Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1424078, 9 pages
https://doi.org/10.1155/2022/1424078

Research Article

Trajectory Privacy Preserving for Continuous LBSs in VANET

Zhihong Li,[1,2] Xiaoshuang Xing,[2] Jin Qian,[3] Hui Li,[3] and Gaofei Sun[2]

[1]*School of Computer Science and Technology, Soochow University, Suzhou, China*
[2]*School of Computer Science and Engineering, Changshu Institute of Technology, Suzhou, China*
[3]*School of Computer Science and Technology, Taizhou University, Taizhou, China*

Correspondence should be addressed to Xiaoshuang Xing; xing@cslg.edu.cn

Received 2 December 2021; Revised 18 January 2022; Accepted 28 January 2022; Published 15 February 2022

Academic Editor: Ruinian Li

Location-based services (LBSs) support various applications in vehicular ad hoc network (VANET). However, location/trajectory privacy becomes a serious concern for LBSs. Existing location/trajectory privacy-preserving schemes rarely take the attack model of adversaries into consideration, and the cost for achieving privacy has not been carefully studied. To deal with these problems, this study proposes a collaborative trajectory obfuscation scheme based on analyzing the attack model and designs a privacy-preserving efficiency metric that balances the achieved privacy and the cost. Through simulation, the effects of the density of vehicles using the same LBSs on the performance of our design and an existing scheme are investigated. The performance comparison results validate the effectiveness and efficiency of our scheme.

1. Introduction

Vehicular ad hoc network (VANET) has become an important framework of the intelligent transportation system (ITS) for applications such as navigation, road safety, and entertainment [1]. Location-based service (LBS) usually acts as the foundation that supports these applications. For example, a vehicle that wants to find a nearby supermarket and decide a suitable driving route should submit its identity, location, and service requests to corresponding LBS providers (LBSPs). While providing LBSs, LBSPs collect vehicles' locations. Once the LBSP is untrustworthy or attacked by malicious adversaries, vehicles' identities and locations will be disclosed. By analyzing frequently visited locations, private information of vehicle drivers, such as personal preferences, work locations, home addresses, and health conditions, can be revealed [2, 3]. To deal with these threats, location privacy preserving for LBSs has long been considered as an important research topic in VANETs.

The basic idea for location privacy preserving is to use pseudonyms instead of real identities when submitting LBS requests to eliminate the link between vehicles' identities and locations [4]. However, simple pseudonym replacement can only provide a single-point location privacy preserving [5]. When continuous LBSs are used, frequently visited location and/or trajectories of vehicles can still be revealed if locations are submitted to LBSPs periodically with the same pseudonym. To deal with this problem, pseudonym-changing schemes such as silent period [6, 7] and mix zone [8] are proposed, but their performance is unsatisfactory when facing the correlation attacks and their real-time performance is expected to be improved.

Location/trajectory obfuscation schemes have been designed to solve the aforementioned problems. In location obfuscation schemes, vehicles change their actual location coordinates within a tolerable error range and submit the changed locations to LBSPs. In this way, location privacy is preserved with the cost of decreased service quality since the LBSs are provided based on changed locations [9]. In trajectory obfuscation schemes, fake LBS requests, whose locations are obtained from collaborators, are submitted together with vehicles' actual LBS requests. In this way, adversaries will be misled and failed to trace the trajectories when proper collaborators are selected [10]. Trajectory obfuscation schemes preserve privacy without decreasing the service quality. However, most work has not taken the attack model of adversaries into consideration for designing trajectory obfuscation schemes. Intuitively, trajectory

privacy can be better preserved if we understand how the attackers trace the trajectories. Besides, the scenario when not all vehicles are using the same LBS has not been carefully studied. When vehicle v selects a vehicle not using the same LBS as the collaborator, the collaborator will not be able to cause significant bifurcation of the trajectory. Consequently, an attacker can still trace the trajectory successfully with high probability. Regarding the performance evaluation, various metrics, such as tracking successful ratio [11], location/trajectory entropy [12], and anonymity set size [13], have been designed to measure the achieved location/trajectory privacy. However, the cost for achieving such privacy-preserving performance has rarely been studied.

In this study, we tackle these challenges by making the following contributions:

(1) We propose a novel trajectory privacy-preserving scheme based on understanding the attack model of adversaries. We first analyze how the adversaries predict/trace the vehicles' trajectories using the Kalman filter. Then, vehicles using LBSs predict their future locations with the Kalman filter and select collaborators based on predicted locations. In this way, collaborators that are most capable of misleading the adversaries can be selected and the trajectory privacy-preserving performance can be ensured.

(2) A privacy-preserving efficiency metric is designed to evaluate the trajectory privacy-preserving performance and the cost for achieving such privacy performance.

(3) Unlike the simulation settings of the existing collaborative solutions, we set the density of vehicles using the same LBSs as a variable to better reproduce the real usage scenario. The effects of the density on our design and an existing scheme are investigated.

The rest of the study is organized as follows. Related works are reviewed in Section 2. Section 3 describes the considered system model. The attack model of adversaries is analyzed in Section 4, based on which a collaborative trajectory obfuscation scheme is proposed in Section 5. The performance of our design is compared with some exiting schemes in Section 6, and this study is concluded in Section 7.

2. Related Work

LBS supports a broad range of applications in VANETs. Locations are submitted to the LBS provider together with vehicle identities, which threaten users' location privacy and trajectory privacy. To deal with these threats, pseudonym-changing-based schemes and location/trajectory obfuscation-based schemes have been extensively studied [14].

In pseudonym-changing-based schemes, vehicles use pseudonyms instead of real identities when submitting LBS requests. In this way, links between locations and identities are broken, and location privacy can be preserved. Moreover, a vehicle changes pseudonyms following designed algorithms. Thus, links among locations of a vehicle at different times are broken and trajectory privacy can be preserved.

The silent period is an early proposed pseudonym-changing algorithm [6, 7]. A time period is defined as silent period, within which vehicles do not submit any LBS requests (i.e., keep silent) and after which vehicles submit LBS requests with changed pseudonyms. This method can mislead the attacker when more than one vehicles change their pseudonyms at the end of a silent period. However, due to the silence period, applications with high real-time requirements cannot be satisfied.

Mix zone is considered as a promising pseudonym-changing algorithm [8, 15] where mutually cooperative vehicles concurrently change their pseudonyms in mix zones created by themselves. The effectiveness of the mix zone depends on factors such as geometry, vehicle density, and geographic location in the road network. In addition, most mix zone schemes cannot avoid the continuous query correlation attack, thus limiting their performance in continuous LBS applications.

Due to the limitations mentioned above, pseudonym changing is usually used together with location/trajectory obfuscation. In location obfuscation, a vehicle changes the actual location coordinates within a tolerable error range. Then, the changed location is submitted to the LBSP. A method called CoPrivacy is proposed in [16], where vehicles form k anonymity groups through collaboration and a vehicle replaces its actual location by the regional density center of the anonymous group it belongs to. Reference [17] explores the minimum amounts of obfuscation and anonymization to block attacks on user's location privacy using an information-theoretic approach with the Markov chains. However, location obfuscation preserves location privacy with the cost of decreased service quality since LBSP provides service based on the changed locations [9].

In 2016, [18] proposed a trajectory obfuscation scheme called mutual obfuscating path (MOP) to preserve privacy without decreasing the service quality. For each vehicle v, it selects a collaborator from vehicles that are currently within its communication range. Vehicles whose trajectories are predicted to converge on v's trajectory within a predefined time threshold form the candidate collaborator set and within which the vehicle being nearest to v is selected as the collaborator. Then, v will send two LBS requests with two different pseudonyms to the LBSP. The locations of these two requests will be v's actual location and the collaborator's predicted location, respectively. From the attacker's perspective, the trajectories continue to bifurcate over time, which impedes the attacker from successful trajectory tracing. Reference [10] pointed out that there may be nefarious vehicles in the internal cooperation of MOP. To deal with this problem, it proposed a non-collaborative approach. A vehicle independently decides whether to exploit the fake location of the surrounding vehicles based on the proposed algorithm.

Despite the real time and guaranteed service quality of these two schemes, there are still open challenges to be

solved. First, the attack model has not been investigated in this work. Intuitively, trajectory privacy can be better preserved if we understand how the attackers trace the trajectories. Therefore, in this work, we will design a trajectory obfuscation scheme based on analyzing the attack model. Second, this work does not consider the scenario when not all vehicles are using the same LBS. When vehicle v selects a vehicle not using the same LBS as the collaborator, the collaborator will not be able to cause significant bifurcation of the trajectory. Consequently, an attacker can still trace the trajectory successfully with high probability. Therefore, we will take the scenario when not all vehicles are using the same LBS into consideration and the effects of the density of vehicles using the same LBS on the trajectory obfuscation performance will be investigated.

3. System Model

For easier following, we summarize the notations introduced throughout the next three sections in Table 1.

In the considered system model, there are n vehicles, denoted by $\mathbf{V} = \{v_1, v_2, \ldots, v_n\}$ on the road, and m LBS providers (LBSPs), denoted by $\mathbf{P} = \{p_1, p_2, \ldots, p_m\}$, providing m different LBSs. Each vehicle can access to $p_i \in \mathbf{P}$ for LBS or do not use LBS depending on its demand. Each vehicle is equipped with an onboard unit (OBU). Through the OBU, a vehicle can obtain real-time information about the surrounding environment and communicate with other vehicles or infrastructures in VANET. When a vehicle wants to use the ith $(1 \le i \le m)$ LBS, it sends a request to p_i via Internet service provider (ISP). ISP is a communication agency between OBU and LBSP via which the LBS requests and responses are sent. The considered LBS model is shown in Figure 1.

In our work, we mainly focus on the continuous LBS scenario, where a vehicle periodically sends the LBS request to LBSP during the service time. A topical continuous LBS is navigation. Assume a vehicle v_k uses the ith LBS from time t_0 to time t_{end}. The LBS request sent by v_k at time $t_0 \le t_j \le t_{end}$ can be denoted by $\{I_k, Loc_k^j, R_k^i, t_j\}$. Here, I_k denotes the identity of v_k, $Loc_k^j = (Lo_k^j, La_k^j)^T$ indicates v_k's location at t_j with Lo_k^j and La_k^j being the longitude and the latitude, R_k^i represents v_k's service request to LBSP p_i, and t_j is the timestamp. Let $T_{int} = t_{j+1} - t_j$ denote the time interval between two continuous LBS requests. Without loss of generality, we let T_{int} be an unit time in this study. To resist the identity-link attacks, vehicles usually use time-changing pseudonyms instead of real identity for LBS request. Therefore, the LBS request can be revised as $\{\tilde{I}_k^j, Loc_k^j, R_k^i, t_j\}$ with $\tilde{I}_k^j$ being the pseudonym of v_k at t_j.

Let $\mathbf{v}_i \in \mathbf{V}$ denote the subset of vehicles using the ith LBS from t_0 to t_{end}. By the end of t_{end}, p_i collects information regarding the pseudonyms and locations of these vehicles at different time points as $CI = \{\{\tilde{I}_k^0, Loc_k^0, t_0\}, \ldots, \{\tilde{I}_k^{end}, Loc_k^{end}, t_{end}\}\}$, $v_k \in \mathbf{v}_i$. If p_i is malicious or it is attacked by malicious adversaries, the possible trajectories of vehicles in $\mathbf{v}_i$ can be predicted using the Kalman filter or other methods. Thus, the private trajectory information will be disclosed. To tackle this problem and preserve the trajectory privacy, a collaborative obfuscation method will be designed based on the understanding of the adversary's attack model.

4. Adversary Model When the Kalman Filter Is Used for Trajectory Prediction

To effectively protect the trajectory information from being disclosed, we should understand how the adversaries predict the trajectories. Therefore, we will analyze the attack model of adversaries when the Kalman filter is used for trajectory prediction. Based on CI collected at the LBSP p_i, the adversary forms a state vector $\mathbf{x}^j = (Lo^j, La^j, Vo^j, Va^j)^T$ to denote the state of a LBS request with timestamp t_j. Here, Lo^j and La^j denote the longitude and the latitude obtained from the LBS request, and Vo^j and Va^j denote the velocity in the longitude direction and the latitude direction. Vo^0 and Va^0 are set to be 0. Vo^j and Va^j $(j \neq 0)$ can be calculated as $Vo^j = (1/T_{int})Do^j$ and $Va^j = (1/T_{int})Da^j$ with Do^j indicating the distance between locations (Lo^j, La^j) and (Lo^{j-1}, La^j) in the longitude direction and Da^j indicating the distance between locations (Lo^j, La^j) and (Lo^{j-1}, La^j) in the latitude direction.

For t_0, the adversary will form $|\mathbf{v}_i|$ state vectors $\mathbf{x}_1^j$, $\mathbf{x}_2^j, \ldots, \mathbf{x}_{|\mathbf{v}_i|}^j$ to denote the initial states of the $|\mathbf{v}_i|$ vehicles that are using p_i's service. For t_j $(j \neq 0)$, the adversary may not be able to identify the state of $v_k \in \mathbf{v}_i$ due to pseudonym changing. Let us take an example for better understanding. Let $\{\{110011, (41.40, 2.17), t_0\}, \{110000, (40.40, 2.31), t_0\}, \{110011, (42.40, 2.17), t_1\}, \{110000, (40.87, 2.40), t_1\}, \{100110, (42.98, 2.39), t_2\}, \{001100, (41.05, 2.37), t_2\}\}$ denote the CI collected by p_i from t_0 to t_2. Two state vectors $\mathbf{x}_1^0 = (41.40, 2.17, 0, 0)^T$ and $\mathbf{x}_2^0 = (40.40, 2.31, 0, 0)^T$ will be formed at t_0. Since the pseudonyms are not changed from t_0 to t_1, the adversary easily identifies $\mathbf{x}_1^1$ and $\mathbf{x}_2^2$ from CI as $\mathbf{x}_1^1 = (42.40, 2.17, Vo_1^1, Va_1^1)^T$ and $\mathbf{x}_2^1 = (40.87, 2.40, Vo_2^1, Va_2^1)^T$. At t_3, the pseudonyms of vehicles are changed. The adversary cannot identify directly from CI whether location $(42.98, 2.39)$ or location $(41.05, 2.37)$ belongs to $\mathbf{x}_1^3$. In this case, the Kalman filter will be applied by the adversary to make a prediction and decide which location belongs to $\mathbf{x}_1^3$.

Vehicle $v_k \in \mathbf{v}_i$ is taken as an example, and we will describe how the adversary predicts v_k's trajectory. Let $\bar{\mathbf{x}}_k^j$ denote the state of v_k at t_j $(j \neq 0)$ estimated using motion model, and it can be given as follows:

$$\bar{\mathbf{x}}_k^j = \mathbf{A}\mathbf{x}_k^{j-1} + \mathbf{B}\mathbf{u}_k^{j-1} + \mathbf{w}, \tag{1}$$

where $\mathbf{x}_k^{j-1}$ is considered by the adversary to be v_k's state at t_{j-1} with probability $p(k, j-1)$ and $p(k, 0)$ is set to be 1. The calculation of $p(k, j-1)$ will be given later in this section. $\mathbf{A}$ is called as the state transition matrix, and $\mathbf{B}$ is called as the input matrix. They are defined as follows:

Table 1: Summary of notations.

Symbol	Meaning
n, m	Number of vehicles on the road, number of LBSPs
$v_k, \mathbf{V}$	The kth vehicle $(1 \le k \le n)$, set of vehicles on the road
$p_i, \mathbf{P}$	The ith LBSP $(1 \le i \le m)$, set of LBSPs
t_j	The jth timestamp of v_k using p_i $(0 \le j \le end)$
$I_k, \widetilde{I}_k^j$	Identity of v_k, pseudonym of v_k at t_j
Lo_k^j, La_k^j, Loc_k^j	v_k's longitude at t_j, v_k's latitude at t_j, v_k's location at t_j
R_k^i	v_k's service request to p_i
T_{int}	The time interval between two continuous LBS requests
$\mathbf{v}_i$	Subset of vehicles using the ith LBS $(1 \le i \le m)$
CI	Set of v_i's information collected by p_i
$\mathbf{x}^j$	State vector formed by the adversary at t_j
$\mathbf{x}_{\lvert\mathbf{v}_i\rvert}^j$	Initial state of the $\lvert\mathbf{v}_i\rvert$th vehicle that is using p_i's service at t_j
Vo^j, Va^j	Velocity in the longitude direction and the latitude direction at t_j
Do^j	Distance between locations (Lo^j, La^j) and (Lo^{j-1}, La^j) in the longitude direction
Da^j	Distance between locations (Lo^j, La^j) and (Lo^{j-1}, La^j) in the latitude direction
$\overline{\mathbf{x}}_k^j, \mathbf{x}_k^j$	Estimated state of v_k at t_j $(j \ne 0)$, v_k's state considered by the adversary at t_j
$\mathbf{u}_k^j$	v_k's acceleration vector at t_j
$\mathbf{A}, \mathbf{B}, \mathbf{H}$	State transition matrix, input matrix, measurement matrix
Ao_k^j, Aa_k^j	v_k's acceleration in the longitude direction and the latitude direction at t_j
$\mathbf{w}, \mathbf{v}$	Disturbance noise, measurement noise
$\mathbf{Q}, \mathbf{R}$	Covariance of $\mathbf{w}$, covariance of $\mathbf{v}$
$\mathbf{E}_n$	Unit matrix with n rows and n columns
$\mathbf{P}_j, \overline{\mathbf{P}}_j$	Error covariance at t_j, error covariance calculated according to $\mathbf{P}_{j-1}$ at t_j
$\overline{Lo}_k^j, \overline{La}_k^j$	Estimated longitude and latitude of v_k at t_j
$\overline{Vo}_k^j, \overline{Va}_k^j$	Estimated velocity in the longitude direction and the latitude direction of v_k at t_j
G, C	Distance threshold, set of vehicles meeting the condition G
$p(k', j)$	Probability that $x_k^{'j}$ is considered by the adversary to be $v_{k'}$'s state at t_j
$\mathbf{Z}_{k'}^j$	Measurement value consisting of $Lo_{k'}^j, La_{k'}^j$
$\mathbf{K}$	Kalman gain
s_k^j, S_k^j	v_k's recorded information regarding coordinate, velocity, and acceleration at t_j, set of s_k^j
V_k^j	Set of v_k's neighbors at time t_j
y	Serial number of the minimum distance between $\overline{Loc}_k^{j+1}$ and
v_{ky}	Collaborator selected by v_k

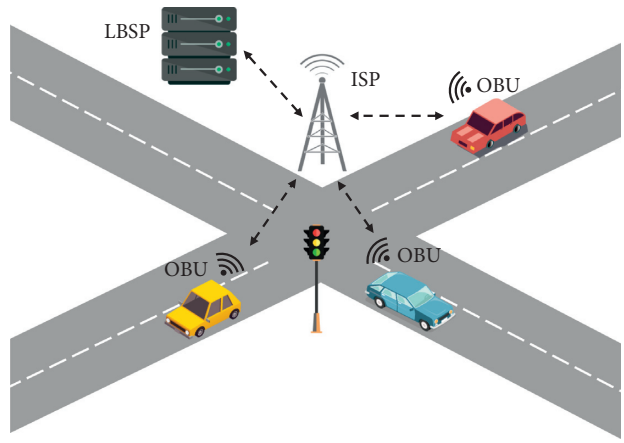

Figure 1: LBS system model in VANET.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \frac{1}{2} & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 1 \end{pmatrix}^{\mathrm{T}}. \tag{2}$$

$\mathbf{u}_k^{j-1}$ is called as the acceleration vector and is defined as $\mathbf{u}_k^{j-1} = (Ao_k^{j-1}, Aa_k^{j-1})^T$ with Ao_k^{j-1} and Aa_k^{j-1} denoting v_k's acceleration in the longitude direction and the latitude direction at time t_{j-1}. Note that Ao_k^0 and Aa_k^0 are set to be 0. Therefore, $\mathbf{A}\mathbf{x}_k^{j-1} + \mathbf{B}\mathbf{u}_k^{j-1}$ can be used to estimate the location and velocity of v_k at time t_j based on the location, velocity, and acceleration at t_{j-1}. Considering the error caused by imprecise estimation, a disturbance noise $\mathbf{w}$ is added in (1). Assume that $\mathbf{w}$ is a Gaussian white noise, and the covariance $\mathbf{Q}$ can be given as follows:

$$\mathbf{Q} = q\mathbf{E}_4. \tag{3}$$

Here, $\mathbf{E}_n$ denotes an unit matrix with n rows and n columns. The error covariance $\overline{\mathbf{P}}$ caused by $\mathbf{w}$ can be calculated as follows:

$$\overline{\mathbf{P}}_j = \mathbf{A}\mathbf{P}_{j-1}\mathbf{A}^T + \mathbf{Q}, \tag{4}$$

where $\mathbf{P}_{j-1}$ is the error covariance at time t_{j-1} and $\mathbf{P}_0$ is set to be $\mathbf{E}_4$.

Getting $\overline{\mathbf{x}}_k^j = (\overline{Lo}_k^j, \overline{La}_k^j, \overline{Vo}_k^j, \overline{Va}_k^j)^T$, the adversary will check CI and find a set of vehicles:

$$C = \{v_{k'} | Di\ s(k') < G\}. \tag{5}$$

Here, $Di\ s(k')$ indicates the distance between locations $(\overline{Lo}_k^j, \overline{La}_k^j)$ and $(Lo_{k'}^j, La_{k'}^j)$, G is a distance threshold set based on the longest distance traveled by the vehicle in one T_{int} under the speed limit on the road. Here, we set G to a fixed value 15. Then, vehicle $v_{k'}$ will be considered to be vehicle v_k by the adversary at time t_j with probability $p(k', j)$, and $p(k', j)$ can be calculated as follows:

$$p(k', j) = \frac{\sum_{v_{k_{11}} \in \in \{C/v_{k'}\}} Dis(k'')}{\sum_{v_{k_{11}} \in \in C} Dis(k11) \times \frac{p(k, j-1)}{|C| - 1}}. \tag{6}$$

Here, $p(k, j-1)$ is the probability that x_k^{j-1} is considered by the adversary to be v_k's state at t_{j-1}. When vehicle $v_{k'}$ is considered by the adversary to be vehicle v_k at time t_j, $\mathbf{Z}_{k'}^j = (Lo_{k'}^j, La_{k'}^j)^T$ will be taken as v_k's location at time t_j. Considering the errors caused by imprecise positioning, the relationship between $\mathbf{Z}_{k'}^j$ and the actual location $\mathbf{H}\mathbf{x}_{k'}^j$ can be described as follows:

$$\mathbf{Z}_{k'}^j = \mathbf{H}\mathbf{x}_{k'}^j + \mathbf{v}. \tag{7}$$

Here, $\mathbf{H}$ is a measurement matrix:

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \tag{8}$$

where $\mathbf{v}$ is the measurement noise caused by imprecise positioning. Assume that $\mathbf{v}$ in (7) is a Gaussian white noise, and the covariances $\mathbf{R}$ can be given as follows:

$$\mathbf{R} = r\mathbf{E}_2, \tag{9}$$

The parameters of the Kalman filter are updated and v_k's state at time t_j is predicted using the following formula:

$$\begin{aligned} \mathbf{K} &= \overline{\mathbf{P}}_j\mathbf{H}^T\left(\mathbf{H}\overline{\mathbf{P}}_j\mathbf{H}^T + \mathbf{R}\right)^{-1}, \\ \mathbf{x}_k^j &= \overline{\mathbf{x}}_k^j + \mathbf{K}\left(\mathbf{Z}_{k'}^j - \mathbf{H}\overline{\mathbf{x}}_k^j\right), \\ \mathbf{P}_j &= (\mathbf{E}_4 - \mathbf{K}\mathbf{H})\overline{\mathbf{P}}_j, \end{aligned} \tag{10}$$

where $\mathbf{K}$ is called as the Kalman gain. It can be found from (10) that $\mathbf{K}$ decreases with $\mathbf{R}$. That is, smaller $\mathbf{R}$ will lead to more trustable estimation using the Kalman filter. Given today's increasingly precise positioning systems, we assume that there is almost no deviation between $\mathbf{Z}_{k'}^j$ and $\mathbf{x}_{k'}^j$. Thus, r is set to be a small value $r = 0.05$. As for the setting of q, we refer to the conclusion in [5] and $q = 0.05$ is set. It should be noted that $\mathbf{x}_k^j$ obtained using (10) denotes v_k's state at time t_j predicted by the adversary when $v_{k'}$ is considered to be v_k. $\mathbf{x}_k^j$ will be substituted into (1) to estimate $\overline{\mathbf{x}}_k^{j+1}$ and correspondingly $p(k, j-1)$ in (6) will be replaced by $p(k', j)$.

5. Collaborative Obfuscation for Trajectory Privacy Preserving

Understanding how the adversary traces the vehicles' trajectories, we will design a collaborative trajectory obfuscation algorithm in this section.

Our main idea is to find collaborators for each vehicle v_k, and the collaborators will help v_k by sending the same LBS request with their pseudonyms and locations. When proper collaborators are selected, the adversary will mistake the collaborators for v_k and thus be misled during trajectory tracing. It should be noted that it is meaningless if a collaborator is driving on the same road with v_k and the adversary is successfully misled by the collaborator since v_k is still predicted to be on that road and the adversary can still get v_k's trajectory. Therefore, this work only selects collaborators at intersections and vehicles that are most capable of misleading the adversary will be selected as collaborators.

While driving, a vehicle v_k keeps recording its locations, velocities, and accelerations for the current time and N most recent historical time points. Assume the current time point is t_j, and the recorded information can be denoted as $S_k^j = \{s_k^j, s_k^{j-1}, \ldots, s_k^{j-N}\}$ with $s_k^j = (Lo_k^j, La_k^j, Vo_k^j, Va_k^j, Ao_k^j, Aa_k^j)$. Here, Lo/La, Vo/Va, and Ao/Aa denote the coordinate, velocity, and acceleration in the longitude/latitude direction, subscript k denotes vehicle v_k, and superscript j denotes time t_j. With the recorded information, v_k will predict its

location in the future time t_{j+1}, denoted by $\overline{Loc}_k^{j+1}$, using the Kalman filter. The prediction process is given in Algorithm 1.

When arriving at an intersection, vehicles predict their future locations at t_{j+1} according to Algorithm 1. Each vehicle adds its predicted location to the beacon message and broadcasts the message to neighbors, which are vehicles within its communication range. In our work, we assume that all vehicles have the same radius of communication range. Therefore, v_k is a neighbor of v_x if vehicle v_x is a neighbor of v_k. For vehicle v_k, let $V_k^j = \{v_{k1}, v_{k2}, \ldots, v_{kY}\}$ denote the set of its neighbors at time t_j. v_k will select $v_{ky} \in V_k^j$ as the collaborator if:

$$y = \arg \min_{1 \le y \le Y} Dis\left(\overline{Loc}_k^{j+1}, \overline{Loc}_{ky}^{j+1}\right), \tag{11}$$

and

$$Dis\left(\overline{Loc}_k^{j+1}, \overline{Loc}_{ky}^{j+1}\right) \le G. \tag{12}$$

Here, G is a distance threshold as described in Section 4. After selecting v_{ky} as the collaborator, v_k sends to v_{ky} a cooperative awareness message containing the LBSP v_k is connected to and the kind of LBS request it is using. Then, v_{ky} will help v_k send fake LBS requests with v_{ky}'s pseudonyms and locations to the LBSP. We assume that all vehicles are willing to collaborate since a selfish vehicle who does not collaborate will threaten its own privacy [18]. Therefore, motivation schemes will not be considered in our work.

As we have mentioned, the scenario where not all vehicles are using the same LBS should be considered when designing algorithms to preserve trajectory privacy. Under this scenario, if the selected collaborator v_{ky} is not using the same LBS as v_k and v_{ky} just sends one fake request to the LBSP, it will be easy for the adversary to identify the fake request. To deal with this problem, the collaborator should keep sending fake LBS requests for a period of time.

In our design, v_{ky}, which is selected as the collaborator at an intersection, is required to keep sending fake LBS requests until arriving at the next intersection. Once the attacker mistakenly tracks v_{ky}, its predicted trajectory may create more divergences at the next intersection. In this way, trajectory privacy of vehicles will be better preserved. The distance between two intersections is generally more than 600 meters in the main urban roads, and the distance between the two gateways of the expressway is even further [19]. According to [20], the average speed of vehicles on weekday in first-tier cities is about $24\,(km/h)$, while it can reach $30 - 45\,(km/h)$ in other major cities. Therefore, we assume that the average speed of vehicles is $10\,(m/s)$ and the time required for a collaborator to keep sending fake requests is $60s$.

Input: $S_k^j = \{s_k^j, s_k^{j-1}, \ldots, s_k^{j-N}\}$, $\mathbf{P}_{j-N} = \mathbf{E}_4$
Output: $\overline{\mathbf{Loc}}_k^{j+1}$
(1) $\mathbf{x}_k^{j-N} = (Lo_k^{j-N}, La_k^{j-N}, Vo_k^{j-N}, Va_k^{j-N})^T$
(2) **for** $l \leftarrow j - N$ to $j - 1$ **do**
(3) $\mathbf{u}_k^l = (Ao_k^j, Aa_k^j)^T$
(4) $\overline{\mathbf{x}}_k^{l+1} = \mathbf{A}\mathbf{x}_k^l + \mathbf{B}\mathbf{u}_k^l$
(5) $\mathbf{Z}_k^{l+1} = (Lo_k^{l+1}, La_k^{l+1})^T$
(6) $\overline{\mathbf{P}}_{l+1} = \mathbf{A}\mathbf{P}_l\mathbf{A}^T + \mathbf{Q}$
(7) $\mathbf{K} = \overline{\mathbf{P}}_{l+1}\mathbf{H}^T(\mathbf{H}\overline{\mathbf{P}}_{l+1}\mathbf{H}^T + \mathbf{R})^{-1}$
(8) $\mathbf{x}_k^{l+1} = \overline{\mathbf{x}}_{l+1} + \mathbf{K}(\mathbf{Z}_{l+1} - \mathbf{H}\overline{\mathbf{x}}_{l+1})$
(9) $\mathbf{P}_{l+1} = (\mathbf{E}_4 - \mathbf{K}\mathbf{H})\overline{\mathbf{P}}_{l+1}$
(10) **end for**
(11) $\overline{\mathbf{Loc}}_k^{j+1} = \mathbf{H}(\mathbf{A}\mathbf{x}_k^j + \mathbf{B}\mathbf{u}_k^j)$

Algorithm 1: Location prediction process of vehicle v_k.

6. Performance Evaluation

In this section, we conduct traffic simulation to evaluate the performance of the proposed trajectory privacy-preserving scheme. 219 vehicles drive on an 6 km × 6 km map of Suzhou, China, as shown in Figure 2, using SUMO [21, 22]. We consider that a LBSP collects the received LBS requests for 5 minutes. Within these 5 minutes, vehicles using this LBS move randomly on the map following traffic rules. The LBS requests are sent by the vehicles every T_{int}. To investigate the effects of the density of vehicles using the LBS on the performance of our design, four situations, where 25%, 50%, 75%, and 100% vehicles on the road are using the LBS, are set.

As described in Section 5, a vehicle v_k records its locations, velocities, and accelerations of N most recent historical time points. This information of N time points will be used to predict v_k's location at the future time and the collaborator will be selected based on the predicted location. A proper set of N will help v_k make an accurate prediction, thus selecting a capable collaborator, with small memory and computation cost. To set a proper value of N, we try to make location predictions using the moving information of the 219 vehicles. As shown in Figure 3, prediction deviation, the average distance difference between predicted locations and real locations, changes with N. Generally, smaller deviation comes with greater N. This is intuitive since more information leads to more precise prediction. However, fluctuation occurs in Figure 3. This is because that vehicles cannot stay on the same moving pattern on different points of the road. For example, the velocity tends to be stable or change slightly between intersections, while the velocity tends to decrease and the acceleration tends to be stable or change slightly when a vehicle approaches the intersections. These complex road conditions cause fluctuations in the results of our predictions of vehicles' full trajectories. Therefore, if N is not set properly, a vehicle will use the moving pattern between intersections to predict the moving pattern approaching intersections leading to imprecise predictions and increased deviations. According to the results shown in Figure 3, N is set to be 5 in our simulation.

6.1. Performance Metrics. Two metrics are designed to evaluate the trajectory privacy-preserving performance. First, tracking success ratio is designed to represent the possibility that the actual trajectory is v_k's trajectory in the

Figure 2: Map used in simulation.

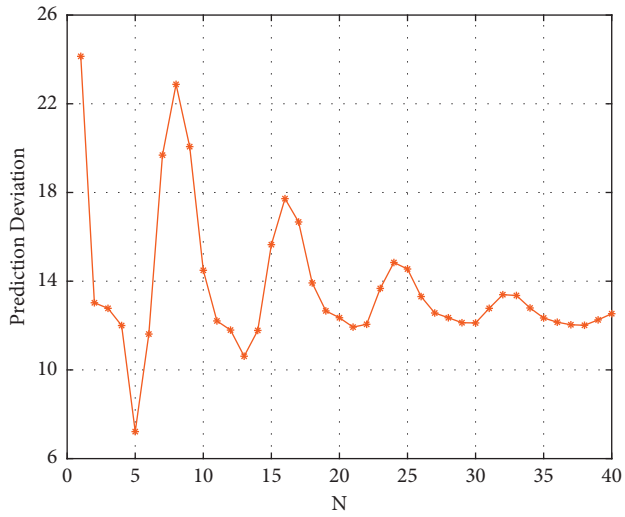

Figure 3: Prediction deviation with different N values.

adversary's eyes. For a time t, the tracking success ratio is defined as follows:

$$SR_t = \prod_{j=0}^{t} p(k, j), \tag{13}$$

where k satisfies that location (Lo_k^j, La_k^j) is on the trajectory.

Then, we propose a metric to investigate the trajectory privacy-preserving efficiency. For collaborative obfuscation, collaborators are required to send fake LBS requests. Therefore, the number of collaborators selected by v_k is considered as the cost to preserve v_k's trajectory privacy. Let NCV_t denote the number of selected collaborators until time t, and the privacy-preserving efficiency E_t is defined as follows:

$$E_t = \frac{1 - SR_t}{NCV_t}. \tag{14}$$

6.2. Performance Analysis and Comparison. We first investigate the average tracking success ratio of each density. As can be seen from Figure 4, tracking success ratio of

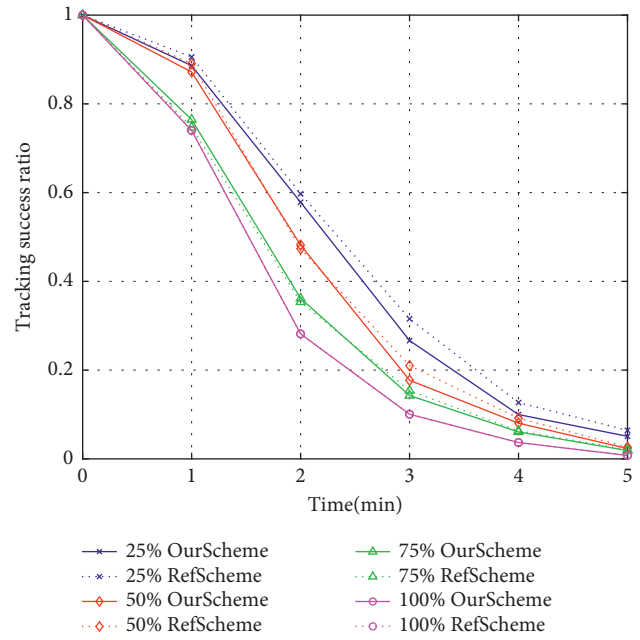

Figure 4: Tracking success ratio *v.s.* The density of vehicles using the LBS.

each density decreases gradually and reduces to almost 0 over time. The tracking success ratio of low density is obviously higher than that of high density. This result confirms the influence of the number of obfuscation locations on privacy preserving. More vehicles using the same LBS will lead to more LBS requests with similar locations submitted at a specific time. These requests can help each other mislead the adversary. That is, the actual LBS requests from other vehicles play the role of fake requests for v_k.

The tracking success ratio of our design is also compared with that of the scheme proposed in [10], referred to RefScheme in the rest of this section. It can be seen from Figure 4 that our scheme holds similar tracking success ratio with RefScheme and our scheme performances slightly worse than RefScheme before 2 minutes under 50% and 75% densities. It is because that vehicles generate fake locations all the time on their way in RefScheme, while our scheme only works at intersections. Less collaborators are selected, and less fake LBS requests are sent leading to a slightly worse performance with less communication cost for collaborator selection and request sending. However, our design outperforms RefScheme when the density of vehicles using the LBS is low. This is because that more confusing trajectories are generated by our scheme. When the density reaches 100%, because the vehicles on the road are using the same LBS and the location updates received by LBSP under these two schemes are the same, the tracking success ratio of the two schemes is the same.

Figure 5 shows the efficiency defined in (14) of two schemes. We can see that the efficiency of both is 0 at first since $SR_0 = 1$. Clearly, our scheme outperforms RefScheme since a similar tracking success ratio can be achieved by our

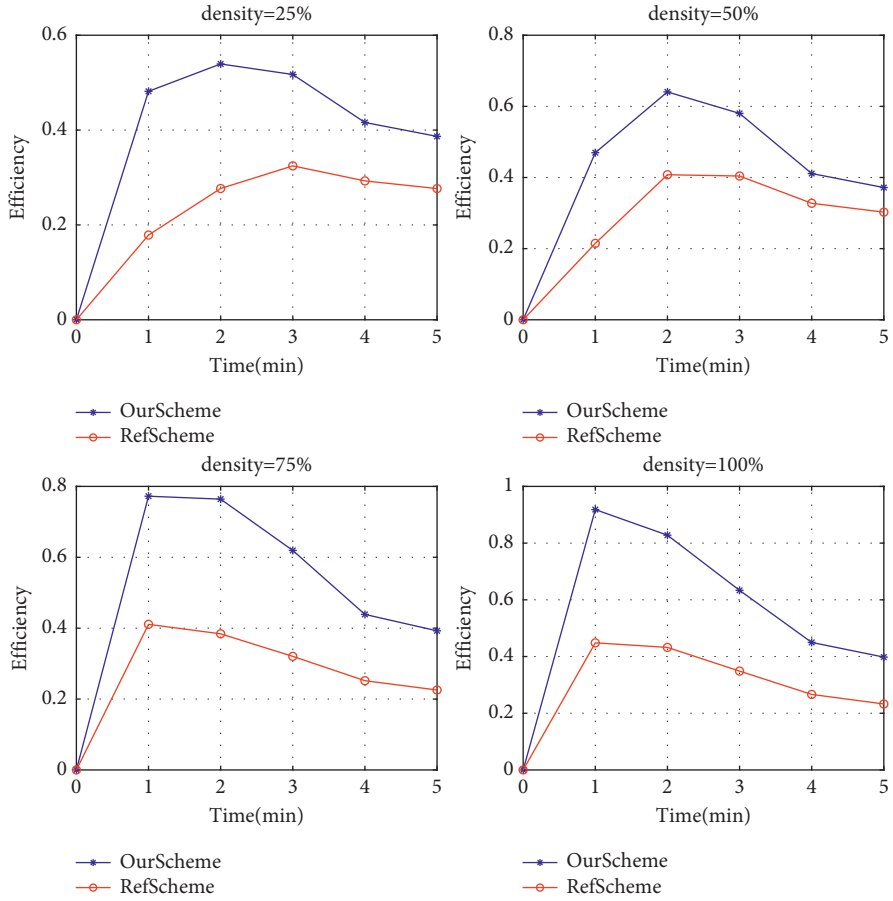

Figure 5: Privacy-preserving efficiency for different densities.

Table 2: Qualitative performance comparison.

Scheme	Service quality	Privacy performance
The proposed scheme	High	High
Silent period [7]	Low	High
Mix zone [15]	High	Low
Location obfuscation [17]	Low	High

scheme with less collaborators. It should be noted that all curves in Figure 5 are not monotonous. This is because that *NCV* does not increase monotonically due to the randomness of vehicle movement.

Moreover, the proposed scheme can be qualitatively compared with other existing methods, such as silent period [7], mix zone [15], and location obfuscation [17] from the perspectives of privacy performance and service quality. The comparison result is shown in Table 2. References [7, 17] provide good privacy protection with compromise on service quality since no service request is sent during the silent period and service is provided based on deviated locations, respectively. The service quality of scheme [15] can be guaranteed, but it is vulnerable to continuous tracking attacks. Our scheme overcomes these drawbacks since fake requests from collaborators protect trajectory privacy, while actual requests ensure service quality.

7. Conclusion

This study proposes a collaborative trajectory obfuscation scheme based on analyzing the attack model of adversaries. Compared with exiting works, our design has high service quality and high privacy strength and can preserve the trajectory privacy with less cost, that is, fewer collaborators. To better reproduce the real usage scenario where not all vehicles are using the same LBS, we introduce the density of vehicles using the same LBS as a variable in our simulation. The results show that low density will increase the risk of trajectory exposure, and therefore, density should be considered when designing trajectory privacy-preserving schemes.

Data Availability

The data that support the findings of this study are derived from the following resources available in the public domain: https://www.openstreetmap.org/and http://sourceforge.net/projects/sumo/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This work was supported by the Natural Science Foundation of Jiangsu Province (BK20211357), Future Network Scientific Research Fund Project (FNSRFP-2021-YB-40), and Natural Science Foundation of China (grant no. 61802274).

References

[1] J. Liang and W. Wang, "Security and privacy in vehicular ad hoc network and vehicle cloud computing: a survey," *Wireless Communications and Mobile Computing*, vol. 2020, p. 2020.

[2] A. Ullah, X. Yao, S. Shaheen, and H. Ning, "Advances in position based routing towards its enabled fog-oriented vanet: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 2, pp. 828–840, 2019.

[3] S. Khan, I. Sharma, M. Aslam, M. Z. Khan, and S. Khan, "Security challenges of location privacy in VANETs and state-of-the-art solutions: a survey," *Future Internet*, vol. 13, no. 4, p. 96, 2021.

[4] T. Gao and L. Zhao, "Pseudonym schemes based on location privacy protection in vanets: a survey," *Innovative Mobile and Internet Services in Ubiquitous Computing*, vol. 55, pp. 597–605, 2020.

[5] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in inter-vehicular networks: why simple pseudonym change is not enough," in *Proceedings of the Seventh International Conference on Wireless On-Demand Network Systems and Services (WONS)*, pp. 176–183, Kranjska Gora, Slovenia, February 2010.

[6] S. Krishna, L. Huang, and M. Li, *Caravan: Providing Location Privacy for Vanet*, Washington University Seattle Department of Electrical Engineering, Washington, DC, USA, 2005.

[7] S. Krishna, M. Li, L. Huang, and R. Poovendran, "Amoeba: robust location privacy scheme for vanet," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1569–1589, 2007.

[8] C. Kalaiarasy, N. Sreenath, and A. Amuthan, "Location privacy preservation in vanet using mix zones: a survey," in *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–5, Tamilnadu, India, December 2019.

[9] X. Zhang, X. Gui, and Z. Wu, "Survey of privacy protection research for location services," *Journal of Software*, vol. 09, pp. 223–245, 2015.

[10] J. Cui, J. Wen, S. Han, and H. Zhong, "Efficient privacy-preserving scheme for real-time location data in vehicular ad-hoc network," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3491–3498, 2018.

[11] I. Wagner and D. Eckhoff, "Technical privacy metrics: a systematic survey," *ACM Computing Surveys*, vol. 51, no. 3, pp. 1–38, 2018.

[12] H. To, K. Nguyen, and C. Shahabi, "Differentially private publication of location entropy," in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 1–10, Redondo Beach, CA, USA, November 2016.

[13] G. P. Corser, H. Fu, and A. Banihani, "Evaluating location privacy in vehicular communications and applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2658–2667, 2016.

[14] T. Hassan, T. Nomani, M. Mohsin, and Saira Sattar, "A survey on location privacy techniques deployed in vehicular networks," in *Proceedings of the 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pp. 604–613, Islamabad, Pakistan, January 2019.

[15] R. Lu, X. Lin, H. Tom, X. Liang, and X. Shen, "Pseudonym changing at social spots: an effective strategy for location privacy in vanets," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2011.

[16] Yi Huang, H. Zheng, and X.-F. Meng, "Coprivacy: a collaborative location privacy-preserving method without cloaking region," *Jisuanji Xuebao*, vol. 34, no. 10, pp. 1976–1985, 2011.

[17] N. Takbiri, A. Houmansadr, D. L. Goeckel, and H. Pishro-Nik, "Limits of location privacy under anonymization and obfuscation," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, pp. 764–768, Victoria, Australia, July 2017.

[18] J. Lim, H. Yu, K. Kim, M. Kim, and S.-B. Lee, "Preserving location privacy of connected vehicles with highly accurate location updates," *IEEE Communications Letters*, vol. 21, no. 3, pp. 540–543, 2016.

[19] Ministry of Housing and PRC Urban-Rural Development, *CJJ129 Specification for Design of Urban Expressway*, China Architecture and Building Press, Beijing, China, 2009.

[20] Amap, *Urban Transportation Report"*, Beijing Traffic Information Center, Beijing, China, 2020.

[21] Steve Coast. "Open street map". [Online]. Available: https://www.openstreetmap.org/.

[22] Eclipse Foundation. "Simulation of urban mobility". [Online]. Available: http://sourceforge.net/projects/sumo/.