Article

A Privacy-Preserving Traffic Monitoring Scheme via Vehicular Crowdsourcing

Chuan Zhang [1], Liehuang Zhu [1,*], Chang Xu [1,*], Xiaojiang Du [2] and Mohsen Guizani [3]

[1]   School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China; chuanz@bit.edu.cn

[2]   Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122, USA.; dxj2005@gmail.com

[3]   Department of Computer Science and Engineering, Qatar University, Doha 2713, Qatar; mguizani@ieee.org

*   Correspondence: liehuangz@bit.edu.cn (L.Z.); xuchang@bit.edu.cn (C.X.)

check for updates

Abstract: The explosive number of vehicles has given rise to a series of traffic problems, such as traffic congestion, road safety, and fuel waste. Collecting vehicles' speed information is an effective way to monitor the traffic conditions and avoid vehicles' congestion, however it may threaten vehicles' location and trajectory privacy. Motivated by the fact that traffic monitoring does not need to know each individual vehicle's speed and the average speed would be sufficient, we propose a privacy-preserving traffic monitoring (PPTM) scheme to aggregate vehicles' speeds at different locations. In PPTM, the roadside unit (RSU) collects vehicles' speed information at multiple road segments, and further cooperates with a service provider to calculate the average speed information for every road segment. To preserve vehicles' privacy, both homomorphic Paillier cryptosystem and super-increasing sequence are adopted. A comprehensive security analysis indicates that the proposed PPTM can preserve vehicles' identities, speeds, locations, and trajectories privacy from being disclosed. In addition, extensive simulations are conducted to validate the effectiveness and efficiency of the proposed PPTM scheme.

Keywords: traffic monitoring; speed; privacy-preserving; vehicular crowdsourcing

1. Introduction

Nowadays, the number of global vehicles has exceeded 1.2 billion and may be headed to 2 billion by 2035 [1]. With such a large number of vehicles, many critical social problems, such as traffic congestions and slow traffic, have emerged, leading to significant time and fuel waste. According to a report released by Harvard Center, for the drivers in the 10 most-congested cities in USA, more than 48 h are wasted in traffic jams, causing \$121 billion loss in time and fuel every year [2]. To deal with these critical problems, both industry and academia are paying great attention to traffic monitoring, and the vehicular ad hoc network (VANET) is considered as one of the most promising ways that can be leveraged in traffic management [3,4].

In VANETs, vehicles, embedded with onboard units (OBUs), can share traffic information (e.g., locations and speeds) to the roadside units (RSUs) through vehicle-to-infrastructure (V2I) communications, and nearby vehicles by vehicle-to-vehicle (V2V) communications [5]. By collecting and analyzing this traffic information, vehicles can easily know different locations' traffic conditions and road safety, and accordingly plot their optimal routes. Recently, several VANET-based traffic monitoring applications have been built. For example, Google and Apple provide real-time navigation services based on current traffic information [6]. Waze has developed an application that can help drivers get the best route with real-time help from other drivers [7]. Although many benefits can be

brought by this emerging network paradigm, its adoption still hinges on how to resolve security and privacy concerns for the users. Since a vehicle's location is tightly bundled with its driver, an attacker can predict a driver's future location based on his vehicle's trajectory, or even infer the drivers' personal information, such as habits, health condition, income, and religious belief, according to their frequently visited places.

To preserve the vehicles' privacy, pseudonyms and anonymous authentication are two effective ways to conceal vehicles' real identities and realize conditional privacy preservation [8–11]. For example, Ni et al. [12] proposed a privacy-preserving real-time navigation system by collecting vehicles' location and speed information, and, with the randomization technique, the sensitive identity privacy is preserved. However, the work in [13] shows that user identities can sometimes be inferred from the location data if users' home and work locations are deduced from the data. Moreover, we observe an attack that, by linking vehicles' speed information, vehicles can also be identified even if they change their pseudonyms. An example is illustrated in Figure 1. At time t_1, a vehicle provides its speed information $PID_A||v_1||L_1||t_1$ to a roadside unit (RSU), and at time t_2, it uploads the speed information $PID_B||v_2||L_2||t_2$, where v denotes the average speed in the road segment, L denotes the location, and t represents the current time. Although the vehicle's pseudonym is changed (i.e., $PID_A \rightarrow PID_B$), attackers can still link the pseudonyms by comparing the estimated passing time (the average passing time can be calculated by using the distance and average speed, and the distance from L_1 to L_2 can be obtained from GPS) and actual passing time (i.e., $t_2 - t_1$) between these two locations. Thus, there still lacks a privacy-preserving traffic monitoring scheme that can protect the vehicles' identities and defend against the linkable attack.

In this paper, to deal with the above challenges, we propose a privacy-preserving traffic monitoring (PPTM) scheme to enable vehicles provide their traffic information without sacrificing their privacy. This scheme uses the homomorphic Paillier cryptosystem to guarantee the privacy of vehicles' speeds, and adopts a well-defined super-increasing sequence to not only protect vehicles' location privacy, but also save tremendous computational costs and communication overhead. Our main contributions can be further summarized below.

- First, inspired by the fact that the average speed would be sufficient to achieve traffic monitoring, we propose PPTM, which uses the super-increasing sequence and homomorphic Paillier cryptosystem to realize privacy-preserving speed aggregation and efficient traffic monitoring. Concretely, each vehicle uses a well defined super-increasing sequence to aggregate its multiple speeds and encrypts the aggregated result before uploading it to the RSU. Then, the RSU will aggregate all reports and cooperate with a service provider to calculate each road segment's average speed. During this process, vehicles' identity, speed, and location privacy will not be disclosed to any other party.

- Second, we find that the anonymous technologies such as pseudonyms and randomizable signature are not suitable for certain VANET-based applications because of the time link attack. To mitigate this attack, we design a privacy-preserving data aggregation approach. Through a comprehensive security analysis, the proposed PPTM is proven to be secure and privacy preservation. Particularly, the proposed scheme can achieve report privacy preservation, report authentication, data integrity, and identity preservation, as well as defend against the collusion attack. The detailed analysis is given in Section 5.

- Third, we conducted extensive simulations to show PPTM is practical and efficient. Compared with a traditional baseline scheme, PPTM could significantly reduce computational costs and communication overhead, indicating that the proposed scheme can indeed realize real-time traffic monitoring.

The rest of this paper is organized as follows. In Section 2, we introduce the system model, security requirement, and design goals of the proposed PPTM scheme. In Section 3, preliminaries including bilinear pairings and Paillier cryptosystem are introduced. The detailed introduction of

PPTM is given in Section 4. In Sections 5 and 6, we analyze the security and performance of PPTM, respectively. In Section 7, some related works are listed, and we draw our conclusion in Section 8.
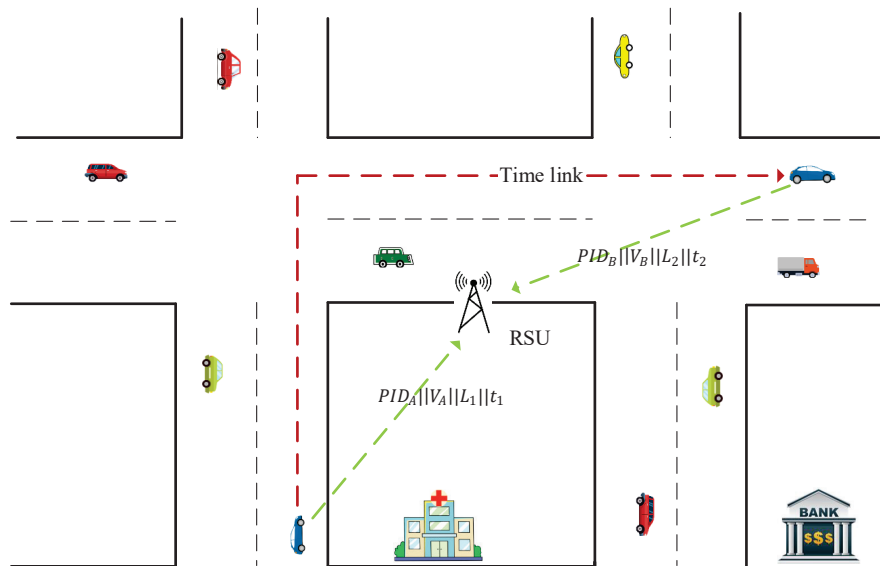

Figure 1. Using passing time to link pseudonyms.

2. System Model, Security Requirements, and Design goals

In this section, we formalize the proposed scheme by giving the system model, threat model, and design goals.

2.1. System Model

In the proposed PPTM scheme, roads are divided into multiple segments and vehicles are expected to provide their average speed for each segment they have passed through. A typical RSU-assisted VANET application is illustrated in Figure 2. In particular, the considered system model consists of the following entities.

- Trust authority (TA): TA is a fully trusted entity that is responsible for the registration of vehicles and RSUs. It builds public/secret key pairs for all entities, and generates sufficient pseudonyms for vehicles before their next registration.
- Vehicles: All vehicles are embedded with OBUs, which enable them to communicate with RSUs and surrounding vehicles through wireless communications. Besides, they also have the ability to generate and run their own homomorphic cryptosystem.
- RSUs: RSUs act as the role of access points, which are widely deployed in the urban area. They can communicate with vehicles via wireless communications and the service provider by wired communications.
- Service provider (SP): SP is a centralized entity that is responsible for providing traffic monitoring services. It connects with all RSUs through fast communication technology, such as wired cables.
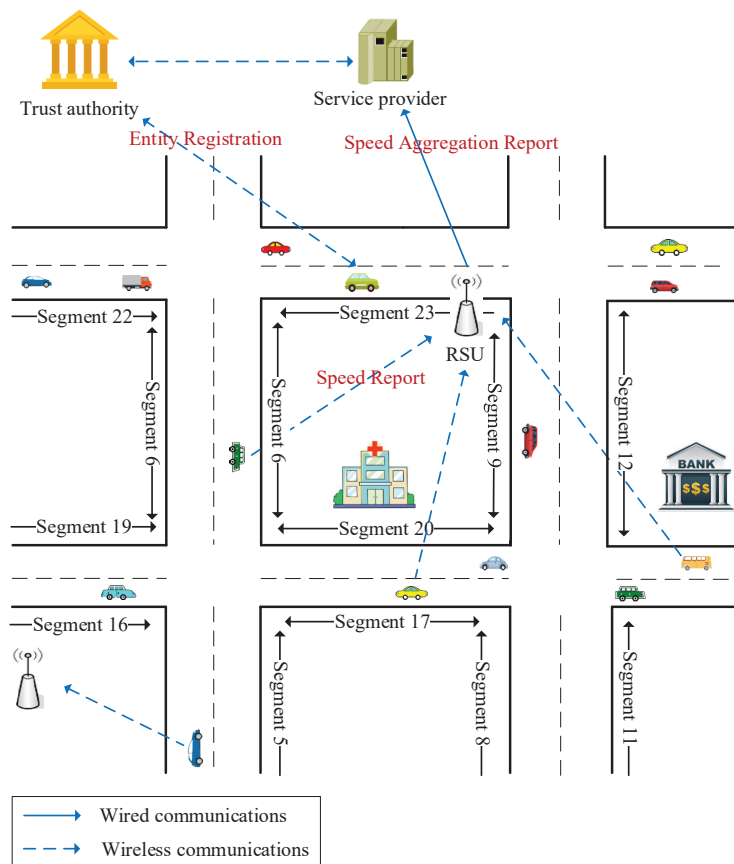
Figure 2. Using trust values to link pseudonyms in a given time period.

2.2. Security Requirements

In our security model, TA is fully trusted as it is responsible for initializing the whole system and generating credentials and public/private keys for all participating entities. SP and RSUs are considered to be honest-but-curious, which means both will strictly follow the designed protocol, but are curious about vehicles' privacy. In particular, we assume there is no collusion between SP and RSUs, which is similar to most existing RSU-assisted scenarios [14,15]. Meanwhile, we assume that vehicles will provide correct speed information to the RSU. This assumption is reasonable in most traffic monitoring scenarios, since: (1) the speed provided by vehicles is in the area where they have passed through, and providing false data would not benefit them; and (2) vehicles want to know the correct traffic conditions, and thus will honestly follow the designed protocol for their mutual benefits. Besides, we also assume there exists an attacker which is curious about drivers' privacy. It may launch attacks, modify speed reports, and threaten data integrity. Based on the above assumptions, the proposed scheme should achieve the following security requirements.

Identity Privacy Preservation. As described above, an attacker can potentially identify drivers even though they adopt pseudonyms and anonymous authentication. Thus, to preserve drivers' identity privacy, attackers cannot infer vehicles' location information (i.e., road segments) based on the given data.

Location Privacy Preservation. Since the speed is location-aware, preserving drivers' location privacy requires preventing their speed from being disclosed. Hence, the proposed scheme should ensure that, even if the RSU or an attacker receives a vehicle's speed information, it cannot recover its speed and further infer its location privacy.

Data Integrity. An attacker may eavesdrop drivers' reports and modify them for its benefits. Thus, the proposed scheme should guarantee data integrity and any malicious operations should be detected.

2.3. Design Goals

Based on the aforementioned security requirements, our goal is to design a privacy-preserving traffic monitoring scheme, which enables vehicles to upload their speeds towards the RSU securely and efficiently. Concretely, the proposed scheme should achieve the following two design goals.

The defined security requirements should be guaranteed. If the proposed scheme fails to realize the aforementioned security requirements, drivers' identity and location privacy may be disclosed, and data reports transmitted to the RSU or other vehicles may be modified. Then, vehicles may be reluctant to provide their speed, and traffic conditions will not be accurately estimated.

High efficiency should be guaranteed. To provide real-time traffic monitoring, vehicles are expected to upload speed information in a short transmission interval. However, to preserve drivers' privacy, sensitive information should be encrypted, which may introduce tremendous computational costs and bandwidth consumption for resource-constrained vehicles. Thus, the proposed scheme should achieve high efficiency in computational costs and communication overhead.

3. Preliminaries

In this section, we review the pairing-based cryptography [16] and the Paillier cryptosystem [17], which serve as the basis of our proposed traffic monitoring scheme.

3.1. Bilinear Pairings

Suppose there are two cyclic groups $\mathbb{G}_1$ and $\mathbb{G}_2$, both of which share a same order q. Then, a bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ has the following properties.

- Bilinearity: $e(aP, bQ) = e(P, Q)^{ab} \in \mathbb{G}_2$, for all $P, Q \in G_1$ and $a, b \in \mathbb{Z}_q^*$.
- Non-degeneracy: $e(P, P) \neq 1$, for all $P \in \mathbb{G}_1$.
- Computability: $e(P, Q)$ can be efficiently computed, for all $P, Q \in \mathbb{G}_1$.

By referring to [18,19], we give two more comprehensive definitions for bilinear pairings.

Definition 1. *Given an input security parameter κ, $\mathcal{G}en$ is a probabilistic algorithm to output a 5-tuple $(q, P, \mathbb{G}_1, \mathbb{G}_2, e)$, in which q is a κ-bit prime, P is a generator, $(\mathbb{G}_1, \mathbb{G}_2)$ are two cyclic groups sharing a same order q, and $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ is an efficient, computable, and non-degenerated bilinear map.*

Definition 2 (Computational Diffie–Hellman (CDH) Problem). *Given elements $(P, aP, bP) \in \mathbb{G}_1$, there exists no effective algorithm can calculate $abP \in \mathbb{G}_1$ for unknown $a, b \in \mathbb{Z}_q^*$ in a probabilistic and polynomial time.*

3.2. Paillier Cryptosystem

As an effective technology to achieve homomorphic properties on the ciphertexts, Paillier cryptosystem has been widely used in various privacy-preserving applications. Concretely, three algorithms are included in the Paillier cryptosystem.

- *Key Generation:* With a security parameter κ_1, select two large κ_1-bit primes p_1, q_1, and calculate $n = p_1 q_1$ and the least common multiple of p_1 and q_1, i.e., $\lambda = lcm(p_1, q_1)$. Then, define a function $L(a) = \frac{a-1}{n}$, and calculate $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n^2$, where $g \in \mathbb{Z}_n^*$. Then, the public/private keys are $pk = (n, g)$ and $sk = (\lambda, mu)$.
- *Message Encryption:* Given a plaintext $m \in \mathbb{Z}_n$, after choosing a random value $r \in \mathbb{Z}_n^*$, the message is encrypted as $c = E(m) = g^m \cdot r^n \bmod n^2$.
- *Ciphertext Decryption:* Given a ciphertext $c = E(m) \in \mathbb{Z}_{n^2}^*$, the message is recovered as $m = D(c) = L(c^{\lambda \bmod n^2}) \cdot \mu \bmod n$.

Note that Paillier cryptosystem has been proven to be correct, secure, and effective against the chosen plaintext attack [17]. Moreover, Paillier cryptosystem allows arithmetic operations on ciphertexts, such as $E(m_1) \cdot E(m_2) = E(m_1 + m_2)$ and $E(m_1)^a = E(a \cdot m_1)$, for all $(m_1, m_2) \in \mathbb{Z}_n^*$.

4. Proposed PPTM Scheme

In this section, we give the details of the proposed PPTM scheme, which includes system initialization, speed request and speed reporting, privacy-preserving report aggregation, secure report reading, and traffic guidance and identity tracing.

4.1. System Initialization

TA initializes the whole system. After selecting two security numbers κ, κ_1, it first runs $\mathcal{Gen}(\kappa)$ to generate a 5-tuple $(q, P, \mathbb{G}_1, \mathbb{G}_2, e)$ and calculates public/private keys of the Paillier cryptosystem, i.e., $pk = (n, g), sk = (\lambda, \mu)$, according to κ_1. Then, TA selects a secure cryptographic hash function H, where $H : \{0,1\}^* \rightarrow \mathbb{G}_1$. Vehicles are required to register themselves periodically. TA chooses a secure key k_0 and generates a secure symmetric encryption algorithm AES_{k_0}. For every registered vehicle with its real identity number ID_i (the real ID can be license number or social secure number), TA generates a group of pseudonyms $\{PID_{ij} = AES_{k_0}(ID_i||x_{ij})\}_{j=1}^n$ by choosing a set of random values $\{x_{ij}\}_{j=1}^n \in \mathbb{Z}_q^*$. Then, TA uses x_{ij} as each vehicle's certified private key and calculates the corresponding public key as $Y_{ij} = x_{ij}P$. For each RSU with its identity number ID_r, TA selects a random number $x_r \in \mathbb{Z}_q^*$ as its private key and calculates the public key as $Y_r = x_r P$. Finally, TA publishes $\{(P, \mathbb{G}_1, \mathbb{G}_2, e, H), (n, g)\}$ and sends $\{\{PID_{ij}, x_{ij}, Y_{ij}\}_{j=1}^n, (IDr, Y_r)\}$ to each vehicle, $\{ID_r, x_r, Y_r\}$ to each RSU, and $(\lambda, \mu, ID_r, Y_r)$ to SP.

In addition, in the coverage of an RSU, roads are divided into multiple segments. Assume that the maximum number of segments within the coverage of each RSU is M, the number of vehicles in every segment is no more than Q, and the maximum speed in every road segment is smaller than V. Then, for the segments located in each RSU's coverage, TA generates a super-increasing sequence $\overrightarrow{a} = (a_1, a_2, \cdots, a_M)$, where a_i denotes the ith segment such that $a_i \in \mathbb{Z}_n^*$ is randomly chosen, $\sum_{i=1}^{j-1} a_i \cdot Q \cdot V < a_j$ for $j = 2, 3, \cdots, M$, and $\sum_{i=1}^{M} a_i \cdot Q \cdot V < n$.

4.2. Speed Request and Speed Reporting

Figure 3 illustrates the system procedure of PPTM. As can be seen, RSU first generates a speed request and all vehicles response it by providing their driving reports. Specifically, the request contains the RSU's ID, the current timestamp TS, the time range TR, and the signature $\sigma_r = x_r H(ID_r||TS||TR)$. Note that, the timestamp is used to defend against the replay attack launched by forged RSUs. Then, the RSU broadcasts the request $R_r = ID_r||TS||TR||\sigma_r$ to vehicles driving in its communication coverage. After receiving this request, vehicles first verify the report by examining whether $e(P, \sigma_r)$ equals to $e(Y_r, H(ID_r||TS||TR))$. If the equation holds, the request will be accepted, since $e(P, \sigma_r) = e(x_r P, H(ID_r||TS||TR)) = e(Y_r, H(ID_r||TS||TR||))$.

If the request is valid, vehicles are expected to provide their speed reports. The format of speed is defined as $\{(i, t_i(j), A_i(j), S_i(j))\}_{i=1,j=1}^{M,V}$, where $t_i(j)$ denotes the time passing by the ith segment for the vehicle $\mathcal{V}_j$, and $\{A_i(j), S_i(j)\}$ are calculated as follows,

$$
\begin{cases}
A_i(j) = \begin{cases} 1 & \text{if } \mathcal{V}_j \text{ has passed by the segment } i \\ 0 & \text{if } \mathcal{V}_j \text{ has not passed by the segment } i \end{cases} \\
S_i(j) = \begin{cases} v_i(j) & \text{if } \mathcal{V}_j \text{ has passed by the segment } i \\ 0 & \text{if } \mathcal{V}_j \text{ has not passed by the segment } i \end{cases}
\end{cases}
$$

where $v_i(j)$ is $\mathcal{V}_j$'s average speed when passing by the segment i. Note that the speed report is time series data, which means the reports will be sorted in time order. For example, if a vehicle has gone

through the segments in order of 1, 4, and 5, with the average speed 50, 75, and 60, its speed reports are represented as $\{(1,5,1,50),(4,3.5,1,75),(5,2,1,60)\}$. Then, given a time range as 8, the vehicle should submit the speed report $\{(5,2,1,60),(4,3.5,1,75)\}$ since $2 + 3.5 < 8$. To preserve the privacy of location and speed privacy, the report should be encrypted before uploaded to the RSU. The vehicle $\mathcal{V}_j$ selects two random values $r_{j1}, r_{j2} \in \mathbb{Z}_n^*$ and calculates the ciphertexts as $C_{j1} = g^{(a_1 \cdot A_1(j) + \cdots + a_M \cdot A_M(j))} \cdot r_{j1}^n \mod n^2$ and $C_{j2} = g^{(a_1 \cdot S_1(j) + \cdots + a_M \cdot S_M(j))} \cdot r_{j2}^n \mod n^2$. Then, the vehicle signs the report with its secret key by computing $\sigma_j = x_j H(PID_j || Y_j || C_{j1} || C_{j2} || TS)$. After that, $\mathcal{V}_j$ delivers the speed report $R_j = PID_j || Y_j || C_{j1} || C_{j2} || TS || \sigma_j$ to the RSU.
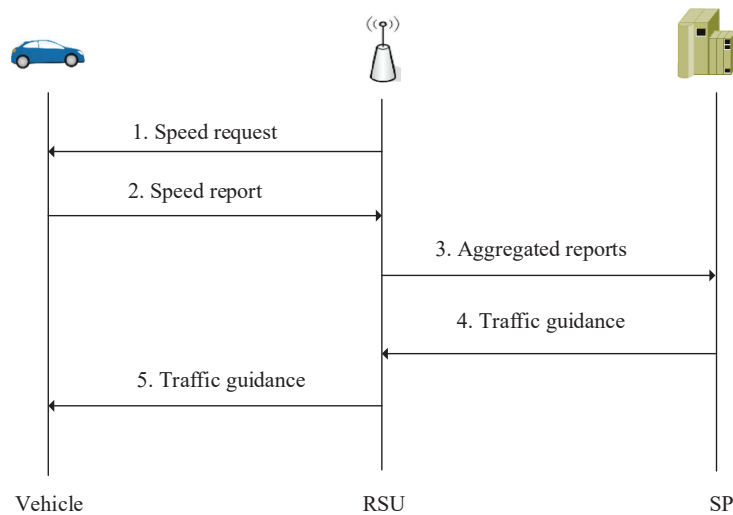

Figure 3. System procedure of PPTM.

4.3. Privacy-Preserving Report Aggregation

Upon receiving the report, RSU first checks the freshness of this report, i.e., to make sure that the difference between request and response is within a certain range. Then, the RSU verifies the vehicle's report by examining $e(P, \sigma_j) \overset{?}{=} e(Y_j, H(PID_j || Y_j || C_{j1} || C_{j2} || TS))$ as $e(P, \sigma_j) = e(x_j P, H(PID_j || Y_j || C_{j1} || C_{j2} || TS)) = e(Y_j, H(PID_j || Y_j || C_{j1} || C_{j2} || TS))$. Especially, to improve efficiency, RSU can perform batch verification to check $e(P, \sum_{j=1}^N \sigma_j) \overset{?}{=} \prod_{j=1}^N e(Y_j, H(PID_j || Y_j || C_{j1} || C_{j2} || TS))$, where N is the number of vehicles passing by every segment. The proof is given below.

$$
\begin{aligned}
e(P, \sum_{j=1}^N \sigma_j) &= e(P, \sum_{j=1}^N x_j H(PID_j || Y_j || C_{j1} || C_{j2} || TS)) \\
&= \prod_{j=1}^N e(P, x_j H(PID_j || Y_j || C_{j1} || C_{j2} || TS)) \\
&= \prod_{j=1}^N e(Y_j, H(PID_j || Y_j || C_{j1} || C_{j2} || TS)).
\end{aligned}
\tag{1}
$$

By performing this operation, fewer time-consuming pairing operations $e(\cdot, \cdot)$ are required (i.e., $2N$ vs. $N + 1$).

After checking the validity of vehicles' reports, the RSU executes the following steps to obtain the aggregated results in a privacy-preserving way.

- *Step 1.* Calculate the aggregated results C_1 and C_2 based on the encrypted data $\{C_{j1}\}_{j=1}^N$ and $\{C_{j2}\}_{j=1}^N$ as follows.

$$\begin{cases} C_1 = \prod_{j=1}^{N} C_{j1} \bmod n^2 \\ \qquad = \prod_{j=1}^{N} g^{a_1 \cdot A_1(j) + \cdots + a_M \cdot A_M(j)} \cdot r_{j1}^n \bmod n^2 \\ \qquad = g^{a_1 \sum_{j=1}^{N} A_1(j) + \cdots + a_M \sum_{j=1}^{N} A_M(j)} \cdot (\prod_{j=1}^{N} r_{j1})^n \bmod n^2 \\ C_2 = \prod_{j=1}^{N} C_{j2} \bmod n^2 \\ \qquad = \prod_{j=1}^{N} g^{a_1 \cdot S_1(j) + \cdots + a_M \cdot S_M(j)} \cdot r_{j2}^n \bmod n^2 \\ \qquad = g^{a_1 \sum_{j=1}^{N} S_1(j) + \cdots + a_M \sum_{j=1}^{N} S_M(j)} \cdot (\prod_{j=1}^{N} r_{j2})^n \bmod n^2 \end{cases} \tag{2}$$

- *Step 2.* Use the secret key x_r to generate a signature as

$$\sigma_r = x_r H(ID_r || C_1 || C_2 || TS). \tag{3}$$

- *Step 3.* Send the aggregated and encrypted data $ID_r || C_1 || C_2 || TS || \sigma_r$ to the SP.

For ease of understanding, we give an example to show how aggregated vehicle and speed are aggregated, as shown in Figure 4. The RSU receives the ciphertexts of four speed reports $\{R_1, R_2, R_3, R_4\}$, each of which contains four segments. After performing the aggregations, the aggregated results of vehicle and speed are the ciphertexts of $a_i \sum_{j=1}^{4} A_i(j)$ and $a_i \sum_{j=1}^{4} S_i(j)$, respectively, where $i \in [1, 4]$. In the following, we show how to recover the aggregated vehicles and speeds for every segment.

	Segment 1		Segment 2		Segment 3		Segment 4
R_1	$\{a_1 * 1, a_1 * 50\}$	+	$\{a_2 * 0, a_2 * 0\}$	+	$\{a_3 * 0, a_3 * 0\}$	+	$\{a_4 * 1, a_4 * 36\}$
R_2	$\{a_1 * 1, a_1 * 60\}$	+	$\{a_2 * 1, a_2 * 80\}$	+	$\{a_3 * 0, a_3 * 0\}$	+	$\{a_4 * 1, a_4 * 30\}$
R_3	$\{a_1 * 0, a_1 * 0\}$	+	$\{a_2 * 1, a_2 * 88\}$	+	$\{a_3 * 1, a_3 * 40\}$	+	$\{a_4 * 1, a_4 * 33\}$
R_4	$\{a_1 * 1, a_1 * 55\}$	+	$\{a_2 * 1, a_2 * 75\}$	+	$\{a_3 * 1, a_3 * 35\}$	+	$\{a_4 * 1, a_4 * 35\}$
Aggregated vehicle	$a_1 * 3$	+	$a_2 * 3$	+	$a_3 * 2$	+	$a_4 * 4$
Aggregated speed	$a_1 * 165$	+	$a_2 * 243$	+	$a_3 * 75$	+	$a_4 * 134$

Figure 4. Traffic aggregation example in PPTM.

4.4. Secure Report Reading

On receiving the aggregated report, SP first checks data validity by examining $e(P, \sigma_r) \stackrel{?}{=} e(Y_r, H(ID_r || C_1 || C_2 || TS))$, and then performs the following steps to recover the aggregated results from the ciphertexts,

- *Step 1.* Considering $M_1 = a_1 \sum_{j=1}^{N} A_1(j) + \cdots + a_M \sum_{j=1}^{N} A_M(j)$, $M_2 = a_1 \sum_{j=1}^{N} S_1(j) + \cdots + a_M \sum_{j=1}^{N} S_M(j)$ and $R_1 = \prod_{j=1}^{N} r_{j1}, R_2 = \prod_{j=1}^{N} r_{j2}$, the aggregated ciphertexts $C_1 = g^{M_1} \cdot$

$R_1^n \bmod n^2, C_2 = g^{M_2} \cdot R_2^n \bmod n^2$ are still valid ciphertexts of Paillier cryptosystem. Hence, the SP can use the secret key (λ, μ) to obtain M_1 and M_2 as

$$
\begin{cases}
M_1 = D(C_1) & = a_1 \sum_{j=1}^{N} A_1(j) + \cdots + a_M \sum_{j=1}^{N} A_M(j) \bmod n^2 \\
M_2 = D(C_2) & = a_1 \sum_{j=1}^{N} S_1(j) + \cdots + a_M \sum_{j=1}^{N} S_M(j) \bmod n^2
\end{cases}
\tag{4}
$$

- *Step 2.* SP then invokes Algorithm 1 to recover the aggregated vehicles $(L_1, L_2, \cdots, L_M)$ and speed $(LS_1, LS_2, \cdots, LS_M)$, where $L_i = \sum_{j=1}^{N} A_i(j)$ and $LS_i = \sum_{j=1}^{N} S_i(j), i \in [1, M]$.

Algorithm 1: Recover the aggregated report.

Input: M_1, M_2, and $\overrightarrow{a}$
Output: $\{L_i\}_{i=1}^{M}$ and $\{LS_i\}_{i=1}^{M}$
1 Set $\mathcal{L}_M = \sum_{i=1}^{M} \sum_{j=1}^{N} a_i A_i(j) \bmod n^2, \mathcal{LS} = \sum_{i=1}^{M} \sum_{j=1}^{N} a_i S_i(j) \bmod n^2$;
2 **for** $i = M, M-1, \cdots 2$ **do**
3 $\quad \mathcal{L}_{i-1} = \mathcal{L}_i \bmod a_i, \ L_i = \frac{\mathcal{L}_i - \mathcal{L}_{i-1}}{a_i}$;
4 $\quad \mathcal{LS}_{i-1} = \mathcal{LS}_i \bmod a_i, \ LS_i = \frac{\mathcal{LS}_i - \mathcal{LS}_{i-1}}{a_i}$;
5 $L_1 = \frac{\mathcal{L}_1}{a_1}; \ LS_1 = \frac{\mathcal{LS}_1}{a_1}$;
6 **return** $\{L_i, LS_i\}_{i=1}^{M}$

The correctness of Algorithm 1. For ease of description, we use the aggregated vehicles to give the correctness analysis. In this algorithm, $\mathcal{L}_M = a_1 \sum_{j=1}^{N} A_1(j) + a_2 \sum_{j=1}^{N} A_2(j) + \cdots a_{M-1} \sum_{j=1}^{N} A_{M-1}(j) + a_M \sum_{j=1}^{N} A_M(j)$. As the number of aggregated vehicles in every segment is smaller than Q, we have

$$
a_1 \sum_{j=1}^{N} A_1(j) + \cdots a_{M-1} \sum_{j=1}^{N} A_{M-1}(j) < (a_1 + \cdots + a_{M-1}) \cdot Q
$$
$$
= \sum_{i=1}^{M-1} Q < a_M.
\tag{5}
$$

Hence, $\mathcal{L}_{M-1} = \mathcal{L}_M \bmod a_M = a_1 \sum_{j=1}^{N} A_1(j) + \cdots + a_{M-1} \sum_{j=1}^{N} A_{M-1}(j)$, and accordingly we have

$$
\frac{\mathcal{L}_M - \mathcal{L}_{M-1}}{a_M} = \frac{a_M \sum_{j=1}^{N} A_M(j)}{a_M} = L_M.
\tag{6}
$$

Following a similar analysis, $L_i = \sum_{j=1}^{N} A_i(j)$ can be proven. In addition, we can prove $LS_i = \sum_{j=1}^{N} S_i(j)$, as it shares the similar procedure as L_i.

4.5. Traffic Guidance and Identity Tracing

After calculating the aggregated vehicles and speeds in all segments, i.e., $(L_1, L_2, \cdots, L_M)$ and $(LS_1, LS_2, \cdots, LS_M)$, the average speed in each segment can be computed as $L_i = \frac{LS_i}{L_i}$. Finally, SP broadcasts the speed information and vehicles can select optimal routes based on the road conditions. In addition, although we assume that all vehicles report their speeds honestly, some vehicles may still upload false traffic data. In this case, the TA can periodically select some speed reports stored in the RSU and recover them to check whether they are truthful. Since vehicles' pseudonyms are generated by using vehicles' real identity ID, malicious vehicles can be easily and quickly identified.

5. Security Analysis

In this section, we give the security analysis of the proposed PPTM scheme. In particular, recalling the aforementioned security requirements, the analysis focuses on how our proposed PPTM scheme can protect each vehicle's report privacy, ensure report authentication and data integrity, and achieve vehicles' identity and location privacy preservation.

The proposed scheme can achieve report privacy preservation. The proposed scheme preserves reports' privacy by using the Paillier cryptosystem. In PPTM, vehicle $\mathcal{V}_j$'s location and speed are formed as C_{j1}, C_{j2}. Since both ciphertexts are valid ciphertexts of Paillier cryptosystem and the Paillier cryptosystem has been proven to be secure under the chosen plaintext attack, the messages are secure and privacy-preserving. That is, although an adversary may eavesdrop a ciphertext, it cannot recover the corresponding message. After receiving all reports from vehicles, instead of recovering each report, the RSU will perform report aggregation and deliver the aggregated ciphertext to the SP. Thus, even though SP holds the secret key, it can only obtain the aggregated result. Therefore, each individual vehicle's report is privacy-preserving in the proposed PPTM scheme.

The proposed scheme can achieve report authentication and data integrity. In our proposed scheme, vehicles' reports and RSU's aggregated report are signed using BLS short signature [20]. Since it has been proven that BSL short signature can defend against the CDH problem [21], our proposed scheme can guarantee the report authentication and data integrity, and any malicious behavior on the vehicles' reports will be detected.

The proposed scheme can protect vehicles' identity privacy. In our proposed scheme, vehicles periodically update their pseudonyms from TA. By changing pseudonyms, vehicles are able to keep themselves anonymous. Moreover, the proposed scheme is also effective to defend against the possible link attack presented in [13], since each vehicle's route (i.e., road segment) is aggregated and encrypted. By this way, attackers cannot infer where vehicles have been based on the given data, and accordingly cannot link their identities. Besides, although SP can obtain the aggregated vehicle information, it is infeasible for it to recover each individual vehicle's route. Therefore, vehicles' identity privacy is preserved in the proposed PPTM scheme.

The proposed scheme can protect vehicles' location privacy. In our proposed scheme, vehicles' location privacy is preserved by aggregating their route reports. Considering the speed is location-aware, attackers may infer vehicles' locations based on the speed information. In this case, our proposed scheme is still effective, since in PPTM each individual speed is also aggregated and encrypted. Similarly, since all speed reports are also aggregated in the RSU, SP cannot obtain each individual vehicle's speed information. Thus, vehicles' location privacy is preserved.

The proposed scheme can resist collusion attacks. The basic idea to mitigate collusion attacks is to ensure the separation of data between different entities. In PPTM, with the assumption that RSU does not collude with SP, neither of them can know each individual vehicle's privacy. More specifically, the RSU cannot know vehicles' reports since they are encrypted by using the SP's public key. The SP can decrypt the summation of vehicles and speed in each segment, while not knowing each individual vehicle's data.

6. Performance Evaluation

In this section, we evaluate the performance of the proposed PPTM scheme in terms of computational costs of vehicles and RSU, and communication overhead of vehicle-to-RSU and RSU-to-SP communications.

6.1. Computational Costs

For the proposed PPTM scheme, when a vehicle $\mathcal{V}_j$ generates an encrypted report $PID_j||Y_j||C_{j1}||C_{j2}||TS||\sigma_j$, it performs two exponentiation operations in $\mathbb{Z}_{n^2}$ to calculate C_{j1} and C_{j2}, and one multiplication in $\mathbb{G}$ to build the vehicle's signature σ_j. After collecting vehicles' reports,

the RSU verifies the received reports with $N + 1$ pairing operations. Besides, the RSU also aggregates vehicles' reports to obtain the aggregated vehicle and speed information, which requires $N - 1$ multiplication operations. However, since the multiplication operations in $\mathbb{Z}_{n^2}$ is negligible compared with the time-consuming exponentiation and pairing operations, the time costs can be omitted. In addition, to generate the signature, it also performs one multiplication operation in $\mathbb{G}$. As for the SP, it needs to verify the aggregated data sent from the RSU and obtain the aggregated data, which cost one pairing operation in $\mathbb{G}$ and two exponentiation operations in $\mathbb{Z}_{n^2}$. Here, we use C_n, C_e, C_m to denote the computational cost of an exponentiation operation in $\mathbb{Z}_{n^2}$, a pairing operation in $\mathbb{G}$, and a multiplication operation in $\mathbb{G}$, respectively. Then, the total computation costs for the vehicle, RSU, and SP will be $2 * C_n + C_m$, $(N + 1) * C_e + C_m$, and $C_e + 2 * C_n$, respectively.

Our proposed PPTM scheme enables each vehicle to embed its multiple speed into one compressed data, and thus large computational costs can be saved. To compare the efficiency of PPTM, a traditional approach denoted by TRPM is considered, which encrypts every individual speed information at the corresponding road segment. Under the same setting, a vehicle has to generate M ciphertexts, consuming M exponentiation operations in $\mathbb{Z}_{n^2}$ to perform the encryption. In addition, for the ciphertexts, the vehicle is required to generate one signature, which needs one multiplication operation in $\mathbb{G}$. Thus, the total time costs will be $M * C_n + C_m$. For the RSU, it performs batch verification to authenticate the reports, which takes $N + 1$ pairing operations. However, since the number of ciphertexts in TRPM is much more than that in PPTM, i.e., ($M * N$ *vs.* $M * 2$), the RSU has to perform more multiplication operations for speed aggregation. Then, the RSU generates a signature and forwards it to the RSU, which will execute M exponentiation operations to recover the aggregated speed in all road segments. Thus, the total computational costs of an individual vehicle, the RSU, and the SP will be $M * C_n + C_m$, $(N + 1) * C_e + C_m$, and $C_e + M * C_n$, respectively.

We list the computational costs of PPTM and TRPM in Table 1. In addition, we conducted extensive experiments to compare the efficiency of our proposed PPTM scheme. We used nodes with 1.5 GHz and 2 GB RAM as resource-constrained nodes (i.e., vehicles), and used a laptop with Intel Core i7-7600U CPU and 16 GB RAM as the entities that hold enhanced computational capacities (i.e., the RSU and cloud). The security number of κ and κ_1 were set as 1024 bits and 160 bits. All experiments were executed 10 times, and the average results were selected. The experimental results indicate that each single multiplication operation in $\mathbb{G}$ took 2 ms, each exponentiation operation in $\mathbb{Z}_{n^2}$ took 5 ms, and each pairing operation in $\mathbb{G}$ took 2 ms. To validate the efficiency of our proposed PPTM, we show the computational costs in terms of the number of vehicles and road segments in Figures 5–8.

In Figure 5, we can see that, as the number of road segments increased, the time TRPM took at the vehicle side increased linearly, while it was relatively stable in PPTM. The reason is that PPTM uses a well defined super-increasing sequence to aggregate the speeds before performing data encryption. We then plot the time cost of the fog node in Figure 6. Since the reports were verified through a batch way, the scheme was efficient at the RSU side. For example, when the number of vehicles reached 500, PPTM only needed 0.994s to finish reports authentication and ciphertexts aggregation. To validate the efficiency of the proposed PPTM, we further conducted a comparative experiment by changing the number of vehicles and road segments. As shown in Figure 7, it was obvious that our proposed PPTM scheme performed much better than the traditional TRPM scheme, which demonstrated the correctness of the complexity analysis in Table 1. The result shown in Figure 8 also demonstrates the efficiency of our proposed scheme in terms of the average speed calculation at the SP side.

Table 1. Comparison of computational complexity.

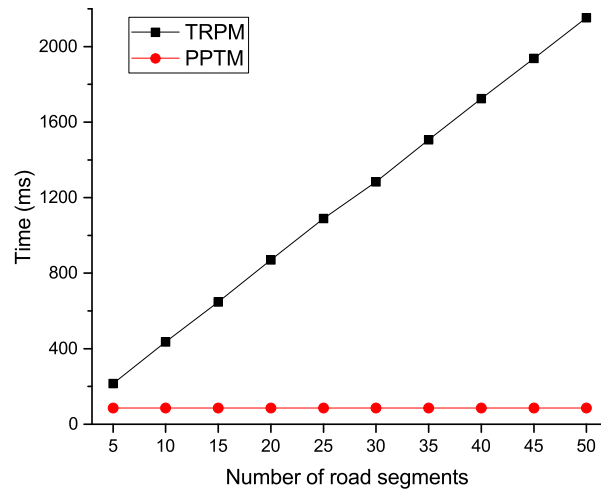|	PPTM	TRPM
Vehicle	$2 * C_n + C_m$	$M * C_n + C_m$
RSU	$(N + 1) * C_e + C_m$	$(N + 1) * C_e + C_m$
SP	$C_e + 2 * C_n$	$C_e + M * C_n$

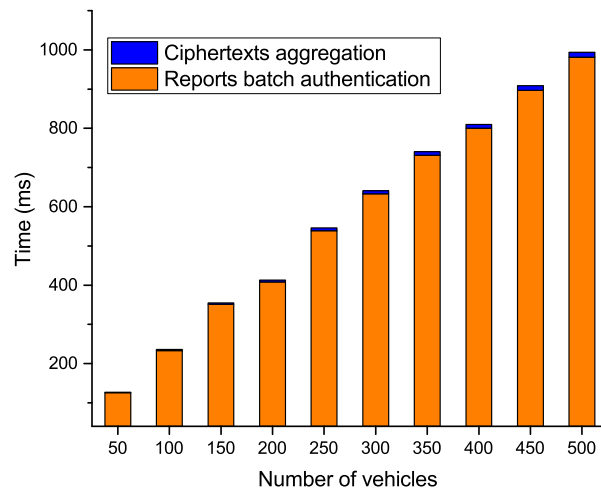**Figure 5.** Computational costs of ciphertexts generation at the vehicle side.

Figure 6. Computational costs of ciphertexts generation and reports batch verification at the RSU side.
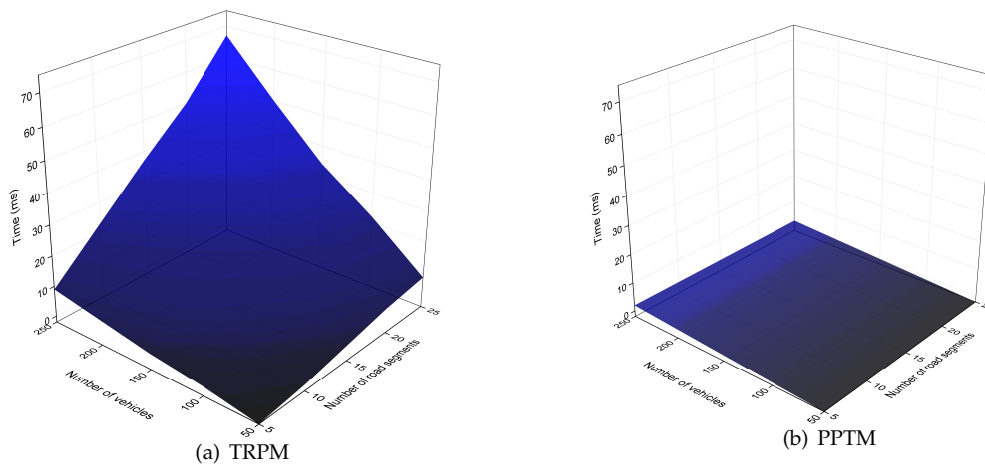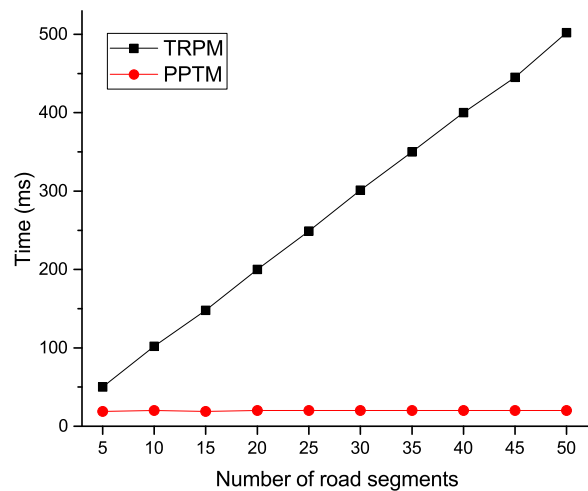

(a) TRPM

(b) PPTM

Figure 7. Computational costs with varying number of vehicles and road segments at the RSU side.

Figure 8. Computational costs of average speed calculation at the SP side.

6.2. Communication Overhead

We then analyzed the communication overhead of the proposed scheme. Generally, the communications of PPTM includes two parts, i.e., vehicle-to-RSU communication and RSU-to-SP communication. For the vehicle-to-RSU communication, each individual vehicle generates its traffic report and transmits it to the RSU. Recalling our previous description, the vehicle's report is defined as $PID_j||Y_j||C_{j1}||C_{j2}||TS||\sigma_j$ and the size is $S_v = |PID_j| + 160 + 2048 * 2 + |TS| + 160$, where the size of n and $\mathbb{G}$ were set as 1024 bits and 160 bits, respectively. RSU is responsible for collecting N reports in its coverage region, thus the total communication cost for the RSU is $S_R = N * S_v$. In the traditional TRPM scheme, each vehicle needs to generate a ciphertext with 2048-bits for every road segment. Then, the total communication cost of vehicle-to-RSU will be $S_v = |PID_j| + 160 + 2048 * M + |TS| + 160$. We then considered the RSU-to-SP communication. In PPTM, RSU transmits the aggregated report $ID_r||C_1||C_2||TS||\sigma_r$ to the SP, which costs $S_S = |ID_r| + 2048 * 2 + |TS| + 160$ bits. Alternatively, TRPM needs to forward each segment's aggregated report to the SP, which requires $|ID_r| + 2048 * M + |TS| + 160$ bits. Based on the analysis presented above, it was obvious that our proposed PPTM could significantly reduce the bandwidth costs compared with the traditional TRPM.

7. Related Works

Recently, traffic monitoring has received considerable attention as it is important to reduce fuel waste, air pollution, and improve drivers' driving experience. By collecting vehicles' traffic information, the traffic conditions can be better identified. Based on this, many schemes and applications have been proposed. However, the security and privacy of vehicles are still major concerns [22,23]. In fact, if drivers' privacy is not being strictly protected, they are usually reluctant to submit their data.

To realize privacy-preserving traffic monitoring, some schemes [12,24–26] have been proposed. For example, Chim et al. [24] presented a secure navigation scheme that uses RSUs to guide vehicles in a distributed way. However, since all vehicles can obtain a same master key, their scheme cannot defend against the insider attacks. By using vehicular cloud and zero-knowledge proof, Sur et al. [25] designed a secure navigation approach. Nevertheless, the credentials cannot be reused, which introduces more computational costs. Ni et al. [12] and Wang et al. [27] realized real-time navigation by collecting vehicles' speed information. With the technology of randomizable signature, their schemes can achieve conditional privacy preservation. Rabieh et al. [26] further proposed a privacy-preserving route

reporting scheme. In their scheme, vehicles' future routes are collected, which would be used to calculate the number of vehicles appearing in next routes.

Although many efforts have been made to realize privacy-preserving traffic monitoring, most of them, nevertheless, ignore the time link attack, as described in Figure 1. Since vehicles are required to report their driving reports periodically or at different road segments, by linking their arriving time, vehicles' trajectories can be easily identified. That is, the traditional technologies to protect drivers' identity privacy, such as pseudonyms or randomizable signature, are not suitable in certain VANET-based applications. Besides, to preserve data privacy, cryptographic primitives such as elliptic curve cryptography [28] and key and key aggregation are proposed [29], which may introduce extensive workloads on the vehicles. Inspired by the work in [30], we applied the use of super-increasing sequence to aggregate vehicles' routes and speed information. In this way, vehicles' identity and location privacy are preserved.

8. Conclusions

Vehicles' speed information is important to monitor the traffic conditions and prevent road congestion, however it threatens drivers' privacy. In this paper, we propose a privacy-preserving traffic monitoring scheme by collecting vehicles' speed and route information. The main idea is to aggregate multiple speeds into one compressed datum so that vehicles' identity and location privacy will not be disclosed. Security analysis indicates that the proposed PPTM scheme is secure and privacy-preserving. Besides, extensive simulations demonstrated its efficiency. In the future, we will try to achieve privacy-preserving traffic monitoring without the assistance of the RSU.

References

1. Voelcker, J. 1.2 Billion Vehicles On World's Roads Now, 2 Billion By 2035: Report. Available online: https://www.greencarreports.com/news/1093560_1-2-billion-vehicles-on-worlds-roads-now-2-billion-by-2035-report/ (accessed on 29 July 2014).
2. Moskvitch, K. Can a City Every Be Traffic Jam-Free? Available online: http://www.bbc.com/future/story/20140611-can-we-ever-end-traffic-jams/ (accessed on 11 June 2014).
3. Wu, C.; Ji, Y.; Liu, F.; Ohzahata, S.; Kato, T. Toward Practical and Intelligent Routing in Vehicular Ad Hoc Networks. *IEEE Trans. Veh. Technol.* **2015**, *64*, 5503–5519. [CrossRef]
4. Abboud, K.; Zhuang, W. Stochastic Modeling of Single-Hop Cluster Stability in Vehicular Ad Hoc Networks. *IEEE Trans. Veh. Technol.* **2016**, *65*, 226–240. [CrossRef]
5. Zhu, L.; Zhang, C.; Xu, C.; Du, X.; Xu, R.; Sharif, K.; Guizani, M. PRIF: A Privacy-Preserving Interest-Based Forwarding Scheme for Social Internet of Vehicles. *IEEE Internet Things J.* **2018**, *5*, 2457–2466. [CrossRef]
6. Wu, D.J.; Zimmerman, J.; Planul, J.; Mitchell, J.C. Privacy-Preserving Shortest Path Computation. In Proceedings of the 23rd Annual Network and Distributed System Security Symposium, San Diego, CA, USA, 21–24 February 2016.
7. Waze. Available online: https://www.waze.com/ (accessed on 1 January 2016).
8. Xu, J.; Xue, K.; Yang, Q.; Hong, P. PSAP: Pseudonym-Based Secure Authentication Protocol for NFC Applications. *IEEE Trans. Consum. Electron.* **2018**, *64*, 83–91. doi: [CrossRef]
9. Yang, Q.; Xue, K.; Xu, J.; Wang, J.; Li, F.; Yu, N. AnFRA: Anonymous and Fast Roaming Authentication for Space Information Network. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 486–497. [CrossRef]

10. Lwamo, N.M.R.; Zhu, L.; Xu, C.; Sharif, K.; Liu, X.; Zhang, C. SUAA: A Secure User Authentication Scheme with Anonymity for the Single & Multi-server Environments. *Inf. Sci.* **2019**, *477*, 369–385. doi: [CrossRef]

11. Du, X.; Chen, H. Security in wireless sensor networks. *IEEE Wirel. Commun.* **2008**, *15*, 60–66.

12. Ni, J.; Lin, X.; Zhang, K.; Shen, X. Privacy-Preserving Real-Time Navigation System Using Vehicular Crowdsourcing. In Proceedings of the IEEE 84th Vehicular Technology Conference, VTC Fall 2016, Montreal, QC, Canada, 18–21 September 2016; pp. 1–5.

13. Golle, P.; Partridge, K. On the Anonymity of Home/Work Location Pairs. In Proceedings of the 7th International Conference on Pervasive Computing, Nara, Japan, 11–14 May 2009; pp. 390–397.

14. Zhang, C.; Zhu, L.; Xu, C.; Sharif, K.; Du, X.; Guizani, M. LPTD: Achieving lightweight and privacy-preserving truth discovery in CIoT. *Future Gener. Comp. Syst.* **2019**, *90*, 175–184. [CrossRef]

15. Xue, K.; Hong, J.; Ma, Y.; Wei, D.S.L.; Hong, P.; Yu, N. Fog-Aided Verifiable Privacy Preserving Access Control for Latency-Sensitive Data Sharing in Vehicular Cloud Computing. *IEEE Netw.* **2018**, *32*, 7–13. [CrossRef]

16. Boneh, D.; Franklin, M.K. Identity-Based Encryption from the Weil Pairing. *SIAM J. Comput.* **2003**, *32*, 586–615. [CrossRef]

17. Paillier, P. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, 2–6 May 1999; pp. 223–238.

18. Abdalla, M.; Bellare, M.; Rogaway, P. The Oracle Diffie-Hellman Assumptions and an Analysis of DHIES. In Proceedings of the RSA Conference on Topics in Cryptology, San Francisco, CA, USA, 8–12 April 2001; pp. 143–158.

19. Xu, C.; Lu, R.; Wang, H.; Zhu, L.; Huang, C. TJET: Ternary Join-Exit-Tree Based Dynamic Key Management for Vehicle Platooning. *IEEE Access* **2017**, *5*, 26973–26989. [CrossRef]

20. Boneh, D.; Lynn, B.; Shacham, H. Short Signatures from the Weil Pairing. *J. Cryptol.* **2004**, *17*, 297–319. [CrossRef]

21. Bellare, M.; Rogaway, P. Random Oracles are Practical: A Paradigm for Designing Efficient Protocols. In Proceedings of the 1st ACM Conference on Computer and Communications Security, Fairfax, VA, USA, 3–5 November 1993; pp. 62–73.

22. Du, X.; Xiao, Y.; Guizani, M.; Chen, H. An effective key management scheme for heterogeneous sensor networks. *Ad Hoc Netw.* **2007**, *5*, 24–34. [CrossRef]

23. Xiao, Y.; Du, X.; Zhang, J.; Hu, F.; Guizani, S. Internet Protocol Television (IPTV): The Killer Application for the Next-Generation Internet. *IEEE Commun. Mag.* **2007**, *45*, 126–134. [CrossRef]

24. Chim, T.W.; Yiu, S.; Hui, L.C.K.; Li, V.O.K. VSPN: VANET-Based Secure and Privacy-Preserving Navigation. *IEEE Trans. Comput.* **2014**, *63*, 510–524. [CrossRef]

25. Sur, C.; Park, Y.; Rhee, K.H. An efficient and secure navigation protocol based on vehicular cloud. *Int. J. Comput. Math.* **2016**, *93*, 325–344. [CrossRef]

26. Rabieh, K.; Mahmoud, M.M.E.A.; Younis, M.F. Privacy-Preserving Route Reporting Schemes for Traffic Management Systems. *IEEE Trans. Veh. Technol.* **2017**, *66*, 2703–2713. [CrossRef]

27. Wang, L.; Liu, G.; Sun, L. A Secure and Privacy-Preserving Navigation Scheme Using Spatial Crowdsourcing in Fog-Based VANETs. *Sensors* **2017**, *17*, 668. doi:10.3390/s17040668. [CrossRef]

28. Du, X.; Guizani, M.; Xiao, Y.; Chen, H. Transactions papers a routing-driven Elliptic Curve Cryptography based key management scheme for Heterogeneous Sensor Networks. *IEEE Trans. Wirel. Commun.* **2009**, *8*, 1223–1229. [CrossRef]

29. Xiao, Y.; Rayi, V.K.; Sun, B.; Du, X.; Hu, F.; Galloway, M. A survey of key management schemes in wireless sensor networks. *Comput. Commun.* **2007**, *30*, 2314–2341. [CrossRef]

30. Lu, R.; Liang, X.; Li, X.; Lin, X.; Shen, X. EPPA: An Efficient and Privacy-Preserving Aggregation Scheme for Secure Smart Grid Communications. *IEEE Trans. Parallel Distrib. Syst.* **2012**, *23*, 1621–1631.