


Article

Achieve Location Privacy-Preserving Range Query in Vehicular Sensing

Qinglei Kong ¹ , Rongxing Lu ^{2,*}, Maode Ma ¹ and Haiyong Bao ³

¹ School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore; qlkong@ntu.edu.sg (Q.K.); emdma@ntu.edu.sg (M.M.)

² Faculty of Computer Science, University of New Brunswick, Fredericton, NB E3B 5A3, Canada

³ School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China; baohy@zjgsu.edu.cn

* Correspondence: rlu1@unb.ca; Tel.: +1-506-451-6966

Received: 12 June 2017; Accepted: 2 August 2017; Published: 8 August 2017

Abstract: Modern vehicles are equipped with a plethora of on-board sensors and large on-board storage, which enables them to gather and store various local-relevant data. However, the wide application of vehicular sensing has its own challenges, among which location-privacy preservation and data query accuracy are two critical problems. In this paper, we propose a novel range query scheme, which helps the data requester to accurately retrieve the sensed data from the distributive on-board storage in vehicular ad hoc networks (VANETs) with location privacy preservation. The proposed scheme exploits structured scalars to denote the locations of data requesters and vehicles, and achieves the privacy-preserving location matching with the homomorphic Paillier cryptosystem technique. Detailed security analysis shows that the proposed range query scheme can successfully preserve the location privacy of the involved data requesters and vehicles, and protect the confidentiality of the sensed data. In addition, performance evaluations are conducted to show the efficiency of the proposed scheme, in terms of computation delay and communication overhead. Specifically, the computation delay and communication overhead are not dependent on the length of the scalar, and they are only proportional to the number of vehicles.

Keywords: range query; location privacy preservation; vehicular sensing

1. Introduction

Nowadays, the fast development of automotive industry and the wide deployment of on-board sensors have created a huge opportunity of vehicular sensing [1]. Other than protecting the normal operations of vehicles, the data generated by on-board sensors (e.g., chemical spill detectors, vibration sensors, and acoustic detectors [2]) can also provide unprecedented spatial-temporal coverage and witness unpredictable incidents with virtually zero investment in the deployment and maintenance of fixed surveillance infrastructures [3]. Meanwhile, since modern vehicles are normally not constrained by energy supply and equipped with adequate on-board storage (up to terabytes), they can keep the continuously harvested data in their on-board storage, which avoids the network congestion caused by the sensed data uploading. Motivated by the practical profits of vehicular sensing, various vehicular sensing based applications have appeared. For example, MobEyes [2] proposes a proactive urban monitoring application with the data collected from cameras and chemical detection sensors, a road surface condition detection application [4] is devised by opportunistically gathering data from vibration and GPS sensors. Meanwhile, to attract sufficient participants to join in the sensing process, an incentive mechanism for mobile crowdsensing has been devised in [5].

However, the wide application of vehicular sensing has met several challenges [6–8]. The first challenge is related to the location privacy of data requesters and data uploading vehicles. To retrieve

the wanted data from the distributive on-board storage, each data query should specify the corresponding time and location. However, if the location information of a data query is disclosed, it may bring social reputation or economic damage to the querying location. Meanwhile, since vehicles are dynamically moving, their data reports should also be spatial-temporal tagged. The location information of a vehicle can be correlated with certain personal affairs (such as churches and hospitals), or it may reveal the identities of the vehicle owners (such as residences and offices). Without location privacy preservation, data requesters and vehicles are reluctant to issue data queries and upload sensed data, which leads to the under utilization of vehicular sensing. Thus, the location privacy of data requesters and vehicles should be preserved.

The second challenge is related to the accuracy of data query results. Due to the sheer volume of on-board data generation and limited transmission bandwidth, it is impossible for vehicles to upload all the sensed data immediately, and some of the sensed data are maintained in their on-board storage (except those real-time data for applications with stringent delay requirements) [9]. Since vehicles are dynamically and opportunistically moving, the sensory data maintained in their on-board storage captured during a past time period may or may not be generated within the target query area. Moreover, it is highly possible that the queried sensory data are partially generated within the target query area. Thus, it is difficult for the data requester to accurately identify and acquire the wanted data from the on-board storage of the massive and dynamically moving vehicles. However, most of the current secure range query schemes are devised for the outsourced central cloud storage [10,11], which cannot be directly applied to the distributive vehicular storage. Thus, a secure and accurate range query scheme is needed for the distributive on-board storage scenario to fully exploit the potential of vehicular sensing.

In this paper, to overcome the above challenges, we propose a novel privacy-preserving range query scheme from the distributive on-board storage in vehicular ad hoc networks (VANETs) in a practical scenario: acquiring the data harvested by the on-board air pollution sensors within the defined query area, i.e., an industrial area, to monitor the air quality. The proposed scheme exploits the Paillier cryptosystem [12] to preserve the location privacy of data requesters and vehicles, and protect the confidentiality of the sensed data. Specifically, the contributions of this paper are threefold.

First, the proposed scheme structures each multi-dimension scalar in one dimension, where the multi-dimension scalars denote the positions of the data requesters and vehicles in the format of grid cells; meanwhile, the proposed scheme supports secure scalar product computation for location matching. Thus, the location-based data query and data reports can be transmitted to the data server with high efficiency; meanwhile, the data harvested within the target query area can still be identified with location privacy preservation.

Second, since the vehicles are dynamically moving, the average of the sensory data captured during a short-time period may contain both the sensory data generated within and outside the query area. The proposed scheme enables the identification of the number of data reports generated within the target query area, and guarantees the accuracy of the sensory data captured within the target area.

Third, we give detailed security analysis to show that the proposed scheme is secure under the defined security model and achieves the security requirements in terms of location privacy preservation and confidentiality. Meanwhile, we conduct comparative performance evaluation to show that the proposed scheme is more efficient than the existing homomorphic secure scalar product schemes in terms of computational complexity and communication overhead.

The remainder of this paper is organized as follows. We describe the system model, the security requirements, and the design goals in Section 2. We recall the preliminaries and propose our privacy-preserving range query scheme in Section 3, followed by our security analysis and performance evaluations in Sections 4 and 5, respectively. We show related work in Section 6, and finally conclude our work in Section 7.

2. System Model, Security Requirement and Design Goal

2.1. System Model

In the system model, we present the proposed range query scheme in a practical vehicular sensing application: help the data requester to acquire the data generated by vehicles (data harvested by the on-board air pollution sensors [13]) located within the target query area (an industrial district) during a past short-time period, which can help to investigate the air quality of the given industrial district. The system model consists of five entities, as shown in Figure 1.

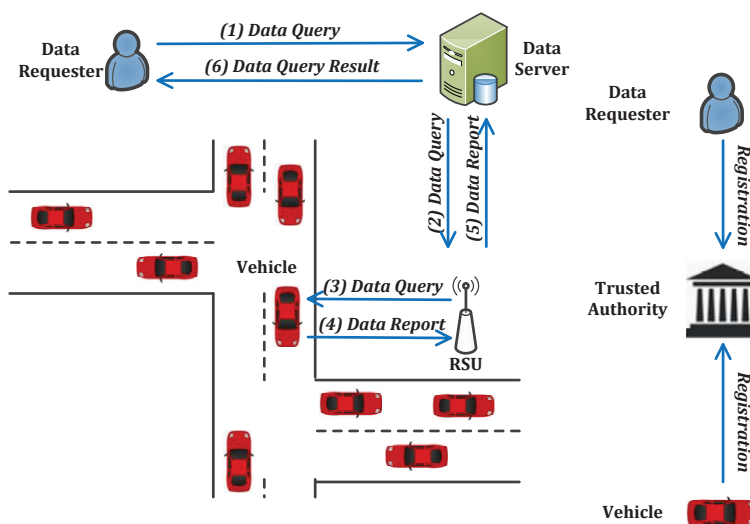


Figure 1. Data query architecture.

- The **data requester** aims to gather the air quality data in a given industrial district from the on-board air pollution sensors. The data requester submits a location-based data query ((1) *Data Query*) towards the data server and waits to hear the reply, as shown in Figure 1.
- When the **data server** receives a data query request from a data requester, it forwards the received data query request ((2) *Data Query*) to all the vehicles through road side units (RSUs). Meanwhile, the data server receives all the on-board sensed data reports from RSUs, performs data filtering according to the data query, and delivers the data query response towards the data requester ((6) *Data Query Result*), as shown in Figure 1.
- Each **RSU** serves as a gateway between the data server and data uploading vehicles, it helps to broadcast received data queries towards all the vehicles under its coverage ((3) *Data Query*) and forward the received on-board sensed data reports towards the data server ((5) *Data Report*), as shown in Figure 1.
- Each **data uploading vehicle** is equipped with the required on-board sensor and enough on-board storage. When a piece of data is harvested by the on-board sensor of a vehicle, it is maintained in the on-board storage with time and location tagged. When a vehicle receives a data query, it first checks whether the queried data is still maintained. If the vehicle owns the queried data, it uploads the location-based sensed data to the nearest RSU((4) *Data Report*), as shown in Figure 1.
- The **trusted authority** is a trusted and powerful entity, which is mainly responsible for the system bootstrap, key materials management, and the registration of new data requesters and vehicles, as shown in Figure 1.

The wireless connection between a vehicle and an RSU is realized through the IEEE 802.11p standard, a short- to medium-range communication technology operating at 5.85–5.925 GHz band with 3–27 Mbps data rates [14], which is mainly designed for the intelligent transportation systems

radio service. The connections between RSUs and the data server, and those between data requesters and the data server, are realized through either wired links or any other links with high bandwidth and low transmission delay.

2.2. Security Requirements

In the security model, we consider the trusted authority is fully trusted, while the data server and RSUs are assumed to be honest-but-curious, that is, they follow the protocols, but they may try to infer the generation location and the content of data queries and data reports, which may violate the location privacy and confidentiality. Therefore, to preserve the location privacy of data requesters and vehicles, and protect the confidentiality of each individual sensed data report, the following security requirements should be satisfied:

Location Privacy. Protect the location privacy indicates that the location information contained in each data query and data report should be protected [6], and the location information in this context indicates the grid cell scalars, which the data query and data generation locations are mapped to. Even if the data server obtains all the possible data queries and data reports, it cannot identify the grid cell scalar contained in any location-based data query or data report. Location privacy protection also includes that the data uploading vehicles cannot learn the grid cell scalar of the data query, or vice versa. In this way, the location privacy of data requesters and data uploading vehicles can be preserved.

Confidentiality. Protect the confidentiality of individual sensed data report means that, even if the data server obtains all the possible on-board sensed data reports, it cannot recover the content of an individual sensed data report [15]. Thus, the individual on-board sensed data can achieve the security requirement of confidentiality.

Note that there may exist other types of attacks such as impersonation, eavesdropping, and violation of data integrity [16]. Since we mainly focus on the privacy-preserving range query from the distributive on-board storage in VANETs, these security threats are beyond our study scope. We also assume that there is no collusion attack in the system, which is in accordance with previous research on secure data query [17].

2.3. Design Goals

Under the aforementioned system model and security requirements, our design goal is to develop a privacy-preserving range query scheme from the distributive on-board storage in VANETs. Specifically, the following three objectives should be achieved.

According to the above statement, if the proposed range query scheme does not take security into consideration, the location privacy of data requesters and data uploading vehicles could be threatened. Then, data requesters and vehicles may not be willing to participate in the range query process, and the system cannot properly operate. Therefore, the proposed scheme should achieve the security goal of location privacy.

To accurately reflect the air quality, it is important for data requesters to accurately obtain the data harvested by vehicles in the given query area. Since vehicles are dynamically moving and their positions are opportunistically changing, each received sensed data report may or may not be able to reflect air quality. Thus, to achieve the goal of the high accuracy in air quality monitoring, the data server should accurately filter all the received data reports.

Since vehicles are featured with the fast-moving characteristic, the connections between RSUs and vehicles are relatively short and intermittent, and the communication overhead introduced by the data query and sensed data reports should also be minimized. To extract the desired data from the massive amount of sensory data reports efficiently, the computational complexities brought to data server should also be deliberately evaluated.

3. Proposed Scheme

In this section, we propose the privacy-preserving range query scheme from the distributive on-board storage in VANETs, which mainly consists of five parts: preliminaries, system initialization, data query generation, data report generation, and data filtering.

3.1. Preliminaries

In our proposed scheme, the Paillier cryptosystem is exploited due to its additive homomorphic property and the homomorphic multiplication property of one plaintext and one ciphertext [12], which has been widely employed in many privacy-preserving data processing applications [15]. Specifically, the Paillier cryptosystem consists of three components: key generation, encryption, and decryption.

- **Key Generation:** Given the security parameter κ , two large prime numbers p_1, q_1 are first chosen, where $|p_1| = |q_1| = \kappa$. Then, the RSA modulus $n = p_1 q_1$ and $\lambda = \text{lcm}(p_1 - 1, q_1 - 1)$ are computed. Define a function $L(u) = \frac{u-1}{n}$, after choosing a generator $g \in \mathbb{Z}_{n^2}^*$, $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n$ is further calculated. Then, the public key is $pk = (n, g)$, and the corresponding private key is $sk = (\lambda, \mu)$.
- **Encryption:** Given a message $m \in \mathbb{Z}_n$, choose a random number $r \in \mathbb{Z}_n^*$, and the ciphertext can be calculated as $c = E(m) = g^m \cdot r^n \bmod n^2$.
- **Decryption:** Given the ciphertext $c \in \mathbb{Z}_{n^2}^*$, the corresponding message can be recovered as $m = D(c) = L(c^\lambda \bmod n^2) \cdot \mu \bmod n$.

3.2. System Initialization

For the range query system under consideration, we assume a trusted authority, located at the management authority of vehicles and traffics, will bootstrap the whole system.

Given a security parameter κ , the trusted authority selects two large prime numbers p_1 and q_1 , where $|p_1| = |q_1| = \kappa$. The trusted authority also calculates the Paillier cryptosystem's public key $(n = p_1 q_1, g)$, and the corresponding private key (λ, μ) . Moreover, the trusted authority chooses a secure cryptographic hash functions $H()$, where $H : \{0, 1\}^* \rightarrow \mathbb{Z}_n^*$. In addition, the trusted authority also chooses a random number α . The trusted authority publishes the system parameter as $params = \{n, g, H()\}$.

The trusted authority assigns the private key (λ, μ) towards the data server, but the trusted authority does not share α with the data server. During the registration of each data uploading vehicle (w vehicles in total) or each data requester, the trusted authority checks its eligibility (whether the vehicle has the on-board air pollution sensors installed) and securely returns α towards it, as shown in Figure 2.

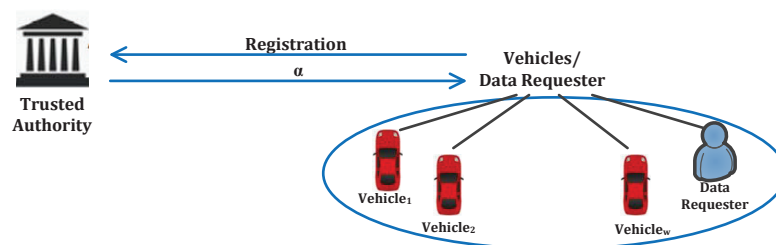


Figure 2. Registration and secret sharing among the data requester and vehicles.

3.3. On-Board Data Query Generation

In this subsection, we first describe the query area construction. Then, we identify the data query generation process.

3.3.1. Query Area Construction

A data requester, the environmental protection agency in this context, aims to monitor the air quality of an industrial district through collecting the data generated by the air pollution sensor (*Data_Type*), during a past time period (from the starting time t_s to the ending time t_e). We exploit the grid cell system defined in [18] to construct a grid cell architecture, which reflects the locations of the data requester and vehicles: the data requester also specifies a large region contains the target industrial district, which is large enough for the data requester and vehicles to be comfortable with revealing the fact that they are somewhere within this query area, and the bottom-left vertex of the large region is with the bottom-left vertex (x_0, y_0) . The large region is constructed into a grid cell structure, which is divided into $k = a \times b$ equal-sized grid cells with the side length of L , i.e., $L = 0.8$ km, when we take an industrial area with the coverage of 2.4 km^2 as an example [19]. Given location L_i with coordinates (x_i, y_i) , the corresponding grid cell identifier can be computed as $\lceil \frac{x_i - x_0}{L} \rceil + \lfloor \frac{y_i - y_0}{L} \rfloor \cdot a$. As shown in Figure 3, the data query area and the moving trajectory of vehicles are denoted with grid cells.

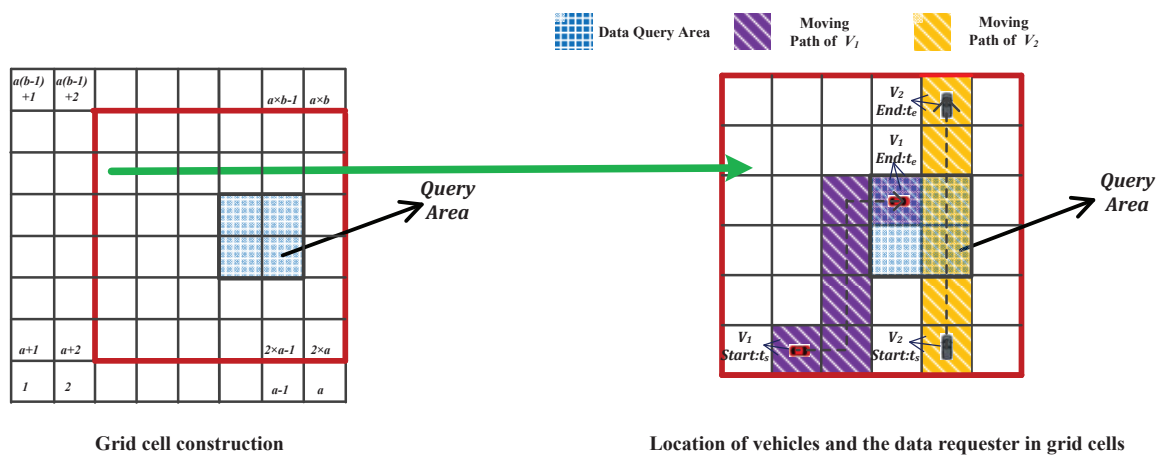


Figure 3. Defined grid cell architecture/locations of vehicles and the data query area in grid cells.

3.3.2. Data Query Generation

The data requester generates a query area that includes a group of grid cells \mathcal{I}_d , which is overlapped with the target industrial district. Meanwhile, the data requester generates a grid cell scalar $\vec{u} = \{u_1, \dots, u_k\}$ to denote the defined grid cell architecture shown in Figure 3, which is initially set to zero. If $i \in \mathcal{I}_d$, i.e., grid cell i belongs to the target industrial district, we set the corresponding position in the scalar \vec{u} to $u_i = 1$. Moreover, the data requester selects a number β , which satisfies the condition that $2 \cdot k < 2^\beta$ and $2^{(2k+1)\beta} < n$. In addition, we structure the multi-dimension grid cell scalar to one composite value by calculating $m_u = \sum_{i=1}^k u_i \cdot 2^{i\beta}$, and generate the ciphertext of m_u , which is

$$\begin{cases} C_{u,1} = g^{m_u + H(\alpha||1||TS)} \cdot r_{u,1}^n \bmod n^2, \\ C_{u,2} = g^{m_u \cdot H(\alpha||1||TS)^{-1} + H(\alpha||2||TS)} \cdot r_{u,2}^n \bmod n^2, \end{cases} \quad (1)$$

where TS is the current timestamp, and $r_{u,1}, r_{u,2} \in \mathbb{Z}_n^*$ are two random numbers.

Finally, the data requester formulates a data request and delivers it to the data server:

$$Request \leftarrow \langle Data_Type || x_0 || y_0 || a || b || L || \beta || t_s || t_e || \Delta t || C_{u,1} || C_{u,2} || TS \rangle, \quad (2)$$

where Δt is the on-board sensed data sampling interval.

After receiving *Request*, the data server further delivers the data request

$$Request_1 \leftarrow \langle Data_Type || x_0 || y_0 || a || b || L || \beta || t_s || t_e || \Delta t || TS \rangle \quad (3)$$

towards the data uploading vehicles through RSUs.

3.4. Data Report Generation

For each registered data uploading vehicle, the previously sensed data (on the scale of a few past days) and the related information (such as number, time, data type, sensed data, location coordinates, etc.) are maintained in its on-board storage in the format of $\langle n, t, Data_Type, data, (x, y) \rangle$. Upon receiving $Request_1$, it checks its on-board storage to identify the data in consistent with the data query. If the queried data exists, the vehicle performs the following steps.

The vehicle identifies the on-board air pollution sensor data ($data_1, data_2, \dots, data_d$) harvested at the past time points $(t_s, t_s + \Delta t, \dots, t_e)$, where $d = \frac{t_e - t_s}{\Delta t} + 1$, $d < k$, Δt is the sampling interval. Then the vehicle calculates the sum of the sensed data reports, which is denoted as $sum_v = \sum_{i=1}^d data_i$ and it satisfies the condition that $\sum_{v=1}^w sum_v < n$, where w denotes the total number of data uploading vehicles. In addition, the vehicle generates the ciphertext of the on-board sensed data sum_v :

$$C_{v,0} = g^{sum_v + H(\alpha||0||TS)} \cdot r_{v,0}^n \bmod n^2, \quad (4)$$

where $r_{v,0} \in \mathbb{Z}_n^*$ is a random number.

Then, the vehicle maps the location coordinates of the on-board sensed data into grid cell architecture shown in Figure 3, and generates a vector $\vec{v} = \{v_1, \dots, v_k\}$ with Algorithm 1.

Algorithm 1 Vehicle Scalar Generation

Data: Given location coordinates $\{(x_1, y_1), (x_2, y_2), \dots, (x_d, y_d)\}$ and sets $\vec{v} = \{v_1, \dots, v_k\} = \vec{0}$.

Compute:

```

1: for  $i = 1 : d$  do
2:   Given location coordinates  $L_i(x_i, y_i)$ , the vehicle calculates  $j = \lceil \frac{x_i - x_0}{L} \rceil + \lfloor \frac{y_i - y_0}{L} \rfloor \cdot a$ ;
3:   if  $j \in \{1, 2, \dots, k\}$  then
4:     Calculates  $v_j = v_j + 1$ ;
5:   end if
6: end for
```

Output: \vec{v}

Then, the vehicle structures the multi-dimensional scalar \vec{v} into one dimension: $m_v = \sum_{j=1}^k v_j \cdot 2^{(k+1-j) \cdot \beta}$, and generates the ciphertext

$$\begin{cases} C_{v,1} = g^{m_v + H(\alpha||1||TS)^{-1}} \cdot r_{v,1}^n \bmod n^2, \\ C_{v,2} = g^{m_v \cdot H(\alpha||1||TS) - H(\alpha||2||TS) - \eta_v} \cdot r_{v,2}^n \bmod n^2, \end{cases} \quad (5)$$

where $r_{v,1}, r_{v,2} \in \mathbb{Z}_n^*$ are two random numbers, and $\eta_v = \eta_{v,1} \cdot 2^\beta + \dots + \eta_{v,k} \cdot 2^{k \cdot \beta} + \eta_{v,k+1} \cdot 2^{(k+2) \cdot \beta}$, such that $\eta_{v,i}$ is a random number that satisfies the condition: $\eta_{v,i} + k < 2^\beta$ when $i \in \{1, \dots, k\}$ and $\eta_{v,k+1} \in [1, k \cdot 2^{(k-2) \cdot \beta}]$.

Finally, the vehicle formulates the *Data_Report* and delivers it to its nearest RSU:

$$Data_Report \leftarrow \langle C_{v,0} || C_{v,1} || C_{v,2} \rangle. \quad (6)$$

3.5. Privacy Preserving Data Filtering

Upon receiving all the data reports, each RSU forwards all the collected on-board sensed data reports to the data server, and the data server filters within all the received data reports to identify the on-board sensed data reports generated within the industrial district.

The data server first computes the scalar product of \vec{u} and \vec{v} to identify the data reports generated within the industrial district, based on the idea that the data uploading vehicle must share at least one

grid cell with the industrial district. Then, the data server decrypts $C_{u,1}$ with its private key sk and recovers the value of $m_u + H(\alpha||1||TS)$. Meanwhile, the data server also computes

$$\begin{aligned} C_{u,v} &= \frac{(C_{v,1})^{m_u + H(\alpha||1||TS)}}{C_{u,2} \cdot C_{v,2}} \bmod n^2 \\ &= \frac{(g^{m_v + H(\alpha||1||TS)^{-1}} \cdot r_{v,1}^n)^{m_u + H(\alpha||1||TS)}}{g^{m_u H(\alpha||1||TS)^{-1} + m_v H(\alpha||1||TS) - \eta_v} \cdot r_{u,2}^n \cdot r_{v,2}^n} \bmod n^2 \\ &= g^{m_u m_v + \eta_v + 1} \cdot \left(\frac{r_{v,1}^{m_u + H(\alpha||1||TS)}}{r_{u,2} \cdot r_{v,2}} \right)^n \bmod n^2. \end{aligned} \quad (7)$$

Based on $C_{u,v}$, the data server recovers the value of $s_{u,v} = m_u m_v + \eta_v + 1$ with the private key sk . With $s_{u,v}$, the data server can calculate the scalar product of $\vec{u} \cdot \vec{v}$,

$$k_v = \vec{u} \cdot \vec{v} = \frac{(s_{u,v} \bmod 2^{(k+2)\beta}) - (s_{u,v} \bmod 2^{(k+1)\beta})}{2^{(k+1)\beta}}, \quad (8)$$

where k_v indicates the pieces of sensed data generated within the industrial district. If $k_v = \vec{u} \cdot \vec{v} \geq 1$, at least one piece of sensed data is generated in industrial district, and the vehicle becomes a member of the group \mathcal{N}_{in} ; otherwise, all the data reports are generated out of the industrial district, and the vehicle becomes a member of the group \mathcal{N}_{out} , such that $|\mathcal{N}_{in}| + |\mathcal{N}_{out}| \leq w$. Finally, the data server calculates $K_{in} = \sum_{v \in \mathcal{N}_{in}} k_v$.

Correctness. The correctness of Equation (8) can be clearly illustrated with the following equation:

$$\begin{aligned} s_{u,v} &= m_u m_v + \eta_v + 1 \\ &= u_1 v_1 \cdot 2^{(k+1)\beta} + u_2 v_1 \cdot 2^{(k+2)\beta} + \dots + u_k v_1 \cdot 2^{2k\beta} \\ &\quad + u_1 v_2 \cdot 2^{k\beta} + u_2 v_2 \cdot 2^{(k+1)\beta} + \dots + u_k v_2 \cdot 2^{(2k-1)\beta} \\ &\quad + \dots \\ &\quad + u_1 v_k \cdot 2^{2\beta} + u_2 v_k \cdot 2^{3\beta} + \dots + u_k v_k \cdot 2^{(k+1)\beta} \\ &\quad + \eta_{v,k+1} 2^{(k+2)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1 \\ &= \left(\sum_{i=1, i < j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1 \right) \\ &\quad + \left(\sum_{i=1, i > j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k+1} 2^{(k+2)\beta} \right) \\ &\quad + \left(\sum_{i=1, i=j}^k u_i v_j \cdot 2^{(k+1)\beta} \right). \end{aligned} \quad (9)$$

As shown in Equation (9), $s_{u,v}$ consists of three parts: (i) when $i < j$, $(\sum_{i=1, i < j}^k u_i v_j \cdot 2^{(k+1+i-j)\beta} + \eta_{v,k} 2^{k\beta} + \dots + \eta_{v,1} 2^\beta + 1) < 2^{(k+1)\beta}$; (ii) when $i > j$, $(\sum_{i=1, i > j}^k u_i v_j \cdot 2^{(i-j-1)\beta} + \eta_{v,k+1}) \cdot 2^{(k+2)\beta}$; and (iii) when $i = j$, $(\sum_{i=1, i=j}^k u_i v_j \cdot 2^{(k+1)\beta})$. Thus, with Equation (8), $\vec{u} \cdot \vec{v}$ can be computed.

3.6. Data Report Aggregation

The data server first aggregates the group of sensed data generated outside the industrial district \mathcal{N}_{out} , which is

$$\begin{aligned}
C_{out} &= \prod_{v \in \mathcal{N}_{out}} C_{v,0} \bmod n^2 \\
&= g^{\sum_{v \in \mathcal{N}_{out}} (sum_v + H(\alpha||0||TS))} \cdot \left(\prod_{v \in \mathcal{N}_{out}} r_{v,0} \right)^n \bmod n^2.
\end{aligned} \tag{10}$$

Then, the data server decrypts C_{out} with the private key sk , and calculates the average of the sensed data generated outside the industrial district, which is $S_{out} = \sum_{v \in \mathcal{N}_{out}} (sum_v + H(\alpha||0||TS))$. Meanwhile, the data server aggregates the sensed data within the group \mathcal{N}_{in} , which is

$$\begin{aligned}
C_{in} &= \prod_{v \in \mathcal{N}_{in}} C_{v,0} \bmod n^2 \\
&= g^{\sum_{v \in \mathcal{N}_{in}} (sum_v + H(\alpha||0||TS))} \cdot \left(\prod_{v \in \mathcal{N}_{in}} r_{v,0} \right)^n \bmod n^2.
\end{aligned} \tag{11}$$

In addition, the data server decrypts $S_{in} = \sum_{v \in \mathcal{N}_{in}} (sum_v + H(\alpha||0||TS))$ with the private key sk ; meanwhile, the data server delivers S_{in} , S_{out} , K_{in} , $|\mathcal{N}_{in}|$ and $|\mathcal{N}_{out}|$ towards the data requester.

After receiving S_{in} , S_{out} , K_{in} , $|\mathcal{N}_{in}|$ and $|\mathcal{N}_{out}|$, the data requester first calculates the average of the sensed data outside the industrial district, which is

$$Ave_{out} = \frac{(S_{out} - |\mathcal{N}_{out}| \cdot H(\alpha||0||TS)) \bmod n}{d \cdot |\mathcal{N}_{out}|}. \tag{12}$$

Then, the data requester calculates the average of the sensed data inside the industrial district, which is

$$Ave_{in} = \frac{(S_{in} - (|\mathcal{N}_{in}| \cdot d - K_{in}) \cdot Ave_{out} - |\mathcal{N}_{in}| \cdot H(\alpha||0||TS)) \bmod n}{K_{in}}. \tag{13}$$

4. Security Analysis

Following the earlier discussed security requirements, we analyze the security properties of the proposed privacy-preserving range query scheme in this section. Specifically, our analysis will focus on how the proposed scheme can achieve the location privacy preservation of data requesters and data uploading vehicles, and protect the confidentiality of the individual on-board sensed data.

The proposed range query scheme can preserve the location privacy of data requester and data uploading vehicles. In the proposed scheme, the data requester first identifies a large region to construct the grid cell structure, which contains the industrial district. Based on the constructed grid cell structure, nothing could be learned besides the fact that the data query occupies one or a few grid cells in the defined grid cell structure. Since the location of the data requester and the vehicles are mapped to the constructed grid cell architecture, protecting the grid cell scalar of the data requester and the vehicles means preserving the location privacy. The following two paragraphs illustrate how the proposed scheme can protect the grid cell scalar during the secure scalar product computation process.

In the proposed scheme, the data requester's multi-dimensional grid cell scalar \vec{u} is structured into one dimension m_u , and then encrypted by the public key of the data server pk to generate $(C_{u,1}, C_{u,2})$. After receiving $(C_{u,1}, C_{u,2})$, the data server decrypts $C_{u,1}$ with its private key sk , and obtains $m_u + H(\alpha||1||TS)$. To prevent the data server from recovering the value of m_u , $H(\alpha||1||TS)$ and $H(\alpha||2||TS)$ are introduced in $(C_{u,1}, C_{u,2})$. Since the secret α is only shared between the data requester and data uploading vehicles and the defined security model does not take collusion attack into consideration, the data server cannot recover the value of m_u . As $(C_{u,1}, C_{u,2})$ are valid ciphertexts of the Paillier cryptosystem, which is known to be semantically secure against the chosen plaintext attack, and the composite grid cell scalar contained in $(C_{u,1}, C_{u,2})$ is also semantically secure. Meanwhile, the locations of data uploading vehicles are also encrypted by the public key of the data server and

protected by the secret α . Since only the data server with the private key sk can decrypt $(C_{u,1}, C_{u,2})$ and $(C_{v,1}, C_{v,2})$, and the data server cannot recover the value of m_u and m_v with the decrypted results, the composite grid cell scalars of the data requester and vehicles can be protected during transmission.

Due to the homomorphic property of the Paillier cryptosystem, the data server first computes $(C_{v,1})^{m_u + H(\alpha||1||TS)}$ and then recovers the value of $s_{u,v} = m_u \cdot m_v + \eta_v + 1$ through decrypting $\frac{(C_{v,1})^{m_u + H(\alpha||1||TS)}}{C_{u,2} \cdot C_{v,2}}$. To avoid the exhaustive attack against $m_u \cdot m_v$, a random number η_v is also included in $s_{u,v}$, but the data server can still obtain the scalar product of $\vec{u} \cdot \vec{v}$ by computing $\frac{(s_{u,v} \bmod 2^{(k+2)\beta}) - (s_{u,v} \bmod 2^{(k+1)\beta})}{2^{(k+1)\beta}}$. In this way, the content of each composite grid cell scalar can still be protected. Meanwhile, the location privacy of the data requester and data uploading vehicles can be preserved during the secure scalar product computation.

The individual on-board sensed data report is confidential in the proposed scheme. In the proposed scheme, each vehicle's on-board sensed data report is formed as $C_{v,0} = g^{sum_v + H(\alpha||0||TS)} \cdot (r_{v,0})^n \bmod n^2$. Since $C_{v,0}$ is a valid ciphertext of the Paillier cryptosystem, the sensed data sum_v in $C_{v,0}$ is also semantically secure. After decrypting $C_{v,0}$ with sk , the data server still cannot recover the value of $sum_v + H(\alpha||0||TS)$, since α is a secret shared between the data requester and vehicles. In addition, the data server aggregates all the sensed data within the group \mathcal{N}_{in} and the group \mathcal{N}_{out} , respectively, and delivers the aggregated results S_{in} and S_{out} to the data requester. Without α , the data server also cannot learn the aggregation of the sensed data reports sum_v , and the confidentiality of the aggregation results can be achieved. Furthermore, the data requester can only obtain the aggregated results contained in S_{in} and S_{out} , but it cannot recover the value of each individual sensed data. Therefore, the confidentiality of individual sensed data can also be protected.

5. Performance Evaluation

In this section, we describe the parameter setup, and evaluate the performance of the proposed scheme in terms of computation complexity and communication overhead.

5.1. Parameter Setup

We conduct the experiments with the Java Paillier Library [20] on a desktop with 3.40 GHz processor and 8.00 GB memory to study the operation costs. The experimental result shows that the cost of a single exponentiation operation in $\mathbb{Z}_{n^2}^*$ ($|n^2| = 2048$) is $C_e = 8.82$ ms. The proposed scheme integrates all the elements in a scalar to one composite value, which requires us to choose the proper parameter β and the length of the scalar k . Given $|n| = 1024$, the maximum value of β_{max} is 7 and the maximum length of the scalar is $k_{max} = 63$, which satisfies the condition that $2 \cdot k_{max} < 2^{\beta_{max}}$ and $(2 \cdot k_{max} + 1) \cdot \beta_{max} < |n|$.

We compare the proposed scheme with the traditional homomorphic secure scalar product when Paillier encryption is exploited, i.e., each element in a scalar is encrypted separately at the data requester side, and then transmitted to the data server, i.e., the ciphertext of the data requester are generated as $C_{u,1,i}, C_{u,2,i}, i = 1, 2, \dots, k$. At the vehicle side, each element in a scalar is also encrypted and delivered individually, i.e., the ciphertext of each vehicle are generated as $C_{v,1,i}, C_{v,2,i}, i = 1, 2, \dots, k$. At the data server side, the product of two corresponding elements in the two scalars are also computed individually, $C_{u,v,i} = \frac{(C_{v,1,i})^{u_i + H(\alpha||1||TS)}}{C_{u,2,i} \cdot C_{v,2,i}}$, and then aggregated to $C_{u,v} = \prod_{i=1}^k C_{u,v,i}$. Finally, the aggregated results are decrypted to derive the final results.

5.2. Computation Complexity

In the proposed scheme, when a data requester generates a data query, it requires 4 exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate $C_{u,1}||C_{u,2}$. Note that the computation of $2^{i \cdot \beta}, i \in \{1, 2, \dots, k\}$ can be conducted in the setup phase, and multiplication operation in $\mathbb{Z}_{n^2}^*$ is negligible in comparison with the exponentiation operation in $\mathbb{Z}_{n^2}^*$. In the traditional scheme, the data requester needs to generate the ciphertext $C_{u,i,1}||C_{u,i,2}$ for each element in the scalar separately. For a scalar with the length of k ,

the data requester needs to consume $4 \times k$ exponentiation operations in $\mathbb{Z}_{n^2}^*$ to generate the ciphertext. In the proposed scheme, when a vehicle receives a data query, it generates an encrypted location-based data report, which requires 6 exponentiation operations in $\mathbb{Z}_{n^2}^*$. However, for the traditional scheme, each vehicle needs to consume $2 + 4 \times k$ exponentiation operations to generate one data report.

After receiving all the data reports from w vehicles, the data server should compute the scalar product of the data requester and each vehicle, to identify whether the on-board data report is harvested within the industrial district. The data server takes one exponentiation operation in $\mathbb{Z}_{n^2}^*$ to recover the value of $m_u + H(\alpha||1||TS)$ by Paillier decryption, and it also consumes two exponentiation operations in $\mathbb{Z}_{n^2}^*$ to obtain the value of $m_u \cdot m_v + 1$. Since the multiplication operation in $\mathbb{Z}_{n^2}^*$ is considered negligible in comparison to the exponentiation operation in $\mathbb{Z}_{n^2}^*$, the computation cost of aggregation is negligible, and it takes two exponentiation operations in $\mathbb{Z}_{n^2}^*$ for the Paillier decryption to recover the aggregated results. For the traditional scheme, it takes $3 \times k$ exponentiation operations in $\mathbb{Z}_{n^2}^*$ to recover the scalar product and costs two exponentiation operations to obtain the aggregated results.

Thus, in the proposed scheme, totally for the data requester, vehicle, and the data server, the computation cost will be $4 \times C_e$, $6 \times C_e$, and $(3 \times w + 2) \times C_e$. For the traditional scheme, the computation cost for the data requester, vehicle, and data server are $4 \times k \times C_e$, $(2 + 4 \times k) \times C_e$, and $(3 \times k \times w + 2) \times C_e$, respectively, since each element in a scalar are encrypted and processed independently and the computation complexity is proportional to the length of the scalar.

Figure 4a,b show the computation complexity of the data server with both schemes in terms of the scalar length and the number of vehicles. Simulation results show that the proposed scheme greatly reduces the computation complexity of the data server in comparison with the traditional scheme. As shown in Figure 4a, the computation complexity increases with both scalar length and the number of vehicles in the traditional scheme, while in the proposed scheme, the computation complexity only increases with the number of vehicles as shown in Figure 4b. Meanwhile, Figure 5a,b present the computation complexity of the data requester and vehicle in terms of the scalar length, and compare the proposed scheme with the traditional scheme, respectively. Figures 4 and 5 indicate that the proposed scheme can achieve lower computation complexity compared to the traditional scheme.

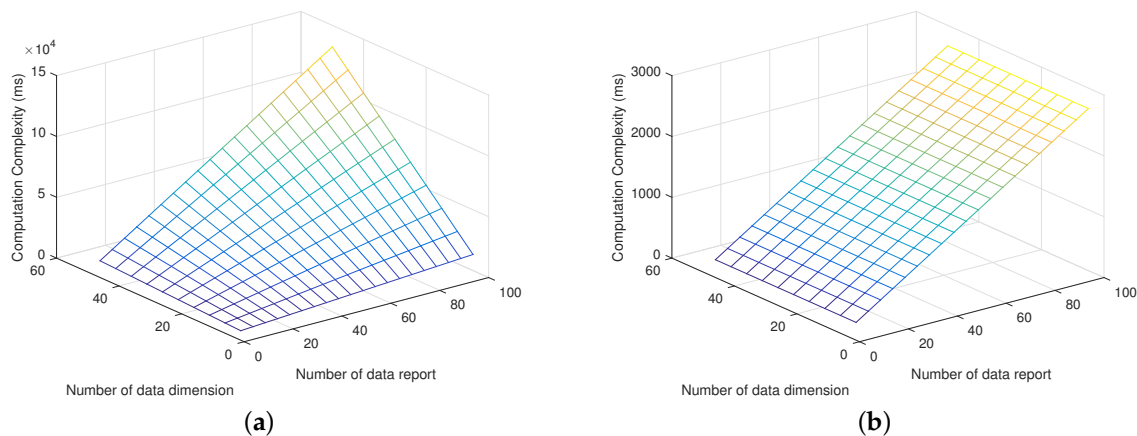


Figure 4. Computation complexity of data server. (a) Computation complexity of the data server with the traditional scheme; (b) Computation complexity of the data server with the proposed scheme.

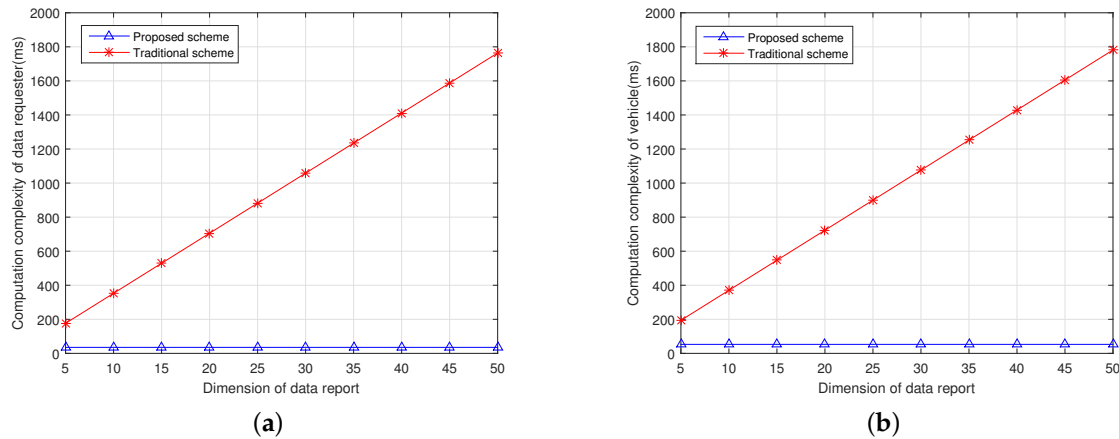


Figure 5. Comparison of the computation complexity of the data requester and vehicle. (a) Computation complexity of the data requester; (b) Computation complexity of one data uploading vehicle.

5.3. Communication Overhead

We evaluate the communication overhead that introduced the data query, which is sent from the data query towards the data server, and the data report sent from the vehicles towards the data server, since the data query sent from the data server towards the RSU and the data query response sent from the data server do not involve the transmission of ciphertexts in the proposed scheme.

We first consider the data requester to data server communication, where the data requester generates a location-based data query and delivers the data query to the data server. The location-based query is in the format of $C_{u,1} || C_{u,2}$, and its size is 2048×2 bits, if we choose 1024-bit n . If the traditional homomorphic scalar product protocol is adopted, the corresponding communication overhead is $2048 \times 2 \times k$ bits, when there are k partitioned grid cells. In Figure 6, we plot the data requester to data server communication overhead in terms of the scalar length, and compare the proposed scheme to the traditional scheme, which shows that the proposed scheme reduces the communication overhead.

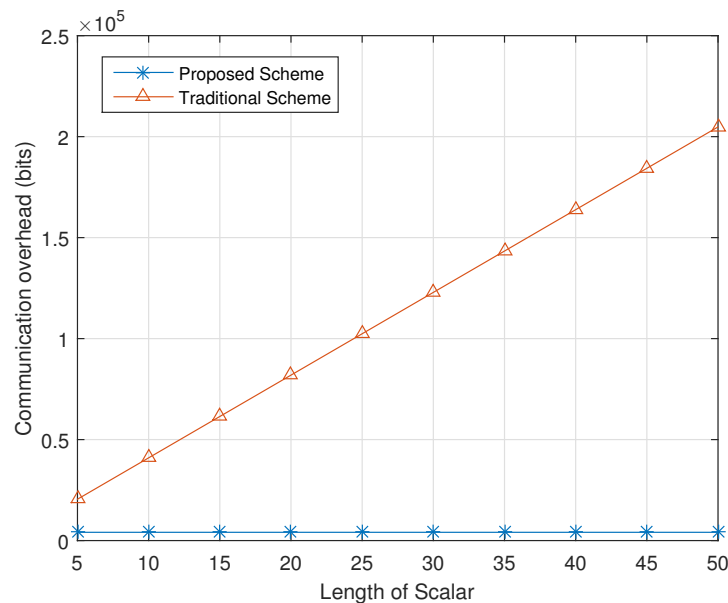


Figure 6. Communication overhead from the data requester to the data server.

During the vehicle to data server communication, each vehicle generates a location-based data report and delivers the data report to the data server via the RSU. The data report is in the format of $C_{u,0}||C_{u,1}||C_{u,2}$, and its size is 2048×3 bits. The data server collects w data reports from the total w vehicles, the communication overhead between the vehicles and the data server is $2048 \times 3 \times w$ bits. For the traditional homomorphic scalar product protocol, the communication between the vehicles and the data server is $2048 \times 3 \times w \times k$ bits, when the scalar length is k . Figure 7a,b plot the communication overhead introduced by vehicles of both schemes in terms of the scalar length and the number of vehicles. Simulation results show that the proposed scheme greatly reduces the vehicle to data server communication overhead in comparison with the traditional scheme. As shown in Figure 7a, the vehicle to data server communication overhead increases with both scalar length and the number of vehicles in the traditional scheme, while in the proposed scheme, vehicle to data server communication overhead only increases with the number of vehicles as shown in Figure 7b.

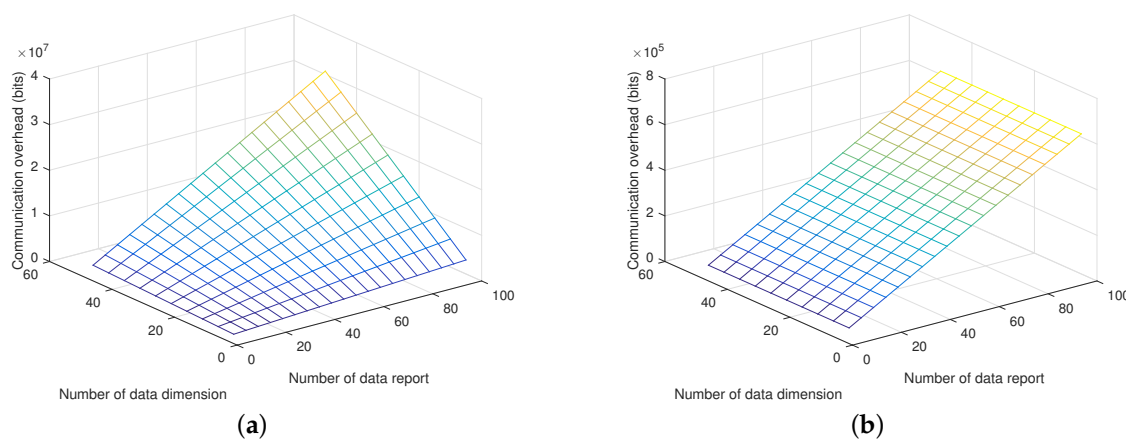


Figure 7. Vehicles to data server communication overhead. (a) Communication overhead of vehicles during data uploading with the traditional scheme; (b) Communication overhead of vehicles during data uploading with the proposed scheme.

6. Related Works

In this section, we briefly review some of the existing related schemes in secure scalar product computation and secure location-based query schemes.

6.1. Secure Scalar Product Computation

Research on secure scalar product computation has been focused on performing privacy-preserving profile matching [21–23], and secure multi-party computation [24,25]. A fine-grained privacy-preserving profile matching scheme is proposed with the Paillier homomorphic cryptosystem in [22], which enables two users to measure the level of similarity in their fine-grained personal files. An efficient privacy-preserving binary scalar product computation protocol is proposed in [23], which computes the similarity in symptom characters. A secure multi-party computation algorithm is proposed in [25] with the BGN homomorphic encryption technique, which allows the cloud server to execute secure scalar product and addition operations without the data content disclosure.

However, in the above schemes, each element in a scalar needs to be encrypted separately, which leads to the result that the size of the ciphertexts is proportional to the length to the scalar. In our proposed scheme, the elements in a given scalar is first structured to one-dimension and then encrypted, which is highly efficient in terms of communication and computation overhead.

6.2. Secure Location-Based Query

There mainly exist three types of secure location-based data query schemes. The first type of solutions are through pseudonyms [15,26]. However, the pseudonym-based solutions are based on the idea of disrupting the connection between the identities of vehicles and their locations, and the locations of the data queries can still be disclosed [11]. The second type of solutions are through location cloaking [27,28], and the identification of the exact location of each participant can be prevented by introducing uncertainty or error into location information. K-Anonymity is a commonly used technique in location cloaking, which preserves the location privacy of one user through hiding one user among a group of K users [29]. Even though location cloaking can achieve a satisfactory level of location privacy preservation, it may not be able to guarantee that there are enough users in the neighborhood and not be able to achieve high accuracy in their query results [30]. The third type of solutions rely on the cryptographic techniques, which can effectively preserve the accuracy of the data query results [31,32]. Various privacy-preserving data query schemes are proposed, such as range query [33]. An efficient spatial range query solution is proposed in [11], which permits data queries over encrypted location based data. The authors in [30] propose a hybrid approach based on location cloaking and the additive homomorphic encryption, which achieves both efficiency and privacy preservation. A spatial range query scheme over ciphertext is proposed in [34], which achieves the location privacy of the user's query and the location-based service confidentiality.

However, these schemes are based on the scenario of the outsourced central cloud storage data query and introduce heavy cryptographic operations, which also can not be applicable to the dynamically moving and distributive on-board storage. In this paper, we propose an efficient privacy-preserving location-based data query scheme, which can be applied to the scenario of the distributive vehicular on-board storage.

Based on the above analysis, in this paper, we propose an efficient location-based vehicular on-board sensory data querying scheme, which greatly reduces the network complexity and preserve the location privacy of both the data requester and vehicles.

7. Conclusions

In this paper, we have proposed a privacy-preserving range query scheme from the distributive on-board storage in VANETs, which achieves the secure scalar product computation with the homomorphic Paillier cryptosystem. The proposed scheme structures the multi-dimensional scalar into one dimension, identifies the data harvested within the industrial district, and computes the aggregation results of the identified sensed data. Security analysis has been conducted to demonstrate its security properties, in terms of location privacy preservation and confidentiality. Performance evaluations have also been done, which indicates that the proposed scheme can significantly reduce the computation complexity and communication overhead. For future work, we will take the possible behaviors of the collusion attack into consideration and design new solutions to resist such attacks.

Acknowledgments: This research was supported in part by the research grant S15-1105-RF-LLF URBAN from the Economic Development Board, Singapore, for the project of "Development Of NTU/NXP Smart Mobility Test-bed". This research was supported in part by Natural Sciences and Engineering Research (NSERC) Discovery Grants (No. Rgpin 04009), NBIF Start-Up Grant (Rif 2017-012), and a URF Grant (No. 124419). This work was supported by the Zhejiang Science Technology Department Project (No. 2016C33173), the Zhejiang Education Department Project (No. Y201533448), and the Natural Science Foundation of Zhejiang Province (LY17F020006).

Author Contributions: Qinglei Kong and Rongxing Lu conceived the project; Maode Ma contributed the system analysis; Haiyong Bao performed the experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abdelhamid, S.; Hassanein, H.S.; Takahara, G. Vehicle as a Mobile Sensor. In Proceedings of the 9th International Conference on Future Networks and Communications, Niagara Falls, ON, Canada 17–20 August 2014; pp. 286–295.
2. Lee, U.; Magistretti, E.; Gerla, M.; Bellavista, P.; Corradi, A. Dissemination and Harvesting of Urban Data Using Vehicular Sensing Platforms. *IEEE Trans. Veh. Technol.* **2009**, *58*, 882–901.
3. Devarakonda, S.; Sevusu, P.; Liu, H.; Liu, R.; Iftode, L.; Nath, B. Real-time air quality monitoring through mobile sensing in metropolitan areas. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, UrbComp@KDD 2013, Chicago, IL, USA, 11 August 2013; pp. 15–1–15–8.
4. Eriksson, J.; Girod, L.; Hull, B.; Newton, R.; Madden, S.; Balakrishnan, H. The pothole patrol: Using a mobile sensor network for road surface monitoring. In Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys 2008), Breckenridge, CO, USA, 17–20 June 2008; pp. 29–39.
5. Yang, G.; He, S.; Shi, Z.; Chen, J. Promoting Cooperation by the Social Incentive Mechanism in Mobile Crowdsensing. *IEEE Commun. Mag.* **2017**, *55*, 86–92.
6. Lu, R.; Lin, X.; Luan, T.H.; Liang, X.; Shen, X.S. Pseudonym Changing at Social Spots: An Effective Strategy for Location Privacy in VANETs. *IEEE Trans. Veh. Technol.* **2012**, *61*, 86–96.
7. Whaiduzzaman, M.; Sookhak, M.; Gani, A.; Buyya, R. A survey on vehicular cloud computing. *J. Netw. Comput. Appl.* **2014**, *40*, 325–344.
8. Wen, M.; Lu, R.; Liang, X.; Lei, J.; Shen, X.S. *Querying over Encrypted Data in Smart Grids*; Springer International Publishing AG: Cham, Switzerland, 2014.
9. Lee, U.; Gerla, M. A survey of urban vehicular sensing platforms. *Comput. Netw.* **2010**, *54*, 527–544.
10. Elmehdwi, Y.; Samanthula, B.K.; Jiang, W. Secure k-Nearest Neighbor Query over Encrypted Data in Outsourced Environments. In Proceedings of the 30th IEEE International Conference on Data Engineering, ICDE 2014, Chicago, IL, USA, 31 March–4 April 2014; pp. 664–675.
11. Li, L.; Lu, R.; Huang, C. EPLQ: Efficient Privacy-Preserving Location-Based Query over Outsourced Encrypted Data. *IEEE Internet Things J.* **2016**, *3*, 206–218.
12. Paillier, P. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In Proceedings of the Advances in Cryptology—EUROCRYPT '99, Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, 2–6 May 1999; pp. 223–238.
13. Mapping the Invisible: Street View Cars Add Air Pollution Sensors. Available online: <https://environment.google/projects/airview/> (accessed on 2 June 2017).
14. 802.11 p-2010-IEEE Standard for Information Technology-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless Lan Medium Access Control (Mac) and Physical layer (Phy) Specifications Amendment 6: Wireless Access in Vehicular Environments, 2010. Available online: <http://standards.ieee.org/findstds/standard/802.11p-2010.html> (accessed on 2 June 2017).
15. 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.
16. Al-kahtani, M.S. Survey on security attacks in Vehicular Ad hoc Networks (VANETs). In Proceedings of the 6th International Conference on Signal Processing and Communication Systems, ICSPCS 2012, Gold Coast, Australia, 12–14 December 2012; pp. 1–9.
17. Zhu, H.; Lu, R.; Huang, C.; Chen, L.; Li, H. An Efficient Privacy-Preserving Location-Based Services Query Scheme in Outsourced Cloud. *IEEE Trans. Veh. Technol.* **2016**, *65*, 7729–7739.
18. Schlegel, R.; Chow, C.; Huang, Q.; Wong, D.S. User-Defined Privacy Grid System for Continuous Location-Based Services. *IEEE Trans. Mob. Comput.* **2015**, *14*, 2158–2172.
19. Beijing Fangshan Industrial Park, 2011. Available online: <https://www.researchandmarkets.com/reports/1946272/beijingfangshanindustrialpark.pdf> (accessed on 2 June 2017).
20. Liu, K. Paillier's Homomorphic Cryptosystem (Java Implementation). Available online: <https://www.csee.umbc.edu/~kunliu1/research/Paillier.html> (accessed on 2 June 2017).
21. Lu, R.; Lin, X.; Liang, X.; Shen, X. A Secure Handshake Scheme with Symptoms-Matching for mHealthcare Social Network. *MONET* **2011**, *16*, 683–694.
22. Zhang, R.; Zhang, J.; Zhang, Y.; Sun, J.; Yan, G. Privacy-Preserving Profile Matching for Proximity-Based Mobile Social Networking. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 656–668.

23. Lu, R.; Lin, X.; Shen, X.S. SPOC: A Secure and Privacy-Preserving Opportunistic Computing Framework for Mobile-Healthcare Emergency. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *24*, 614–624.
24. Goethals, B.; Laur, S.; Lipmaa, H.; Mielikäinen, T. On Private Scalar Product Computation for Privacy-Preserving Data Mining. In Proceedings of the 7th International Conference on Information Security and Cryptology—ICISC 2004, Seoul, Korea, 2–3 December 2004; pp. 104–120.
25. Yuan, J.; Yu, S. Privacy Preserving Back-Propagation Neural Network Learning Made Practical with Cloud Computing. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 212–221.
26. Shilton, K. Four billion little brothers? privacy, mobile phones, and ubiquitous data collection. *Commun. ACM* **2009**, *52*, 48–53.
27. Xu, T.; Cai, Y. Exploring Historical Location Data for Anonymity Preservation in Location-Based Services. In Proceedings of the 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, Phoenix, AZ, USA, 13–18 April 2008; pp. 547–555.
28. Ni, W.; Gu, M.; Chen, X. Location privacy-preserving k nearest neighbor query under user's preference. *Knowl.-Based Syst.* **2016**, *103*, 19–27.
29. Huang, K.L.; Kanhere, S.S.; Hu, W. Preserving privacy in participatory sensing systems. *Comput. Commun.* **2010**, *33*, 1266–1280.
30. Tian, S.; Cai, Y.; Zheng, Q. A hybrid approach for privacy-preserving processing of knn queries in mobile database systems. In Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, CIKM'13, San Francisco, CA, USA, October 27–November 1 2013; pp. 1161–1164.
31. Bilogrevic, I.; Jadliwala, M.; Kalkan, K.; Hubaux, J.; Aad, I. Privacy in Mobile Computing for Location-Sharing-Based Services. In Proceedings of the Privacy Enhancing Technologies—11th International Symposium, PETS 2011, Waterloo, ON, Canada, 27–29 July 2011; pp. 77–96.
32. Puttaswamy, K.P.N.; Wang, S.; Steinbauer, T.; Agrawal, D.; El Abbadi, A.; Kruegel, C.; Zhao, B.Y. Preserving Location Privacy in Geosocial Applications. *IEEE Trans. Mob. Comput.* **2014**, *13*, 159–173.
33. Hore, B.; Mehrotra, S.; Canim, M.; Kantarcioglu, M. Secure multidimensional range queries over outsourced data. *VLDB J.* **2012**, *21*, 333–358.
34. Zhu, H.; Liu, F.; Li, H. Efficient and Privacy-Preserving Polygons Spatial Query Framework for Location-Based Services. *IEEE Internet Things J.* **2017**, *4*, 536–545.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).