A Road Truncation-Based Location Privacy-Preserving Method against Side-Weight Inference Attack

Qingyuan Li 1, Hao Wu 1,2,*, Xiang Wu 1, Ning Zhao 3 and Lan Dong 4

1 State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China; qingyuanli@bjtu.edu.cn (Q.L.); 18120149@bjtu.edu.cn (X.W.)
2 Frontiers Science Center for Smart High-Speed Railway System, Beijing Jiaotong University, Beijing 100044, China
3 National Computer Network Emergency Response Technical Team, Coordination Center of China, Beijing 100029, China; work_zn2022@163.com
4 School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China; donglan@bjtu.edu.cn
* Correspondence: hwu@bjtu.edu.cn; Tel.: +86-10-51688260-602

Abstract: Taking advantage of precise positioning technology, location-based service (LBS) has brought a lot of convenience to people’s daily life and made the city smarter. However, the LBS applications also bring some challenges to personal location privacy protection. In order to obtain services from LBS providers, users have to upload their queries including sensitive information, such as identities and locations. This information may be leaked out by the LBS providers or even eavesdropped on by malicious adversaries, which may cause privacy leakage. To tackle this problem, many solutions have been investigated under the assumption that users are uniformly distributed. However, the users are not always uniformly distributed in real-world situations. For a side-weight inference attack, the adversary would infer that the target user is more likely to belong to the road section with more users, resulting in performance deterioration. In this paper, we investigate the issue of location privacy preservation against side-weight inference attack for non-uniform distributed road network. Meanwhile, we consider the cost function of LBS and formulate the object as a mixed integer programming problem. Then, we propose a road truncation-based scheme to protect location privacy. The road section with high user density is designed to be truncated. Finally, simulation results show that our scheme meets the demand for privacy protection at a low cost. As a result, our scheme is proven to protect users’ location privacy effectively and efficiently.

Keywords: location privacy; l-diversity; road truncation; side-weight inference attack

1. Introduction

With the development of wireless communication technology and smart mobile devices, many applications and services have emerged and enriched our daily life [1]. Among them, location-based service (LBS) is one kind of the most important and popular services that are widely used [2]. LBS can provide various services such as navigation, query of points of interest (POIs), ride-sharing, location-based games and some derivative applications [3]. Because of these conveniences, people’s demand for LBS has greatly increased.

LBS brings benefits and convenience to daily life whereas it poses a risk to users’ privacy [4]. In order to obtain LBS, mobile users have to send query requests, which are sensitive as they may contain the users’ personal information, such as user ID, precise location and query content, to LBS providers [5]. Based on a user’s personal information, it is easy to find out private information such as a home address, working place, habits and health status [6,7]. People are willing to use LBS and provide sensitive information, as they believe that the service providers are trusted. However, the LBS providers are honest but curious about the information. They would like to collect the information uploaded
by the users to make optimization of their query algorithm or to sell to other companies. Even if the LBS provider can be fully trusted, there are still some malicious users that may eavesdrop on the channel and get the information of the users. What is worse, once the LBS provider is hacked, the adversary would get the uploaded information of users. The leakage of location privacy has become an issue that cannot be ignored [8]. In 2017, the records of global data leakage and theft reached 1.6 billion, causing huge economic losses and causing consumers’ extreme concerns about privacy. In 2019, the overall economic loss of Chinese users was more than 10 billion dollars due to personal information disclosure, fraud and other reasons. In recent years, data abuse events such as Facebook have aroused widespread concern in the society [9]. Due to security system vulnerabilities, up to 87 million Facebook users’ private information was leaked, which contains name, contact information, search records, login location, etc. The disclosure of this sensitive information may lead to fraud, theft and other crimes [10]. Therefore, the problem of privacy disclosure needs to be solved urgently as it hinders the development of LBS [11].

In order to solve this problem, many studies focused on addressing how to protect privacy in LBS during the past several years based on location obfuscation and encryption [12], such as \( k \)-anonymity [13], mixed zone [14], caching [15], dummy locations [16], homomorphic encryption [17] and differential privacy [18–20]. Among them, the most classic location privacy preservation scheme is \( k \)-anonymity. In \( k \)-anonymity scheme, the location of a user sent to the server is a region, namely anonymous set, instead of precise coordinates. There are at least \( k \) users in the anonymous set. The adversary can not distinguish the real location of a specific user from the other \( k - 1 \) obfuscation locations. The recent researches mostly focus on Euclidean geometric space, where users can be anywhere in the space. However, in real-world situations, a user’s activity is restricted by road network topology, especially for vehicle users, which brings new challenges to location privacy protection. As a user is on a specific road, the adversary can guess the road to which the user belongs. Similar to \( k \)-anonymity, Machanavajjhala introduced the idea of \( l \)-diversity [21]. The anonymous set should consist of not only \( k \) users but also \( l \) different road sections.

Although the \( k \)-anonymity and \( l \)-diversity schemes protect location privacy, the uneven distribution of users will lead to the performance degradation of the privacy protection scheme. In the urban scene, there are a large number of vehicles on the main roads, while there are a few vehicles on other roads. Specific query users are distributed on high-density roads with high probability, so attackers infer that users are located on high-density roads with high probability. The specific query user is more likely to lie on the road with more users. Based on the background knowledge, the attacker obtains the user’s distribution information and infers which road the user is on with high probability, namely side-weight inference attack, which reduces the performance of location privacy protection. However, the user distribution is ignored. Motivated by this, we investigate the issue of location privacy preservation for the non-uniform distributed road network.

In this paper, we propose a privacy preserving method based on road truncation, considering non-uniform distributed road network. In our scheme, the road section with high user density is designed to be truncated to reduce the density and then added to the anonymous set. As a result, the divergence between different road sections decreases, which is harder for the adversary to track the user. On the premise of privacy protection, we also consider reducing the query cost of the algorithm. In \( k \)-anonymity based schemes, the location information uploaded by the user is an area containing several users and roads, which affects the quality of LBS. Our aim is to enable the LBS to protect users’ location privacy while keeping the location query service relatively efficient. In summary, the contributions of this work are as follows:

- We investigate the issue of location privacy preservation for non-uniform distributed road network. We also consider the cost function and formulate the object as a mixed integer programming problem.
- We propose a privacy-preserving scheme that achieves location privacy protection based on road truncation. The proposed scheme meets the demand of privacy protec-
tion for non-uniform distributed road network at a low cost. As a result, the proposed scheme provides good quality of LBS.

- We analyze the scheme in terms of privacy and efficiency with theoretical analysis and the simulation results show that our scheme works well with both privacy and efficiency. The proposed scheme achieves \( k \)-anonymity and \( l \)-diversity and resists side-weight inference attack. Meanwhile, the proposed scheme maintains a low cost level.

The rest of this paper is organized as follows. Section 2 reviews the related works of location privacy preserving schemes. In Section 3, we give some preliminaries and explain how privacy leakage happens. In Section 4, the motivations and aims are described and the problem is formulated. Then, we introduce our proposed scheme. Privacy analysis and numerical results are shown in Section 5 to evaluate the performance of the proposed scheme. Finally, we present the conclusion in Section 6.

2. Related Work

In order to provide reliable service, many algorithms have been suggested to protect location privacy in LBS during the past several years. These algorithms are divided into the following categories [22]. One popular strategy is based on spatial and temporal obfuscation such as \( k \)-anonymity, caching schemes and dummy location schemes. In \( k \)-anonymity schemes, the real location is obfuscated and the user is hidden among many users. These algorithms are simple and easy to implement and suitable for most cases. The precision of location, on the other hand, is sacrificed [23]. Caching schemes provide accurate location service, but they cannot meet the demands of real-time services [24]. In dummy location schemes, many false location data is generated and sent to the server as well as the real one, which causes a waste of resources [25]. Another strategy to protect location privacy is the cryptography method, such as homomorphic encryption [26,27]. These algorithms need high-performance computational capacity to encrypt and decrypt the messages, which takes a lot of time. So the algorithms are not suitable for low-performance equipment and real-time services. In Table 1, we summarise the location privacy preserving schemes.

Among all the location privacy preserving algorithms, the most widely used algorithm is \( k \)-anonymity. The concept \( k \)-anonymity is first proposed in the field of data publishing for data desensitization [28]. If any arbitrary attribute is contained in at least \( k \) records, the dataset achieves \( k \)-anonymity. This ensures that any individual cannot be easily distinguished by linking attacks. Later, this idea is applied to protect location privacy [29]. In each anonymous set, there are at least \( k \) users. By this means, it protects location privacy from the LBS provider and the probability of tracking a specific user is reduced to \( 1/k \).

Most of the previous studies are based on Euclidean geometric space. However, a user’s activity is restricted by road network topology, which brings new challenges in protecting location privacy. A method named \( l \)-diversity is proposed for query privacy preservation based on road network topology [21], where the anonymous set should consist of not only \( k \) users but also \( l \) different road sections. Based on \( l \)-diversity, Chow et al., proposed a new cost function that balanced between the query execution cost and the query quality [30]. An algorithm is proposed to minimize the query execution cost. Liu and Wang introduced an X-Star scheme [31]. It achieves \( l \)-diversity and has good extendibility. To tackle the problem of low anonymous success rate, Mouratidis and Man proposed a new X-Star scheme based on Hilbert index, named H-Star [32]. In [33], Rubner proposed a method that transformed Euclidean geometric space into the road network environment. Considering the road network environment, Xue et al., proposed an anonymous ring scheme to realize location privacy protection [34]. In [35] a voronoi graph based algorithm is proposed. It considers the problem when users are not uniformly distributed and resists the side-weight inference attack.
Table 1. Summary of Privacy Preserving Schemes that address Location Privacy.

| Privacy Preservation Mechanism | Reference | Characteristic |
|-------------------------------|-----------|----------------|
| k-anonymity                   | [13,28]   | flexible and easy to implement; lack of location accuracy |
| l-diversity                   | [31,32,35]| k-anonymity method considering road network |
| Caching scheme                | [15,24]   | accurate location; not applicable for real-time service |
| Dummy location                | [16,25]   | accurate location; false location data causing resource waste |
| Homomorphic encryption        | [17,26,27]| ciphertext can be calculated; high computation cost |
| Differential privacy          | [18–20]   | strict proof of privacy assessment |

3. Preliminaries and Problem Definition

In this section, we review the background knowledge of location privacy and explain how privacy leakage happens.

3.1. System Model

In this paper, we consider the location privacy preserving issue for LBS in a city scenario. The mathematical symbols are shown in Table 2.

Table 2. Mathematical symbols.

| Parameter | Description |
|-----------|-------------|
| C         | the user’s query content |
| e_i       | road section i |
| E(S)      | the number of edges in anonymity set S |
| l         | l-diversity |
| loc       | the user’s location information |
| k         | k-anonymity |
| n         | the number of query objects required by the user |
| Q         | LBS query message |
| S         | anonymity set |
| t_0       | the starting time of the algorithm |
| t_n       | the current time of the algorithm |
| u_i       | a user’s identity |
| V_o(S)    | the number of marginal nodes |
| w(e_i)    | the weight of edge e_i |
| (x, y)    | location coordinates |
| α         | number of roads in candidate set C_i |
| β         | number of roads that are truncated in C_i |
| δ_t       | the tolerant time for generating an anonymity set |
| λ         | distribution degree |
| ε         | the ratio between the weight of edge e_i and the whole of the anonymous set |

As shown in Figure 1, the system consists of three parts, users, base stations (BS) and LBS providers. The users’ activities are restricted by road network topology. With embedded GPS sensors, the users can get precise location coordinates and velocity. The base stations are connected by optical fibers and communicate with all users. They have background knowledge of road network topology and user distribution.

When a user acquires the LBS, it needs to upload its personal location to the LBS providers through the base station. Navigation and query of POIs are basic LBS applications, which facilitate people’s daily life. For example, the POI helps the user discover unknown environmental resources. A user can easily get the location information of interested buildings such as shops, banks and gas stations in the surrounding area. In order to protect location privacy, the user will communicate with the base station to get the information of user distribution and generate an anonymous set. The query message is uploaded to the base station and finally forwarded to LBS providers. The LBS provider deals with the query request and responds to the user with query results according to the anonymous set.
3.2. LBS Query Message

An original LBS query message is defined as \( Q = (u_{id}, \text{loc}, C) \), where \( u_{id} \) denotes a user’s identity; \( \text{loc} \) represents the user’s location information, usually precise location coordinates \((x, y)\); \( C \) denotes the user’s query content.

To preserve location privacy, the message is transformed. The location information can be a circular area, a polygonal area composed of multiple points, or a collection of multiple edges.

3.3. Adversary Model

In general, the adversary may be a malicious user or an LBS provider. The LBS provider is supposed to be honest but curious about the data. It would collect the information uploaded by the users to make optimization of their algorithm or to sell the information to other companies. In addition, some malicious users eavesdrop on the channel and get query information. The adversary may infer the user’s location based on the background information.

3.4. Side-Weight Inference Attack

Before illustrating the scheme in this paper, we first discuss the problem of location privacy disclosure and the corresponding protection schemes. Then, we explain how a side-weight inference attack happens, under the scenario of non-uniform distribution of users.

In the \( k \)-anonymity scheme, there are at least \( k \) users in each anonymous set. As shown in Figure 2, there are four users in area A. When a user (marked as a triangle) in area A wants to get the service, it sends the query request with region A instead of its precise coordinates. The adversary gets the query message and knows the user belongs to area A. However, there is only an obfuscated area. The adversary cannot distinguish the target user from other users. Without considering the distribution of users, the probability of the adversary targeting the specific user is \( 1/k \). As a result, the successful tracking probability of the target user is \( 1/4 \). Similarly, for anonymous set B, the successful tracking probability is \( 1/3 \).
Most of the early researches are based on Euclidean geometric space. In fact, many users of LBS are vehicles and restricted in road sections. As shown in Figure 3, there are $3 \times 3$ blocks with 12 streets ignoring the marginal road sections. For both anonymous sets A and B, there are four users. There is one road across the anonymous set A. Meanwhile, there are four roads across the anonymous set B. Obviously, the users in anonymous set B achieves higher privacy. Similar to the definition of $k$-anonymity, the anonymous set should contain at least $l$ different road sections to fulfill the $l$-diversity of query requirement. Regardless of the differences of road sections, the probability of the adversary targeting the specific road section is $1/l$.

In real-world situations, the users are not uniformly distributed. Hence, the side-weight inference attack happens as the adversary would infer that the target user belongs to a high-density road section with high probability. As shown in Figure 4, the anonymous set is the blue circle which contains four road sections $\{e_1, \ldots, e_4\}$. There are 13 users in the anonymous set. We can guess the target user is more likely to belong to the edge $e_1$, as there are more users in $e_1$. 

**Figure 2.** Location privacy protection based on $k$-anonymity.

**Figure 3.** Location privacy protection based on road network.
In order to evaluate the performance of the proposed scheme against side-weight inference attack, we introduce two definitions to describe the tracking probability and distribution degree of users. First, we define the tracking probability $\epsilon$.

$$\epsilon = \frac{w(e_i)}{\sum_{i=1}^{n} w(e_i)}$$

(1)

where $w(e_i)$ stands for the weight of edge $e_i$, namely, the number of users on road $e_i$. Parameter $\epsilon$ is the ratio between the weight of edge $e_i$ and the whole of the anonymous set, which depicts the tracking probability of the adversary targeting on the specific road section.

Then, distribution degree $\lambda$ is defined as follow,

$$\lambda = \frac{\max w(e) - \min w(e)}{\overline{w}(e)}$$

(2)

where $\max w(e)$ and $\min w(e)$ represent the maximum and minimum values respectively. $\overline{w}(e)$ stands for the average weight of all road sections. Parameter $\lambda$ is a normalized factor, which describes the biggest difference of distribution in different sections.

4. Problem Formulation and Proposed Scheme

In this section, we introduce the motivation of this paper and the problem of location privacy protection is formulated. Then we propose a privacy preserving scheme based on road truncation.

4.1. Design Goal

In order to achieve location privacy protection, the algorithm has to satisfy three goals simultaneously.

4.1.1. k-Anonymity

First of all, the algorithm has to satisfy $k$-anonymity demand. Suppose, there are at least $k$ users in an anonymous set, the probability that the attacker can track the specific user is at most $1/k$. If the privacy request of user $i$ is $k_0$, the anonymous set size $k$ has to satisfy $k \geq k_0$.

The anonymous set size $k$ can be described as

$$k = \Sigma a_i \ast w(e_i),$$

(3)
where \( a_i \) depicts whether road \( e_i \) is in the anonymous set \( S \).

\[
a_i = \begin{cases} 
0, & e_i \notin S \\
1, & e_i \in S 
\end{cases}
\]

4.1.2. \( l \)-Diversity

In order to realize privacy protection in road network topology, the algorithm has to satisfy \( l \)-diversity demand. If the privacy request of user \( i \) is \( l_0 \), the road number \( l \) of anonymous set has to satisfy \( l \geq l_0 \).

The road number \( l \) of anonymous set can be described as

\[
l = \Sigma a_i.
\]

4.1.3. Resist Side-Weight Inference Attack

Since the users are not uniformly distributed. The adversary would infer that the target user is more likely to belong to the road section with more users. Considering this, the distribution degree has to satisfy \( \lambda \leq \lambda_0 \), where \( \lambda_0 \) stands for the privacy requirement.

4.2. Cost Model

In this section, we discuss the cost model for LBS. This cost model will be used to evaluate the efficiency of the privacy preserving algorithm. Take the POI as an example, the cost is calculated in detail. A typical example of POI is to find the \( k \) nearest objects, namely k-nearest-neighbor (k-NN) query. In order to protect location privacy, the location information is an anonymous area containing the user. The user is either in the internal area or on the border of the anonymous set. Based on this, we divide the query cost into two parts. The cost of the inner search is the size of the anonymous set. The outer search is to find the objects closest to the marginal nodes. We define the query execution cost of a private query as:

\[
Cost(S, n) = E(S) + V_o(S) \times n,
\]

where \( E(S) \) stands for the number of edges in anonymity set \( S \), \( V_o(S) \) stands for the number of marginal nodes and \( n \) stands for the number of query objects required by the user.

In order to protect location privacy, many users and edges are added to the anonymous set \( S \), which will increase the query cost. Therefore, the algorithm should reduce the cost as much as possible on the premise of privacy protection. Based on the discussion above, we formulate the problem as,

\[
\begin{align*}
\min_{\{a_i\}} \quad & \text{Cost}(S, n) \\
\text{s.t.} \quad & k \geq k_0, \\
& l \geq l_0, \\
& \lambda \leq \lambda_0,
\end{align*}
\]

where constraint Equations (6b) and (6c) are integers, and Equation (6d) is fractional. Thus, the problem is a mixed integer program problem. We give a heuristic algorithm in the next section, as the problem is a non-convex optimization problem.

4.3. Proposed Scheme

According to the road network topology, we propose a road truncation-based scheme to generate the anonymous set. The proposed scheme should protect location privacy while reducing the cost as much as possible. As the users are not always uniformly distributed. Our scheme is based on road truncation, that is, the road sections with higher user density
would be truncated before generating anonymous sets. Firstly, we introduce the flow of our scheme and then describe the anonymous generation algorithm in detail.

4.3.1. Workflow

The workflow of the proposed scheme is shown in Figure 5 and the detailed scheme is shown as follows. When a user launches a query request, it firstly communicates with the base station to get the background knowledge of road network topology and user distribution. Then the user runs the Road Truncation (R-T) Algorithm 1 to generate an anonymous set. As our proposed scheme achieves not only \( k \)-anonymity and \( l \)-diversity, but also resist the side-weight inference attack, it may take a long time to generate an anonymous set. In order to ensure the quality of service (QoS), we define a tolerant time \( \delta_t \). If the anonymous set is not successfully generated within the tolerant time \( \delta_t \), the constraint Equation (6d) is neglected and the anonymous set is generated by algorithm GA [30], which meets the demand of \( k \)-anonymity and \( l \)-diversity. The tolerant time \( \delta_t \) can be described as

\[
t_n - t_0 < \delta_t,
\]

where \( t_n \) is the current time and \( t_0 \) is the starting time of the algorithm.

Algorithm 1 The R-T algorithm.

| Input: \( e_i, k_0, l_0, \lambda_0 \) |
| Output: anonymous set \( S_i \) |
| 1: | add \( e_i \) to candidate set \( C_i \) |
| 2: | while \( N(C_i) < a \) do |
| 3: | if \( e_j \) is adjacent to \( C_i \) then |
| 4: | add \( e_j \) to candidate set \( C_i \) |
| 5: | end if |
| 6: | end while |
| 7: | compute the average edge weight \( \bar{w} \) |
| 8: | sort all edges in \( C_i \) by the weight |
| 9: | add \( e_i \) to candidate set \( S_i \) |
| 10: | while \( k < k_0 \) do |
| 11: | if \( e_h (e_h \in C_i) \) is adjacent to \( S_i \) then |
| 12: | add \( e_h \) to anonymous set \( S_i \) |
| 13: | end if |
| 14: | end while |
| 15: | while \( k < k_0 \) do |
| 16: | if \( e_k \) is the bigger \( \beta \) edges in \( C_i \) then |
| 17: | truncate \( e_k \) |
| 18: | refresh \( C_i \) |
| 19: | end if |
| 20: | end while |

4.3.2. Road Truncation

In this section, we propose an algorithm for generating an anonymous set. Our algorithm is based on road truncation (R-T). The detailed information is shown in Algorithm 1.

In the first step, we compute the average edge weight \( \bar{w} \). We locate the user \( u_i \) and adds edges \( e_i \) to the candidate set \( C_i \). Then we find out all edges that are adjacent to \( C_i \) and add them to \( C_i \). If there are few edges in \( C_i \), repeat this procedure until the number of roads in \( C_i \) reaches the threshold \( a \). Then we get the average edge weight,

\[
\bar{w} = \frac{1}{a} \sum_{i=1}^{a} w(e_i).
\]
In the second step, the system sorts all edges in \( C_i \) by the weight and adds \( e_h \) to anonymous set \( S_i \), until the parameter have to satisfy \( k \geq k_0, l \geq l_0 \). Then we makes road \( e_k \) truncated with highest edge weight according to the rule that

\[
e_k = \left| \frac{1}{n} \sum_{i=1}^{n} w(e_i) \right|.
\]

(8)

Repeat the procedure and make sure that \( \lambda \leq \lambda_0 \). Finally, the output anonymous set satisfies the \( k \)-anonymity demand and the \( l \)-diversity demand and the algorithm has a good ability to resist the side-weight inference attack.

![Diagram](image.png)

**Figure 5.** The workflow of our algorithm based on road truncation.

5. Performance Evaluation

In this section, we evaluate the performance of the scheme with both theoretical analysis and simulation results.

5.1. Privacy Analysis

In this part, we analyze the privacy protection performance with privacy degree to make sure that the adversary can not infer the real user’s location among the users of an anonymous set. We quantify the privacy protection degree with anonymity parameters and distribution parameter \( \lambda \). The anonymous parameters contain the number of users and edges in the anonymous set, which are frequently used to evaluate the degree of privacy protection. Obviously, it would be easy to track the user when there are few elements in the anonymous set. Usually, the user will set a threshold to realize location privacy protection \( \{k_0, l_0, \lambda_0\} \).

With the background information, the adversary can get the locations of users and utilize the non-uniform distribution in the road network to determine the correct road section where a user lies in. We consider the distribution degree \( \lambda \), which describes the biggest difference in the distribution of users in different sections. Our method is based on road truncation. The road sections with more users are truncated. As a result, the distribution degree \( \lambda \) decreases. The divergence between different road sections decreases, which is harder
for the adversary to track the user. Our scheme performs well and resists the side-weight inference attack.

5.2. Simulation Results

In this section, we evaluate the performance through simulation results. The experiment was written in Matlab, running in Intel Core i7-4610M 3.00 GHz CPU with 8 GB DDR memory of 64-bit Windows 7 operating system. We consider an environment that consists of tens of road sections. The user follows universal distribution and the user distribution degree $\lambda$ varies from 2 to 4. We assume $\alpha = 3 \times l$, where $\alpha$ stands for the number of roads in the candidate set.

We evaluate the performance of our proposed scheme. Moreover, we compare the performance of our proposed scheme with some existing methods, such as the GA in [30], AR in [34], VW in [35] and exhaustive method. In [30], the cost function is considered a greedy approach (GA) is proposed. In [34], the adjacent road sections are preferentially added to the anonymous set. In [35], the tracking probability is considered and a voronoi map (VW) based algorithm is proposed. First, we show the performance of location privacy preserving. Then the cost of LBS and time consumption of generating an anonymous set are evaluated.

From Figure 6, we can see the anonymous set parameters, such as the anonymous set size $k$, road section number $l$ and distribution degree $\lambda$. It is obvious that all of the four schemes meet the demand of $k$-anonymity and $l$-diversity ($k_0 = 10$, $l_0 = 3$). In other words, all the schemes can provide primary privacy protection for users. The anonymous set size and road section number of VW and our scheme are bigger than in GA and AR. It is because VW and our scheme consider the user distribution. In order to meet the needs of more strict privacy protection, more users and road sections are added to the anonymous set. Only our proposed scheme meets all the demands of privacy protection, including $k$-anonymity and $l$-diversity and distribution degree $\lambda$. The proposed scheme can resist the side-weight inference attack and provide better privacy protection for the non-uniform distributed road network.

![Figure 6. Anonymous set parameters. ($k_0 = 10$, $l_0 = 3$, $\lambda_0 = 2$).](image)

Figure 7 gives the analysis of the tracking probability $\epsilon$. We observe that the tracking probability would decrease with the increase of $l_0$. With the increase of $l_0$, there are more road sectors in the anonymous set. It becomes more difficult for the adversary to track the user. The tracking probability of VW and our scheme are smaller than that of GA and AR. It is because there are more road sectors in the anonymous set in VW and our scheme.
than in GA and AR, so as to meet the needs of more strict privacy protection. As the cost function is considered in our algorithm, the anonymous set size of our scheme is a bit smaller than in VW. Therefore, the tracking probability of our algorithm is a bit bigger than in VW. Our scheme achieves privacy protection with a smaller anonymous set.

Figure 7. The tracking probability $\epsilon$ comparison with different query requirement $l_0$. ($k_0 = 10$, $\lambda_0 = 2$).

Figure 8 shows the distribution degree of anonymous set in different schemes. It is obvious that the scheme proposed in this paper meets the privacy protection requirements in all scenarios. In our proposed scheme, the road section with high user density is designed to be truncated. As a result, the distribution degree $\lambda$ of the anonymous set decreases. Since the distribution of users tends to become uniform, it is hard for the adversary to track the user. Our proposed scheme can resist the side-weight inference attack. On the contrary, the other three schemes do not change the user distribution. The distribution degree $\lambda$ of the anonymous set increase with the distribution degree of users in the map. The adversary would infer that the target user belongs to a high-density road section with high probability. The location privacy is not well protected against side-weight inference attacks.

Figure 8. The distribution degree of anonymous set comparison with different user distribution degree. ($k_0 = 10$, $l_0 = 3$, $\lambda_0 = 2$).
The scheme proposed in this paper is designed to protect location privacy against side-weight inference attacks while reducing the cost as much as possible. In the following, the performance of the cost of LBS and time consumption of generating an anonymous set is evaluated. We compare the performance of our proposed scheme with VW and exhaustive method, which consider the non-uniform distributed users.

Figure 9 shows the impact of privacy parameter $l_0$ on LBS query cost. The LBS query cost increases along with the increase of privacy parameter $l_0$. With the increase of $l_0$, there are more users and road sectors in the anonymous set. Simultaneously, we can see that the cost of our algorithm is lower than VW. The scheme proposed in this paper is based on road truncation, where the road section with high user density is designed to be truncated. Based on Equation (5), the anonymous set in our algorithm is smaller than in VW and the LBS provider would respond fewer messages with a smaller anonymous set input. As a result, the cost of our proposed scheme is lower than in VW. Obviously, the cost of the exhaustive method is the smallest as it provides the optimal solution. The cost of our scheme is about 10% bigger than the exhaustive method. Our proposed scheme has reduced the cost of LBS effectively.

Figure 9. LBS query cost with different query requirement $l_0$. ($k_0 = 10$, $\lambda_0 = 2$).

Figure 10 shows the impact of privacy parameter $l_0$ on the average time computation of generating an anonymous set. We can see that the time consumption would increase with the increase of $l_0$. When the privacy requirement $l_0$ increases, there will be more road sectors in the anonymous set. As a result, the time consumption to generate an anonymous set increases as it takes more time to decide which road sections is added to the anonymous set. The time consumption of our proposed scheme is lower than VW algorithm and the exhaustive method. The algorithm VW is based on the voronoi map of a road network, which consists of two stages. In the first stage, the road map is divided into serials of continuous polygons according to the user location. In the second stage, the anonymous set is generated. As a result, the algorithm VW takes a lot of time. As for the exhaustive method, it has to iterate through all the feasible solutions which satisfies the privacy protection requirements. Hence, it takes more time than our proposed scheme. Our scheme provides privacy protection within limited time, which is applicable in LBS.

In conclusion, based on the above analysis, our proposed scheme not only satisfies the demand of $k$-anonymity and $l$-diversity but also has the ability to resist side-weight inference attacks. Meanwhile, the cost of LBS and time for generating an anonymous set is small. The scheme provides location privacy for users at a low cost.
6. Conclusions and Future Work

In this paper, we investigate the issue of location privacy preservation for non-uniform distributed road network. According to the user’s distribution, the side-weight inference attack happens as the adversary would infer that the target user belongs to a high-density road section with high probability. In other words, the adversary would track a target user with a high probability, and location privacy is poorly protected. We propose a privacy-preserving method based on road truncation. The road section with high user density is designed to be truncated to reduce the divergence between different road sections. We also consider the cost function and formulate the object as a mixed integer programming problem. The simulation results show that our scheme not only achieves $k$-anonymity and $l$-diversity, but also resists the attack of side-weight inference. Meanwhile, the proposed scheme runs fast at a low cost. As a result, our scheme is proved to protect users’ location privacy effectively and efficiently for non-uniform distributed road network.

In the future, we are planning to study the problem of trajectory privacy in continuous LBS. Moreover, machine learning is widely used in many aspects and the Generative Adversarial Networks (GAN) is a promising tool to protect privacy.

**Author Contributions:** Conceptualization, Q.L. and H.W.; methodology, Q.L. and H.W.; software, Q.L.; validation, Q.L., H.W., X.W., N.Z. and L.D.; formal analysis, Q.L.; investigation, Q.L.; resources, Q.L.; data curation, Q.L.; writing—original draft preparation, Q.L.; writing—review and editing, H.W., X.W., N.Z. and L.D.; visualization, Q.L.; supervision, H.W.; project administration, H.W.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Key R&D Program of China (2018YFE0207600) and the Fundamental Research Funds for the Central Universities, China (2020JBZD005).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

The following abbreviations are used in this manuscript:

| Abbreviation | Description               |
|--------------|---------------------------|
| BS           | Base Station              |
| GAN          | Generative Adversarial Networks |
| GPS          | Global Positioning System |
| k-NN         | k-Nearest Neighbor        |
| LBS          | Location-based Service    |
| PoI          | Point of Interest         |
| QoS          | Quality of Service        |

References

1. Zhao, P.; Li, J.; Zeng, F.; Xiao, F.; Wang, C.; Jiang, H. ILLIA: Enabling k-Anonymity-Based Privacy Preserving Against Location Injection Attacks in Continuous LBS Queries. *IEEE Internet Things J.* 2018, 5, 1033–1042. [CrossRef]
2. Wang, S.; Hu, Q.; Sun, Y.; Huang, J. Privacy Preservation in Location-Based Services. *IEEE Commun. Mag.* 2018, 56, 134–140. [CrossRef]
3. Yassin, A.; Nasser, Y.; Awad, M.; Al-Dubai, A.; Liu, R.; Yuen, C.; Raulefs, R.; Aboutanios, E. Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications. *IEEE Commun. Surv. Tutor.* 2017, 19, 1327–1346. [CrossRef]
4. Lopes, H.; Pires, I.M.; Sánchez San Blas, H.; García-Ovejero, R.; Leithardt, V. PriADA: Management and Adaptation of Information Based on Data Privacy in Public Environments. *Computers* 2020, 9, 77. [CrossRef]
5. Obaidat, M.A.; Obeidat, S.; Holst, J.; Al Hayajneh, A.; Brown, J. A Comprehensive and Systematic Survey on the Internet of Things: Security and Privacy Challenges, Security Frameworks, Enabling Technologies, Threats, Vulnerabilities and Countermeasures. *Computers* 2020, 9, 44. [CrossRef]
6. Ma, C.; Yan, Z.; Chen, C.W. SSPA-LBS: Scalable and Social-Friendly Privacy-Aware Location-Based Services. *IEEE Trans. Multimed.* 2019, 21, 2146–2156. [CrossRef]
7. He, Y.; Ni, J.; Wang, X.; Niu, B.; Li, F.; Shen, X. Privacy-Preserving Partner Selection for Ride-Sharing Services. *IEEE Trans. Veh. Technol.* 2018, 67, 5994–6005. [CrossRef]
8. Albelaihy, A.; Cazalas, J. A survey of the current trends of privacy techniques employed in protecting the Location privacy of users in LBSs. In Proceedings of the 2017 2nd International Conference on Anti-Cyber Crimes (ICACC), Abha, Saudi Arabia, 26–27 March 2017; pp. 19–24. [CrossRef]
9. Tian, Z.; Wang, Y.; Sun, Y.; Qiu, J. Location Privacy Challenges in Mobile Edge Computing: Classification and Exploration. *IEEE Netw.* 2020, 34, 52–56. [CrossRef]
10. Zheng, Y.; Lu, R.; Guan, Y.; Shao, J.; Zhu, H. Efficient and Privacy-Preserving Similarity Range Query over Encrypted Time Series Data. *IEEE Trans. Dependable Secur. Comput.* 2021, 1. [CrossRef]
11. Wu, X.; Xu, X.; Bilal, M. Towards Privacy Protection Composition Framework on Internet of Vehicles. *IEEE Consum. Electron. Mag.* 2021, 1. [CrossRef]
12. Asuquo, P.; Cruickshank, H.; Morley, J.; Ogah, C.P.A.; Lei, A.; Hathal, B.; Bao, S.; Sun, Z. Security and Privacy in Location-Based Services for Vehicular and Mobile Communications: An Overview, Challenges, and Countermeasures. *IEEE Internet Things J.* 2018, 5, 4778–4802. [CrossRef]
13. Wang, J.; Cai, Z.; Yu, J. Achieving Personalized k-Anonymity-Based Content Privacy for Autonomous Vehicles in CPS. *IEEE Trans. Ind. Inform.* 2020, 16, 4424–4425. [CrossRef]
14. Ying, B.; Nayak, A. Anonymous and Lightweight Authentication for Secure Vehicular Networks. *IEEE Trans. Veh. Technol.* 2017, 66, 10626–10636. [CrossRef]
15. Srinivasan, R.; Mohan, A.; Srinivasan, P. Privacy conscious architecture for improving emergency response in smart cities. In Proceedings of the 2016 Smart City Security and Privacy Workshop (SCSP-W), Vienna, Austria, 11 April 2016; pp. 1–5. [CrossRef]
16. Xu, X.; Chen, H.; Xie, L. A Location Privacy Preservation Method Based on Dummy Locations in Internet of Vehicles. *Appl. Sci.* 2021, 11, 4594. [CrossRef]
17. Ni, J.; Zhang, K.; Lin, X.; Yang, H.; Shen, X.S. AMA: Anonymous mutual authentication with traceability in carpooling systems. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6. [CrossRef]
18. Du, M.; Wang, K.; Xia, Z.; Zhang, Y. Differential Privacy Preserving of Training Model in Wireless Big Data with Edge Computing. *IEEE Trans. Big Data* 2020, 6, 283–295. [CrossRef]
19. Zhou, L.; Yu, L.; Du, S.; Zhu, H.; Chen, C. Achieving Differentially Private Location Privacy in Edge-Assistant Connected Vehicles. *IEEE Internet Things J.* 2019, 6, 4472–4481. [CrossRef]
20. Wang, L.; Zhang, D.; Yang, D.; Lim, B.Y.; Han, X.; Ma, X. Sparse Mobile Crowdsensing with Differential and Distortion Location Privacy. *IEEE Trans. Inf. Forensics Secur.* 2020, 15, 2735–2749. [CrossRef]
21. Machanavajjhala, A.; Gehrke, J.; Kifer, D.; Venkitasubramaniam, M. L-diversity: Privacy beyond k-anonymity. In Proceedings of the 22nd International Conference on Data Engineering (ICDE’06), Atlanta, GA, USA, 3–7 April 2006; pp. 24. [CrossRef]
22. Lu, Z.; Qu, G.; Liu, Z. A Survey on Recent Advances in Vehicular Network Security, Trust, and Privacy. *IEEE Trans. Intell. Transp. Syst.* 2019, 20, 760–776. [CrossRef]

23. Gardner, Z.; Leibovici, D.; Basiri, A.; Foody, G. Trading-off location accuracy and service quality: Privacy concerns and user profiles. In Proceedings of the 2017 International Conference on Localization and GNSS (ICL-GNSS), Nottingham, UK, 27–29 June 2017; pp. 1–5. [CrossRef]

24. Hu, L.; Qian, Y.; Chen, M.; Hossain, M.S.; Muhammad, G. Proactive Cache-Based Location Privacy Preserving for Vehicle Networks. *IEEE Wirel. Commun.* 2018, 25, 77–83. [CrossRef]

25. Tang, J.; Zhu, H.; Lu, R.; Lin, X.; Li, H.; Wang, F. DLP: Achieve Customizable Location Privacy with Deceptive Dummy Techniques in LBS Applications. *IEEE Internet Things J.* 2021, 1. [CrossRef]

26. Yu, H.; Jia, X.; Zhang, H.; Yu, X.; Shu, J. PSRide: Privacy-Preserving Shared Ride Matching for Online Ride Hailing Systems. *IEEE Trans. Dependable Secur. Comput.* 2019, 18, 1425–1440. [CrossRef]

27. Han, S.; Lin, J.; Zhao, S.; Xu, G.; Ren, S.; He, D.; Wang, L.; Shi, L. Location Privacy-Preserving Distance Computation for Spatial Crowdsourcing. *IEEE Internet Things J.* 2020, 7, 7550–7563. [CrossRef]

28. Samarati, P. Protecting respondents identities in microdata release. *IEEE Trans. Knowl. Data Eng.* 2001, 13, 1010–1027. [CrossRef]

29. Gruteser, M.; Grunwald, D. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In Proceedings of the MobiSys ’03: The 1st International Conference on Mobile Systems, Applications and Services, San Francisco, CA, USA, 5–8 May 2003; Association for Computing Machinery: New York, NY, USA, 2003; pp. 31–42. [CrossRef]

30. Chow, C.Y.; Mokbel, M.; Bao, J.; Liu, X. Query-aware location anonymization for road networks. *GeoInformatica* 2011, 15, 571–607. [CrossRef]

31. Wang, T.; Liu, L. Privacy-Aware Mobile Services over Road Networks. *Proc. VLDB Endow.* 2009, 2, 1042–1053. [CrossRef]

32. Mouratidis, K.; Yiu, M.L. Anonymous Query Processing in Road Networks. *IEEE Trans. Knowl. Data Eng.* 2010, 22, 2–15. [CrossRef]

33. Rubner, Y.; Tomasi, C.; Guibas, L.J. The Earth Mover’s Distance as a Metric for Image Retrieval. *Int. J. Comput. Vision* 2000, 40, 99–121. [CrossRef]

34. Xue, J.; Liu, X.Y.; Yang, X.C.; Wang, B. A Location Privacy Preserving Approach on Road Network. *Chin. J. Comput.* 2011, 34, 865–878. [CrossRef]

35. Fan, X.; Tu, J.; Ye, C.; Zhou, F. The research for protecting location privacy based on V-W algorithm. *Eurasiap J. Wirel. Commun. Netw.* 2014, 2014, 202. [CrossRef]