city. We have installed a set of VLPZs (from 3 to 5) on random positions of the map.

The privacy level values and the reputation values of vehicles are initialized according to a normal distribution $\mathcal{N}$ (μ,σ) with a mean equal to $\mu = 1.5$. In addition, as shown in Table II, the required levels of all vehicles and their sensitivity parameters are assumed equals. In our evaluation,

TABLE II
Simulation Parameters

Parameter	Value
Simulation duration	1 hour
Transmission Range	500 m
Traffic density (ρ)	50, 75 veh/km^2
The capacity of VLPZ (K)	10
The number of VLPZs (N_{vlpz})	3, 4, 5
The required level (A_d)	K/3
The reputation threshold (ω)	K/4
The sensitivity parameter (λ)	0.1, 0.2, 0.3
Changing peudonym frequency	1 min

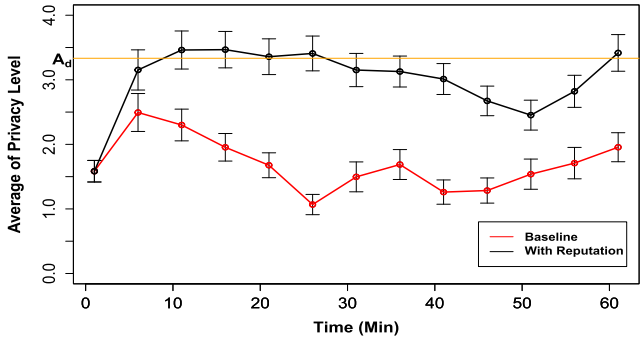


Fig. 4. Comparison between two variants of the framework in terms of the average privacy protection levels obtained by vehicles over time. ($\alpha = 0.1$, $\rho = 75\ veh/km^2$, $N_{vlpz} = 4$).

we run simulation several times with different random seeds and calculate the average value of 95% confidence interval.

Figure 4 illustrates a comparison between two variants of the framework in terms of the evolution of vehicles' privacy protection levels over time: (i) the baseline version that does not use the reputation-based mechanism to motivate vehicles to enter VLPZs, and (ii) the full version of the framework. As we can see in the Figure, the average of privacy protection levels obtained using the full version exceed the ones obtained using the baseline. Indeed, thanks to the motivation mechanism vehicles enter to VLPZs even though they reached their required level to cooperate with others vehicles, which resets the average privacy levels to A_d each time it drops down.

In Figure 5, we study the impact of the number of deployed VLPZs on the average of location privacy protection levels obtained by vehicles over time. As we can see in the Figure, the average level of privacy increases with the number of deployed VLPZs. The reason for this is the fact that with a large number of deployed VLPZs, vehicles have more chance to enter a VLPZ. We also study the impact of traffic density on the average of privacy protection levels obtained by vehicles over time. For this reason, we considered two traffic density values 50 veh/km^2 and 75 veh/km^2. As we can see in Figure 6, the average levels of privacy sightly increases with traffic density. This can be explained that with the existence of vehicles on roads, it is more likely that vehicles pass through VLPZs. Indeed, Figure 7 generally shows the occupancy of VLPZs increases with traffic density.

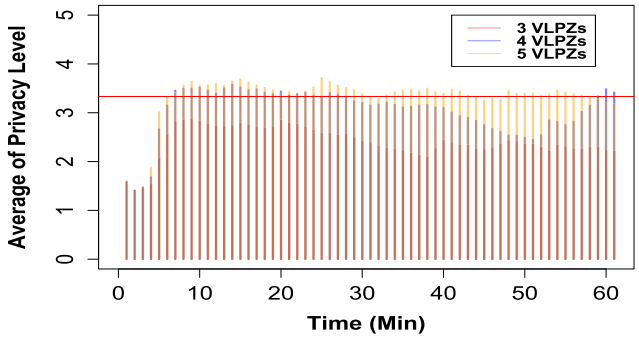


Fig. 5. The average of privacy protection level obtained by vehicles over time as function of the number deployed VLPZs. ($\alpha = 0.1$, $\rho = 75\ veh/km^2$).

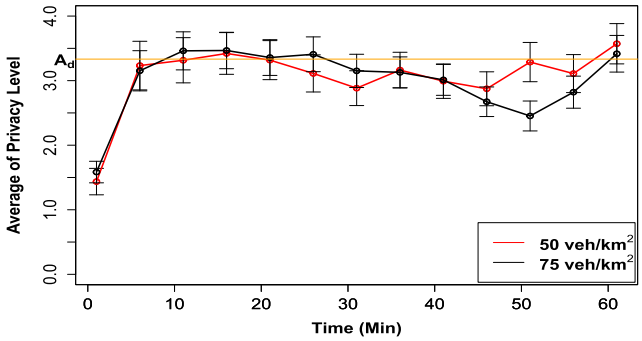


Fig. 6. The average of privacy protection level obtained by vehicles over time as a function of traffic density. ($\alpha = 0.1$, $N_{vlpz} = 4$).

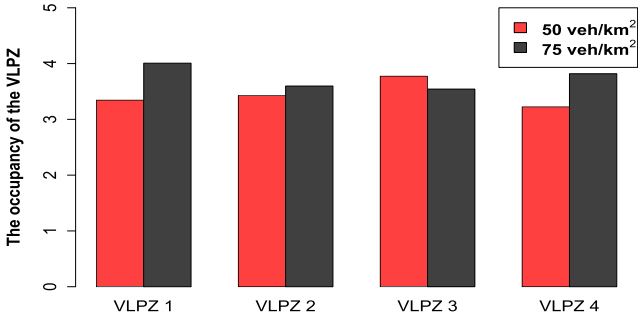


Fig. 7. The occupancy of VLPZs as a function of traffic density. ($\alpha = 0.1$, $N_{vlpz} = 4$).

In Figure 8, we study the impact of the adversary's power on the average of location privacy protection level obtained by vehicles over time. As we can see, the average of level of privacy decreases with the increase of the adversary's power. These results are expected due to the linear nature of the considered privacy loss model. Indeed, the adopted model is abstract and does not rely on important parameters such as the number of sent messages and the mobility of vehicles.

B. Expected Number of Required VLPZs

In this section, we illustrate the utility of the proposed model to calculate the number of required VLPZs at a given cell through a simple analytical evaluation. In this evaluation, we analyze the integral given in the formula (6) and calculate the required VLPZs using formula (7). In the following, we assume that vehicles arrive at the cell with a constant rate 1000 veh/h. We also assume that the residency times

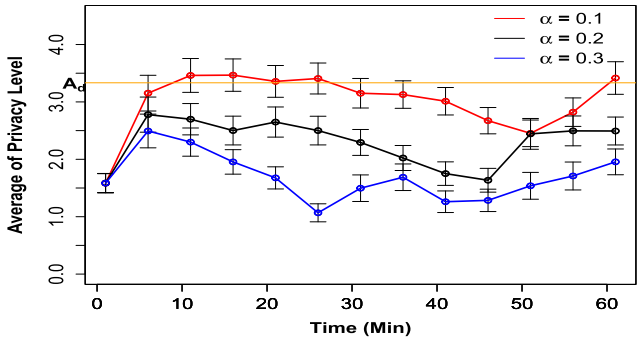


Fig. 8. The average of privacy protection level obtained by vehicles over time as a function of the attacker power. ($\rho = 75\ veh/km^2$, $N_{vlpz} = 4$).

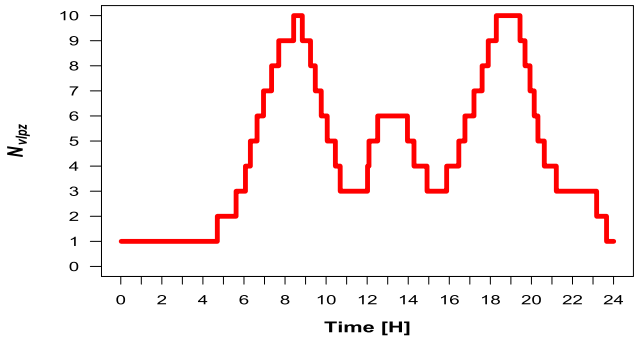


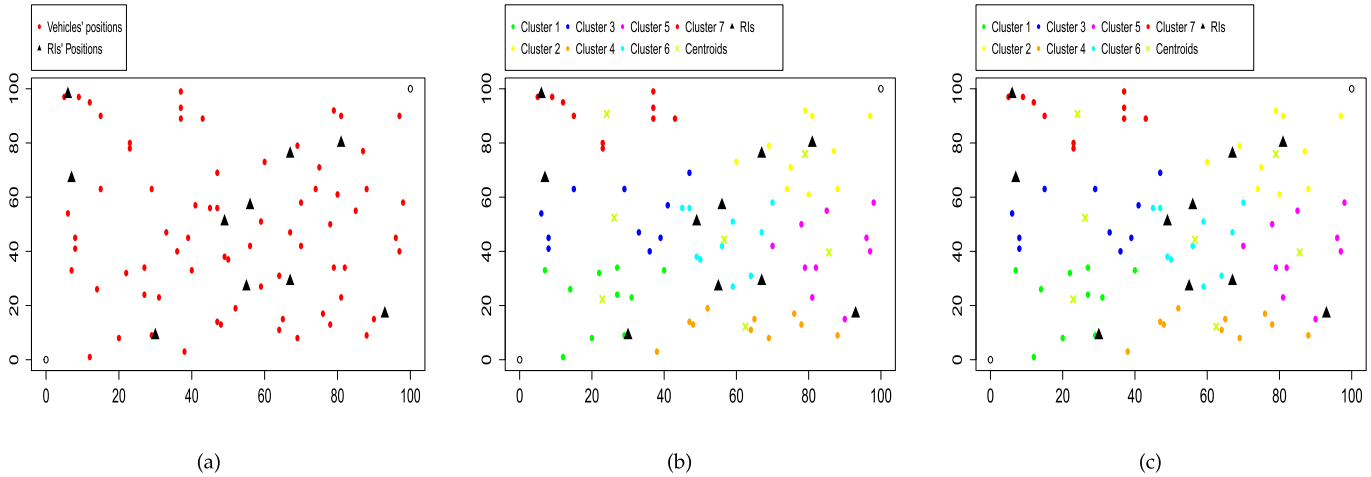
Fig. 9. The expected number of VLPZs required at a given cell as a function of time.

of vehicles inside the cell are exponentially distributed with the distribution function $F_G = 1 - e^{-\mu t}$ with a mean $1/\lambda$ is equal to 1 hour. M is a time-depended random process that marks vehicles according to the traffic density in the cell. For this reason, to model M, we use real measurements, recently published, of 24 hours of traffic mobility in the City of Luxembourg [31]. We assume that, in peak hours, only 10% of vehicles intend to enter the VLPZ and the optimal capacity of a VLPZ (K_{opt}) is equal to 10.

Figure 9 illustrates the expected number of VLPZs ($N_{vlpz}(t)$) required as a function of 24 day hours. From the Figure, we can see that $N_{vlpz}(t)$ starts to increase gradually from about 5 am until it reaches its maximum value ($N_{max} = 10$) VLPZs at 8 am. After that, we notice that $N_{vlpz}(t)$ decreases until about 11 am. However, between 12 pm and 3 pm, $N_{vlpz}(t)$ slightly increases and decreases again. Form 4 pm, $N_{vlpz}(t)$ restarts to increase gradually towards the maximum value, which is reached at about 7 pm. Finally, from 8 pm, $N_{vlpz}(t)$ gradually decreases until the end of day.

C. RIs Selection

The selection of the best RIs to host VLPZs is Np-hard problem. In this section, we propose a simple algorithm to select the best roadside infrastructures to host VLPZs. In order to solve this problem, we first propose to group vehicles into $N_{vlpz}(t)$ clusters that have the same size K_{opt} using a variation of k-means clustering algorithm. The algorithm proposed by ELKI Framework in [32] can be used for this purpose. We then create a list that contains the distances between each centroid of a cluster and each candidate roadside infrastructure j.



(a) (b) (c)

Fig. 10. RIs selection: example.

This list is sorted from the lowest to the highest distance. After that the selection of the list of N_{vlpz}(t) roadside infrastructure (Ψ) begins. In each round, we (i) pick the first element of the list; (ii) include the roadside infrastructure j into Ψ; and finally (iii) update the list by removing the distances from the centroid to the other roadside infrastructures and the distances from j to the other centroids. The algorithm runs until the selection of the N_{vlpz}(t) roadside infrastructures is done.

An example allows illustrating the RIs selection process is given as follows. As shown in Figure 10a, we assume a cell that contains 10 RIs candidates and 70 vehicles. The positions of both vehicles and RIs are randomly generated using Matlab and the capacity of each RI (K_{opt}) is equal to 10. The number of RIs that should then be selected is equal to 7. Figure 10b illustrates the 7 clusters that have been created using the k-means modified version proposed by ELKI Framework and their centroids. Figure 10c illustrates the selected RIs after running the selection algorithm.

VII. Discussion

In this subsection, we provide a short discussion on some advantages and limitations of PRIVANET. Indeed, PRIVANET mainly brings new solutions which are not yet implemented in the current security standards (IEEE 1609.2 [5] and ETSI 102941-v1.1.1 [6]). Specifically, it proposes an effective pseudonym changing strategy and a new IP/MAC generation mechanism from pseudonyms. PRIVANET also proposes improvements for the solutions that are already proposed in security standards. For example, it includes a pseudonym sets distribution method that has low overhead compared to the one proposed by the security standards. In addition, the hierarchical structure of PRIVANET allows the ease of management and its modularity and extensibility make it well adapted for the future deployment of VANETs. However, PRIVANET has some limitations, especially against an active adversary model. This type of adversary can make the proposed privacy protection approach ineffective. For example, an active adversary could know all the distributed pseudonym sets if it controls RSU_{VLPZS}. As a result, the adversary still tracks vehicles even if they change their pseudonyms. In addition, if an active adversary controls all the VLPZs, the changing of pseudonyms will become ineffective since the adversary can easily match between the vehicles that entered the VLPZs and those that exit them.

Fortunately, the impact of the active adversary in PRIVANET depends on its coverage. Indeed, the control of all the VLPZs is more serious than the control of one or two VLPZs. To overcome this issue, PRIVANET should then be built on a strong security architecture that allows to detect active attackers and revoke them from the system.

VIII. Conclusion

In this paper, we proposed PRIVANET, a complete framework that can easily be deployed in the real-world and adopted by ITS community.This framework has a hierarchical structure and is mainly based on logical zones called Vehicular Location Privacy Zones (VLPZs). It considers that the vehicular geographic area is partitioned as a grid, where each cell contains one or many VLPZs. These zones can easily be deployed over the widespread Roadside Infrastructures (RIs) such as gas stations and electric vehicles charging stations. The proposed framework includes different building blocks to address the key issues of the pseudonym changing approach such as effective pseudonym changing and management and a mechanism for overcoming the selfishness behavior of vehicles. An optimal deployment of VLPZs over RIs is also proposed and the framework is evaluated and validated through extensive simulations. The obtained results are promising. They show the ability of the proposed framework to provide an efficient and secure privacy protection for VANETs'users.

References

[1] G. Karagiannis *et al.*, "Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 584–616, Nov. 2011.

[2] Z. Doukha and S. Moussaoui, "An SDMA-based mechanism for accurate and efficient neighborhood-discovery link-layer service," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 603–613, Feb. 2016.

[3] D. Eckhoff, R. German, C. Sommer, F. Dressler, and T. Gansen, "SlotSwap: Strong and affordable location privacy in intelligent transportation systems," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 126–133, Nov. 2011.

[4] F. Dressler, F. Kargl, J. Ott, O. K. Tonguz, and L. Wischhof, "Research challenges in intervehicular communication: Lessons of the 2010 Dagstuhl seminar," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 158–164, May 2011.

[5] *IEEE Standard for Wireless Access in Vehicular Environments–Security Services for Applications and Management Messages*, IEEE Standard 1609.2-2013 (Revision IEEE Std 1609.2-2006), Apr. 2013, pp. 1–289.

[6] *Intelligent Transport Systems (ITS); Security; Trust and Privacy Management*, document TS 102 941, v1.1.1, ETSI, 2012.

[7] E. Schoch, F. Kargl, T. Leinmüller, S. Schlott, and P. Papadimitratos, "Impact of pseudonym changes on geographic routing in VANETs," in *Proc. 3rd Eur. Conf. Secur. Privacy Ad-Hoc Sensor Netw.*, Berlin, Germany: Springer-Verlag, 2006, pp. 43–57.

[8] P. Papadimitratos, L. Buttyan, J.-P. Hubaux, F. Kargl, A. Kung, and M. Raya, "Architecture for secure and private vehicular communications," in *Proc. 7th Int. Conf. ITS Telecommun.*, Jun. 2007, pp. 1–6.

[9] J. Petit, F. Schaub, M. Feiri, and F. Kargl, "Pseudonym schemes in vehicular networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 228–255, 4th Quart., 2014.

[10] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in inter-vehicular networks: Why simple pseudonym change is not enough," in *Proc. 7th Int. Conf. Wireless On-Demand Netw. Syst. Services*, Feb. 2010, pp. 176–183.

[11] R. Lu, X. Lin, T. Luan, X. Liang, and X. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in VANETs," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 86–96, Jan. 2012.

[12] B. Ying and D. Makrakis, "Pseudonym changes scheme based on candidate-location-list in vehicular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 7292–7297.

[13] B. Ying, D. Makrakis, and Z. Hou, "Motivation for protecting selfish vehicles' location privacy in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5631–5641, Dec. 2015.

[14] A. Boualouache and S. Moussaoui, "S2SI: A practical pseudonym changing strategy for location privacy in VANETs," in *Proc. Int. Conf. Adv. Netw. Distrib. Syst. Appl. (INDS)*, Jun. 2014, pp. 70–75.

[15] D. Eckhoff and C. Sommer, "Driving for big data? privacy concerns in vehicular networking," *IEEE Security Privacy*, vol. 12, no. 1, pp. 77–79, Jan. 2014.

[16] J. Freudiger, M. H. Manshaei, J. Hubaux, and D. C. Parkes, "Non-cooperative location privacy," *IEEE Trans. Dependable Secure Comput.*, vol. 10, no. 2, pp. 84–98, Mar./Apr. 2013.

[17] Z. Ma, F. Kargl, and M. Weber, "Pseudonym-on-demand: A new pseudonym refill strategy for vehicular communications," in *Proc. IEEE 68th Veh. Technol. Conf.*, Sep. 2008, pp. 1–5.

[18] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "Towards an efficient pseudonym management and changing scheme for vehicular ad-hoc networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[19] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis," *IEEE Trans. Mobile Comput.*, vol. 10, no. 1, pp. 3–15, Jan. 2011.

[20] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO—Simulation of Urban MObility," *Int. J. Adv. Syst. Meas.*, vol. 5, nos. 3–4, pp. 128–138, Dec. 2012.

[21] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "A survey on pseudonym changing strategies for vehicular ad-hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 770–790, 1st Quart., 2017.

[22] S. Lefèvre, J. Petit, R. Bajcsy, C. Laugier, and F. Kargl, "Impact of v2x privacy strategies on intersection collision avoidance systems," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Dec. 2013, pp. 71–78.

[23] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "VLPZ: The vehicular location privacy zone," *Procedia Comput. Sci.*, vol. 83, pp. 369–376, Jan. 2016.

[24] H. Artail and N. Abbani, "A pseudonym management system to achieve anonymity in vehicular ad hoc networks," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 1, pp. 106–119, Jan. 2016.

[25] A. Boualouache, S.-M. Senouci, and S. Moussaoui, "HPDM: A hybrid pseudonym distribution method for vehicular ad-hoc networks," *Procedia Comput. Sci.*, vol. 83, pp. 377–384, Jan. 2016.

[26] G. Yan, S. Olariu, J. Wang, and S. Arif, "Towards providing scalable and robust privacy in vehicular networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 7, pp. 1896–1906, Jul. 2014.

[27] R. Lu, X. Lin, Z. Shi, and X. S. Shen, "A lightweight conditional privacy-preservation protocol for vehicular traffic-monitoring systems," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 62–65, May/Jun. 2013.

[28] P. P. Papadimitratos, G. Mezzour, and J.-P. Hubaux, "Certificate revocation list distribution in vehicular communication systems," in *Proc. 5th ACM Int. Workshop Veh. Inter-NETworking*, Sep. 2008, pp. 86–87.

[29] T. Aura, *Cryptographically Generated Addresses (CGA), Internet Requests for Comments*, document RFC 3972, Mar. 2005. [Online]. Available: https://tools.ietf.org/html/rfc3972

[30] S. Qadir and M. U. Siddiqi, "Cryptographically generated addresses (CGAs): A survey and an analysis of performance for use in mobile environment," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 11, no. 2, pp. 24–31, Feb. 2011.

[31] L. Codeca, R. Frank, and T. Engel, "Luxembourg SUMO traffic (LuST) scenario: 24 hours of mobility for vehicular networking research," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Kyoto, Japan, Dec. 2015, pp. 1–8.

[32] *Same-Size K-Means Variation*. Accessed: Apr. 16, 2017. [Online]. Available: http://elki.dbs.ifi.lmu.de/wiki/Tutorial/SameSizeKMeans