entropy MDPI

Article

Successive Trajectory Privacy Protection with Semantics Prediction Differential Privacy

Jing Zhang [1,2,*], Yanzi Li [1,2], Qian Ding [1,2], Liwei Lin [1,2] and Xiucai Ye [3]

[1] School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China
[2] Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fuzhou 350118, China
[3] Department of Computer Science, University of Tsukuba, Tsukuba 305-8577, Japan
* Correspondence: jingzhang@fjut.edu.cn

Abstract: The publication of trajectory data provides critical information for various location-based services, and it is critical to publish trajectory data safely while ensuring its availability. Differential privacy is a promising privacy protection technology for publishing trajectory data securely. Most of the existing trajectory privacy protection schemes do not take into account the user's preference for location and the influence of semantic location. Besides, differential privacy for trajectory protection still has the problem of balance between the privacy budget and service quality. In this paper, a semantics- and prediction-based differential privacy protection scheme for trajectory data is proposed. Firstly, trajectory data are transformed into a prefix tree structure to ensure that they satisfy differential privacy. Secondly, considering the influence of semantic location on trajectory, semantic sensitivity combined with location check-in frequency is used to calculate the sensitivity of each position in the trajectory. The privacy level of the position is classified by setting thresholds. Moreover, the corresponding privacy budget is allocated according to the location privacy level. Finally, a Markov chain is used to predict the attack probability of each position in the trajectory. On this basis, the allocation of the privacy budget is further adjusted and its utilization rate is improved. Thus, the problem of the balance between the privacy budget and service quality is solved. Experimental results show that the proposed scheme is able to ensure data availability while protecting data privacy.

Keywords: trajectory publishing; differential privacy; prediction; sensitivity; Markov chain

check for updates

Citation: Zhang, J.; Li, Y.; Ding, Q.; Lin, L.; Ye, X. Successive Trajectory Privacy Protection with Semantics Prediction Differential Privacy. *Entropy* **2022**, *24*, 1172. https://doi.org/10.3390/e24091172

Academic Editors: Alexandre G. Evsukoff and Yilun Shang

Received: 15 July 2022
Accepted: 18 August 2022
Published: 23 August 2022

1. Introduction

Location-based service (LBS) has become increasingly popular in people's daily lives due to the proliferation of mobile devices [1]. At present, LBS has covered all aspects of national economy and social life, such as navigation, query and recommendation of interest points, takeout, check-in, social networking [2], etc. Moreover, the implementation of LBS depends on published trajectory data [3]. However, when releasing trajectory data, there is a probability of being attacked by attackers, resulting in the disclosure of users' trajectory information. The disclosure of trajectory information may lead to the exposure of more personal privacy information, so trajectory privacy has become one of the most important privacies of people.

Traditional trajectory privacy protection technologies include K-anonymity, encryption and differential privacy [4]. The K-anonymity model and its derivative model provide a means of quantitative evaluation, which makes different types of schemes comparable, but cannot provide strict mathematical proof [5]. Meanwhile, the security depends on the background knowledge grasped by the attacker. In addition, cryptography-based privacy protection methods can provide strict protection on data confidentiality, but their disadvantages and challenges lie in weak scalability and low implementation efficiency [6]. This is mainly because the current homomorphic encryption mechanisms inevitably have large computational complexity overhead. The emergence of differential privacy technology makes up for the above problems effectively. It hides sensitive raw data by attaching a noise value that obeys a certain distribution to the raw data. On the one hand, the differential privacy model makes the maximum assumption about the attacker's ability, and does not depend on the background knowledge the attacker has mastered [7]. On the other hand, the differential privacy model is built on a solid mathematical basis, and gives a quantitative model of the degree of privacy leakage, which is simple to implement and efficient to calculate. However, existing studies on trajectory differential privacy protection still have problems in the following three aspects:

(1) The existing trajectory privacy protection mechanism does not take into account the problem of excessive overhead of real-time sensitivity calculation. It is difficult to obtain accurate sensitivity of each position in the trajectory, although the amount of calculation is reduced offline.
(2) The impact of semantic location on trajectory is not considered in the previous scheme. Semantic location is likely to increase the risk of user privacy information disclosure. For example, users' preferences and economic level can be inferred according to the frequency of users' access to certain semantic location points.
(3) In the publishing process of the differential privacy trajectory data set, the allocation of the privacy budget is one of the key factors determining the final amount of noise added. If the privacy budget is not allocated properly, it can cause serious waste and add too much overall noise. However, the current method of privacy budget allocation still stays at average allocation or simple balance allocation, and there is still a certain degree of waste. How to design a more reasonable way of privacy budget allocation according to the characteristics of trajectory data sets is still lacking in relevant research.

If the sensitivity can only be calculated in real time, the calculation cost is too large, which will increase the time cost of the scheme and reduce the operation efficiency. In this paper, a sensitivity map is defined so that the sensitivity of each position point of trajectory can be queried offline. If the impact of semantic location is not taken into account, it is likely to increase the risk of privacy leakage. For example, a user's trajectory is between home and school every day. School is a special semantic location. After acquiring the user's trajectory, the attacker can infer his occupation or even economic status easily. This paper takes into account the impact of semantic location on user location sensitivity to improve the privacy protection effect. In addition, if the allocation of privacy budget is not reasonable, the added noise will be too large or too small. This can result in reduced data availability or insufficient privacy. Therefore, the allocation method of privacy budget is improved in this paper. A semantics- and prediction-based differential privacy protection scheme for trajectory data (SPDP) is proposed in this paper. The contributions are summarized as follows:

(1) A sensitivity map is defined so that the sensitivity of the current position can be accurately confirmed even offline. Thus, the computational overhead is reduced and the operating efficiency of this scheme is improved. The differentiation protection mechanism of location privacy based on a sensitivity map is designed. By allowing users to customize the sensitivity of semantic locations, the privacy budget can be tailored to further improve its utilization.
(2) The differentiation protection mechanism of location privacy based on semantic location is designed. Considering the influence of semantic location sensitivity, sensitivity is determined by the number of trajectories containing the node prefix and semantic sensitivity. The privacy levels are divided according to the location sensitivity. Then the sensitivity ratio and privacy levels are used to allocate the privacy budget of each location to further improve its utilization.
(3) A privacy budget adjustment algorithm based on a Markov chain is proposed. After the privacy budget is allocated based on sensitivity and privacy level, the attack probability of the nodes in the prefix tree is calculated by using the property of the Markov process. Then, the sensitivity and privacy level are adjusted by attack

probability, so as to adjust the allocation of privacy budget and make the allocation of privacy budget more reasonable.

The rest of the article is organized as follows: the related work is given in Section 2; the preliminaries are given in Section 3; the privacy protection method is designed in Section 4; the simulation analysis is discussed in Section 5; and finally, the conclusion is given in Section 6.

2. Related Work

The relevant technologies involved in this paper include trajectory differential privacy protection [8–10] and location recommendation mechanism [11–13]. Therefore, the typical methods of trajectory differential privacy protection and location recommendation mechanism are analyzed, respectively.

Due to the gradual increase of location service applications, the research on privacy protection of location trajectory data has become a hot research topic. In recent years, the differential privacy model based on false data technology has been rapidly applied to protect the privacy of data release after being proposed. This model realizes privacy protection by adding noise to real data sets [14]. In data release, differential privacy realizes different privacy protection degrees and data release accuracy by adjusting privacy parameter ε. Generally speaking, the higher the value of ε is, the lower the degree of privacy protection is, and the higher the accuracy of published data sets is. Differential privacy is mainly realized through a noise mechanism. The first universal differential privacy mechanism is the Laplace mechanism proposed in [15], which is mainly aimed at numerical query. For non-numerical queries, the exponent mechanism is proposed in [16], which is the second universal mechanism to realize differential privacy.

In the privacy protection of trajectory data set release, the prefix method based on the differential privacy model is proposed for the first time in [17]. This method uses a hierarchical framework to construct a prefix tree, divides the trajectories with the same prefix into the same branch of the tree, and realizes differential privacy by adding noise to the node count. However, as the tree grows, the prefix will form a large number of leaf nodes, resulting in too much noise and reducing the accuracy of the published data set. Later, on the basis of prefix method, location trajectory and check-in frequency are used to set thresholds in [18], so as to classify the level of location sensitivity. Then, the corresponding privacy budget is allocated according to the sensitivity, which makes the allocation of privacy budget more reasonable and reduces the amount of noise data.

The work [19] proposes the method of merging similar trajectories. By dividing the trajectory coverage area into grids, the trajectory position points falling into the same grid are represented by the center points of the grid, thus improving the counting value of position points greatly. In [20], the regional division is improved by adopting a multi-level grid model to divide position points at different speeds in the trajectory according to different granularity, so as to maintain the original sequence information of the trajectory to the maximum extent. However, these methods have the problem of low data availability due to excessive information loss rate, and fail to fully consider the semantic location information of users, resulting in semantic inference attacks [21], which leads to the disclosure of users' sensitive privacy.

Published trajectory data can be used in various location services. Location recommendation service in LBS is frequently used. For example, Nur [22] presents a new problem of user identification of top-K social space co-participation location selection (SSLS) in social graphs. Two exact solutions and two approximate solutions are developed to solve this NP-hard problem. Thus, the best set of K positions can be selected for the user from a large number of candidate positions. Location recommendation methods can be divided into three categories generally: content-based recommendation system, collaborative filtering recommendation and mixed recommendation [23]. A content-based recommendation system mainly selects items with high similarity to them as recommendations according to the items users like. Collaborative filtering technology determines a group of recommender users with similar behaviors according to the evaluation behavior of the target users, and takes the evaluation of the recommender users on the project as the recommendation value of the target users. Mixed recommendation is mainly to solve the deficiency of single recommendation technology. Different recommendation technologies can be combined according to different mixing strategies to complete the recommendation.

Lian first proposes a collaborative filtering algorithm based on implicit feedback and content perception [24], which gives a lower preference value to the locations that users have not visited, and a higher preference value to the locations that users have visited according to their historical access frequency. Then, Lian combines the matrix factor decomposition method and puts forward the improved schemes Geo MF [25] and Geo MF++ [26], which improve the accuracy of the recommendation system effectively. In recent years, with the development of deep learning theory, neural network technology has also been used to solve the problem of location recommendation [27–29]. Shyamali [30] proposes the fault tolerance technology of the relevant sensitive random logic circuit to reduce the system error. Lalli [31] reduces operational risk by training four neural networks to detect and handle errors before they cause harm. However, the technology needs a lot of data support. In addition, the above recommendation schemes only focus on the recommendation effect and ignore the user's privacy and security issues. The lack of protection of trajectory data may cause the disclosure of user's privacy information easily.

The existing studies on trajectory differential privacy protection do not take into account the impact of semantic features on trajectory, and the privacy budget allocation is not precise enough. In addition, the existing location recommendation mechanisms ignore the privacy protection of user data. Therefore, a semantics- and prediction-based differential privacy protection scheme for trajectory data is proposed in this paper. The semantic sensitivity and the Markov technology are introduced to improve the utilization rate of the privacy budget. Meanwhile, the location recommendation mechanism is combined with the differential privacy technology to protect the security of the trajectory data while ensuring the location recommendation effect.

3. Preliminaries

The system model of semantics and prediction based differential privacy protection scheme for trajectory data (SPDP) is presented in Figure 1. The system model consists of three parts: mobile smart device, privacy server and location server. Among them, privacy server is a trusted third party anonymous server. This paper focuses on the privacy protection of the system, so it ignores the details of the internal network connection. The location information to be protected is the trajectory data published by the mobile smart device, including the user check-in time, location identification (ID), longitude and latitude. These assumptions are used in most previous works, such as [7,17,18]. In addition, the SPDP scheme proposed in this paper uses differential privacy technology, prefix tree structure, Markov chain and so on to protect the trajectory data. Therefore, the definitions of related concepts are quoted and designed. The detailed definitions involved are shown below.

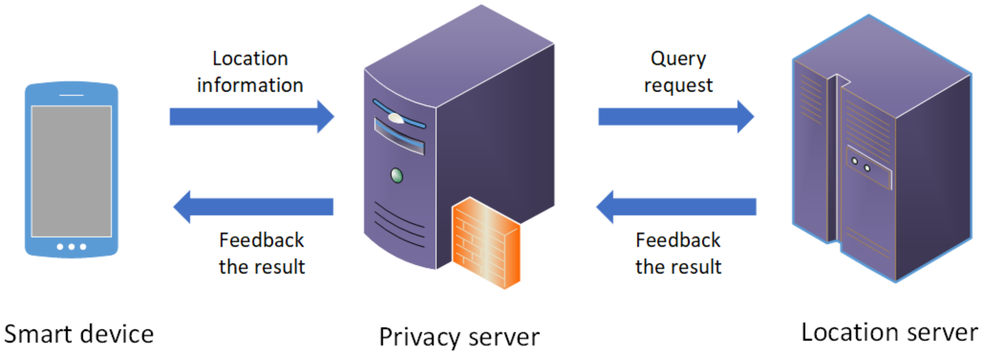

Figure 1. The system model of SPDP.

Definition 1. ***ε-Differential Privacy*** *[14]. Given a query algorithm* $M : D \to R^d$ *that supports a random mechanism, if for any data set D and its adjacent data set D′, algorithm M satisfies Formula (1) for any output O, then the random algorithm M satisfies ε-differential privacy.*

$$\Pr(M(D) \in O) \leq e^{\varepsilon} \cdot \Pr\left(M\left(D'\right) \in O\right) \tag{1}$$

There is only one record difference between adjacent data sets, that is, $\|D - D'\|_1 = 1$. *ε is the privacy budget, which determines the degree of privacy protection and the accuracy of released data sets. The lower the privacy budget is, the closer the probability ratio of algorithm M outputting the same result on D and* D' *is to 1, and the higher the degree of privacy protection is, the lower the accuracy of the corresponding published data set is. When ε = 0, M will output the result with the same probability distribution on D and* D'*, and the degree of privacy protection will reach the highest at this moment, but the published data will not reflect any useful information.*

Definition 2. ***Global Sensitivity*** *[16]. For any query function* $f : D \to R^d$ *, the global sensitivity of f is*

$$\Delta f = max_{D,D'} \|f(D) - f\left(D'\right)\|_1 \tag{2}$$

Global sensitivity is the maximum range of output value variation of a particular query function f on all possible adjacent datasets D and D'*, and its measure is theL1 distance between the two.*

Definition 3. ***Laplacian Mechanism*** *[8]. For any function f on data set D, if the output result of function f satisfies Equation (3), then the random algorithm M satisfies ε-differential privacy.*

$$M(D) = f(D) + \mathrm{Lap}\left(\frac{\Delta f}{\varepsilon}\right)^d \tag{3}$$

where, Δf *is the sensitivity of the query function. The location parameter of the Laplace distribution is 0, and the scale parameter is* $\frac{\Delta f}{\varepsilon}$.

Definition 4. ***Trajectory Prefix*** *[17]. A trajectory* $S = s_1 \to s_2 \to \cdots \to s_{|S|}$ *is a prefix of a trajectory* $T = t_1 \to t_2 \to \cdots \to t_{|T|}$ *, denoted by* $S \preccurlyeq T$*, if and only if* $|S| \leq |T|$ *and* $\forall 1 \leq i \leq |S|$, $s_i = t_i$.

For example, a trajectory and the corresponding trajectory sequence of user u are shown in Figure 2 and Table 1. For trajectory 1: $l_1 \to l_2 \to l_3 \to l_4$, it can be seen that $l_1, l_1 \to l_2$, $l_1 \to l_2 \to l_3$ and $l_1 \to l_2 \to l_3 \to l_4$ are their prefixes, but $l_2 \to l_3$ is not a prefix.

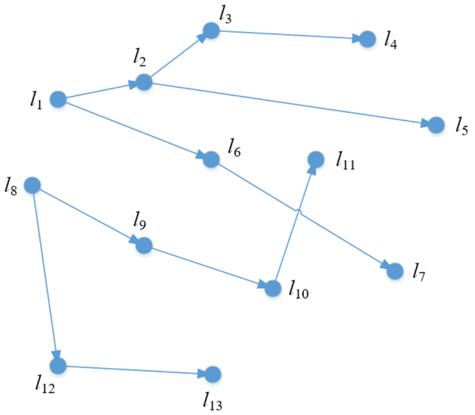

Figure 2. One trajectory of user u.

Table 1. The trajectory sequence of user u.

The Serial Number	The Trajectory	The Serial Number	The Trajectory
1	$l_1 \to l_2 \to l_3 \to l_4$	4	$l_8 \to l_9 \to l_{10} \to l_{11}$
2	$l_1 \to l_2 \to l_5$	5	$l_8 \to l_{12} \to l_{13}$
3	$l_1 \to l_6 \to l_7$	6	$l_{12} \to l_{13}$

Definition 5. ***Prefix Tree*** *[17]. A prefix tree TT of a trajectory database D is a triplet $TT = (V, E, Root(TT))$, where V is the set of nodes labeled with locations, each corresponding to a unique trajectory prefix in D; E is the set of edges, representing transitions between nodes; $Root(TT) \in V$ is the virtual root of TT. The unique trajectory prefix represented by a node $v \in V$, denoted by prefix (v, TT), is an ordered list of locations starting from $Root(TT)$ to v.*

Each node $v \in V$ of TT keeps a doublet in the form of (S_i, pl_i), where S_i is the location sensitivity, and pl_i is the privacy level of the location. Figure 3 illustrates the prefix tree of the sample database in Table 1, where each node v is labeled with its location, sensitivity and privacy level.

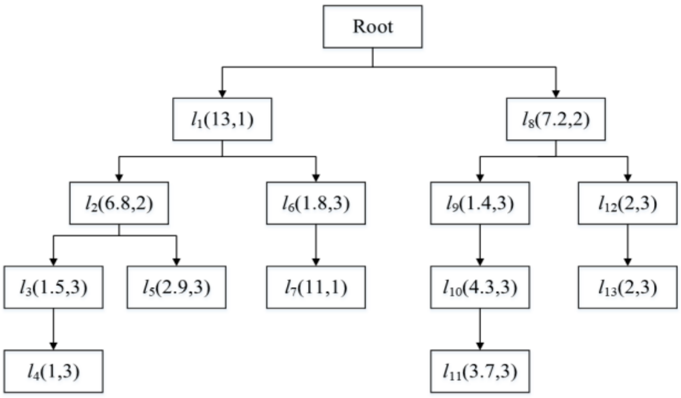

Figure 3. The prefix tree structure of user u.

Definition 6. ***Semantic location****. Semantic location refers to the location that conforms to the characteristics of semantic location type, denoted as SL. In this paper, semantic location types are divided into 10 categories according to geographic tags, including science, education and culture, catering, leisure and entertainment, medical care and so on. Semantic locations can be obtained from map information. Each semantic location has a certain semantic sensitivity, and will affect the location sensitivity within a certain range. Therefore, each location l_i has a certain semantic sensitivity, denoted as Sem_i.*

Definition 7. ***Sensitivity****. The check-in times of user u at location l_i can indicate the user's preference for this location. It is assumed that the more times users check in, the higher the preference degree of users for this location. Attackers can easily master users' preferences by calculating check-in statistics of specific locations, so users' privacy is vulnerable to leakage. To solve this problem, this paper defines the sensitivity of the user's check-in location: $S_i = \alpha_i + Sem_i$. Where, α_i represents the check-in times of user u at position l_i, and Sem_i represents the semantic sensitivity of user at position l_i. As shown in Figure 3 and Table 2. The user's check-in times α_i and semantic sensitivity Sem_i are combined as the location sensitivity S_i of node l_i. The more times a user checks in to a location, the more sensitive that location is.*

Definition 8. ***Privacy Level****. The location privacy level is defined as $pl = r$ $(r = 1, 2, \cdots, n)$ in this paper. It is determined by the sensitivity of user u to the location. Three privacy levels are set in this paper, namely insensitive, normal and sensitive. Then the thresholds are set for position sensitivity. When sensitivity reaches the thresholds, the privacy level of this location changes.*

Table 2. The position sensitivity of user u.

Location	S_i	α_i	Sem_i
l_1	13	3	10
l_2	6.8	2	4.8
l_3	1.5	1	0.5
l_4	1	1	0
l_5	2.9	1	1.9
l_6	1.8	1	0.8
l_7	11	1	10
l_8	6.2	2	4.2
l_9	1.4	1	0.4
l_{10}	4.3	1	3.3
l_{11}	3.7	1	2.7
l_{12}	2	2	0
l_{13}	2	2	0

It is defined as the highest privacy level when $pl = 1$ in this paper. As sensitivity increases, the privacy level of a location decreases. In other words, the position is most sensitive when $pl = 1$, and the position is less sensitive when the value of pl is larger. If pl is small, the sensitivity of the position is relatively high, that is, the more sensitive position l_i is, the less weight it will have. Therefore, the privacy budget allocated to location l_i is small. In differential privacy protection, the smaller the privacy budget allocated to position l_i, the greater the added noise and the higher the privacy protection intensity.

For example, divide the privacy level for the trajectory example of user u shown in Figure 2 and Table 1. Suppose the position is least sensitive when the assumed sensitivity is less than 5. Then assume that the threshold interval is 5, and when the sensitivity exceeds the threshold, the privacy level will change. When the sensitivity exceeds 10, the privacy level is the highest and the location is the most sensitive. Accordingly, the privacy level division of the trajectory example is obtained, as shown in Table 3.

Table 3. The privacy level and sensitivity of user u.

Location	S_i	α_i	Sem_i	pl_i
l_1	13	3	10	1
l_2	6.8	2	4.8	2
l_3	1.5	1	0.5	3
l_4	1	1	0	3
l_5	2.9	1	1.9	3
l_6	1.8	1	0.8	3
l_7	11	1	10	1
l_8	6.2	2	4.2	2
l_9	1.4	1	0.4	3
l_{10}	4.3	1	3.3	3
l_{11}	3.7	1	2.7	3
l_{12}	2	2	0	3
l_{13}	2	2	0	3

Definition 9. ***Markov Process*** *[32]. Assume that the time parameter set of random process* $X = \{X_t, t \in T\}$ *is* $T = \{0, 1, \cdots\}$ *and the state space* E *is discrete,* $E = \{i_0, i_1, i_2, \cdots\}$*. For any* $t \in R$, $i_0, i_1, i_2, \cdots \in E$*, then:*

$$P(X_t = i_t | X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \cdots, X_0 = i_0) = P(X_t = i_t | X_{t-1} = i_{t-1}) \tag{4}$$

If the random process X *satisfies Equation (4), the random process is a Markov process. Where,* $\{X_t = i\}$ *represents the state of random process* X *at time* t *is* i*. The property of Markov processes is that the future state is only related to the present state, not to the past state.*

4. Semantics- and Prediction-Based Differential Privacy Protection Scheme for Trajectory Data (SPDP)

The specific process of the semantics- and prediction-based differential privacy protection scheme for trajectory data (SPDP) proposed in this paper is shown in Figure 4.

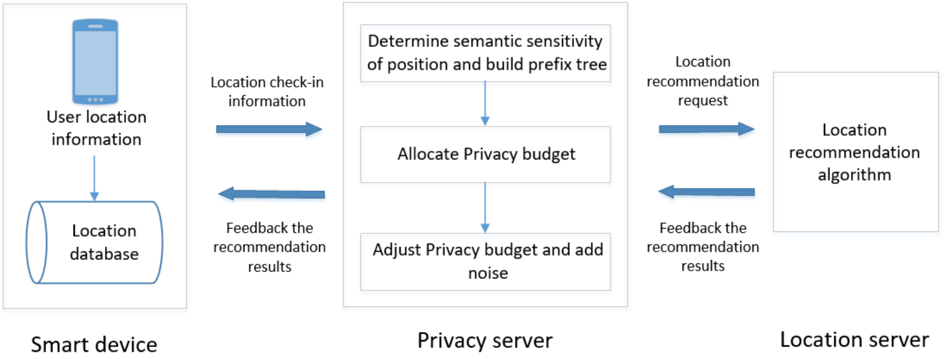
User location information
Location database
Smart device
Location check-in information
Feedback the recommendation results
Determine semantic sensitivity of position and build prefix tree
Allocate Privacy budget
Adjust Privacy budget and add noise
Privacy server
Location recommendation request
Feedback the recommendation results
Location recommendation algorithm
Location server

Figure 4. The framework of SPDP.

Step 1. Sensitivity processing based on semantic location: Allocate different privacy budgets for different semantic locations, determine the semantic sensitivity of the location through the generated semantic sensitivity map, and obtain the location sensitivity and privacy level by combining the check-in times of the location.

Step 2. Privacy budget allocation based on prefix tree: A single location satisfying ε-differential privacy cannot ensure trajectory privacy security. Therefore, the user trajectory is transformed into a prefix tree structure to ensure that the trajectory meets ε-differential privacy, and the privacy budget is allocated according to the sensitivity of the location.

Step 3. Privacy budget adjustment based on Markov chain: The attack probability of the location is predicted by a Markov chain, and the allocated privacy budget is adjusted according to the attack probability to further improve its utilization rate.

Step 4. Location recommendation under differential privacy protection: Add corresponding noise to the location, and reflect the validity and availability of trajectory data under differential privacy protection through the recommendation effect of location recommendation service.

4.1. Sensitivity Processing Based on Semantic Location

Not only semantic locations directly connected to sensitive locations are sensitive. From the perspective of random disturbance distribution, those semantic locations close to sensitive locations still have the risk of exposing sensitive locations even if they are not connected to sensitive locations directly. Therefore, certain semantic sensitivity should also be assigned. This paper considers the global connectivity between location points and radiates the semantic sensitivity of semantically sensitive locations to nearby nodes according to the distance and access degree.

As shown in Figure 5, the semantic location node set $\mathcal{A}$ with a privacy level near any location l_i is first obtained. Then, the map is transformed into an undirected graph. According to the distance and access degree, the equivalent distance between any location l_i and semantic location SL_j is $D_{ij} = d_{SL_j}(c_j - 1)$. Where, d_{SL_j} is the Euclidean distance between l_i and SL_j, and c_j is the number of nodes traversed by the shortest path between the two nodes. Finally, the semantic sensitivity of semantic location radiation in $\mathcal{A}$ of any location l_i is obtained, as shown in Equation (5).

$$Sem_i = \sum_{SL_j \in \mathcal{A}} \frac{r - D_{ij}}{r} \cdot Sem_{SL_j} \tag{5}$$

where, Sem_i represents the semantic sensitivity of location l_i. $\mathcal{A} = \left\{ SL \middle| d_{SL_j} < r \right\}$, r indicates the threshold set by the user.

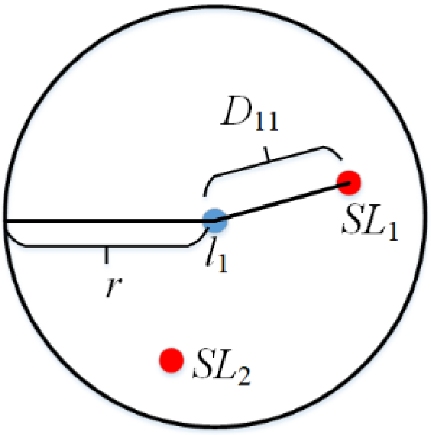

Figure 5. Semantic sensitivity of position l_i.

For the convenience of calculation, we use this paper grid map. Then, the semantic sensitivity of each region in the map is calculated using the above process, and the semantic sensitivity map map_{sen} is generated.

In Algorithm 1, the check-in times α_i and semantic sensitivity Sem_i of each node point in data set T are calculated firstly, and the two are combined as the sensitivity S_i of node (1–6 lines of Algorithm 1). Lines 7–12 of Algorithm 1 divide privacy levels according to node sensitivity. Based on the experimental data, this paper divides the privacy level into three categories. When $pl = 1$, the position is the most sensitive, when $pl = 2$, it is classified as normal, and when $pl = 3$, it is classified as insensitive. If the sensitivity of the node is less than 10, the privacy level is set to level 3. If the sensitivity is between 10 and 50, the privacy level is set to level 2. If sensitivity is greater than or equal to 50, the privacy level is set to level 1. Finally, a prefix tree is constructed and the sensitivity map map_{sen} is generated according to the sensitivity and privacy level of nodes (13–15 lines of Algorithm 1).

4.2. Privacy Budget Allocation Based on Prefix Tree

Because the root node in the prefix tree is not the actual check-in location, the root node does not consume the privacy budget. The privacy budget allocation scheme in this paper is mainly divided into two steps: the privacy budget allocation of each trajectory subsequence and the privacy budget allocation of each child node on the trajectory subsequence. Firstly, the average sensitivity of each trajectory subsequence is calculated to calculate the access probability of each subsequence. Then, the privacy budget is assigned to the trajectory subsequence according to the access probability. Since the higher the access probability, the higher the sensitivity, the allocated privacy budget should be inversely proportional to the access probability. Secondly, the privacy budget is allocated to each node according to the proportion of each node's privacy level in the sum of the privacy level of each trajectory subsequence. Finally, because part of the location points appear in multiple trajectory subsequences, the repeated privacy budget is merged. The privacy budget allocation algorithm based on location sensitivity is shown as follows:

In Algorithm 2, the privacy budget (lines 1–4 of Algorithm 2) is first assigned to the trajectory subsequence. The average sensitivity of each trajectory in dataset T is calculated. Then, the access probability of each trajectory is calculated according to the proportion of sensitivity, and the privacy budget is allocated according to the inverse relationship between the access probability and the privacy budget. In lines 5–7 of Algorithm 2, the privacy budget is allocated to each location in the trajectory according to the location's privacy level, and finally, the privacy budget of the location in multiple trajectories is combined.

Algorithm 1: Sensitivity Processing Algorithm Based on Semantic Location

```
Input: User check-in location data set T
Output: Sensitivity map $map_{sen}(l_i, S_i, pl_i)$, prefix tree TT
begin
$l_i \leftarrow T,\ S_i \leftarrow \varnothing,\ pl_i \leftarrow \varnothing$;
for every position $l_i$ in T do
    $\alpha_i \leftarrow$ T;
    $\mathcal{A} \leftarrow \left\{ SL_j \middle| d_{SL_j} < r \right\}$;
    $Sem_i \leftarrow \sum_{SL_j \in \mathcal{A}} \frac{r - D_{ij}}{r} \cdot Sem_{SL_j}$;
    $S_i \leftarrow \alpha_i + Sem_i$;
    if $S_i < 10$
        $pl_i = 3$;
    else $10 \leq S_i < 50$
        $pl_i = 2$;
        else $S_i \geq 50$
        $pl_i$ =1;
        TT $\leftarrow l_i(S_i, pl_i)$;
    end for
return $map_{sen}(l_i, S_i, pl_i)$, TT
end
```

Algorithm 2: Privacy Budget Allocation Algorithm Based on Sensitivity

```
Input: Privacy budget $\varepsilon$, prefix tree TT
Output: Trajectory set TB after allocating privacy budget
Begin
for every trajectory $T_i$ in TT do;
    $S_{T_i} \leftarrow \frac{\sum_{j=1}^{M} S_{lj}}{M_i}$;
    $P_{T_i} \leftarrow \frac{S_{T_i}}{S_{T_1} + \cdots + S_{T_N}}$;
    $\varepsilon_{T_i} \leftarrow \varepsilon \cdot \frac{\frac{1}{P_{T_i}}}{\sum_{i=1}^{N} \frac{1}{P_{T_i}}}$;
    for every position $l_j$ in $T_i$ do;
        $\varepsilon_{l_{ij}} \leftarrow \varepsilon_{T_i} \cdot \frac{p_{l_j}}{\sum_{j=1}^{M} p_{l_j}}$;
    $\varepsilon_j \leftarrow \sum_{i=1}^{N} \varepsilon_{l_{ij}}$;
    end for
end for
return TB
end
```

4.3. Privacy Budget Adjustment Based on Markov Chain

A trajectory consists of a series of position points that are continuous. The property of the Markov chain corresponds to the trajectory, that the next position depends only on the previous position. The two most important components of the Markov chain are the initial state probability distribution and state transition matrix.

Assume that the possible location set generated by the user at the moment is $L^{(t-1)} = l_1^{(t-1)}, l_2^{(t-1)}, \cdots, l_m^{(t-1)}$, and its probability value is $P^{(t-1)} = p_1^{(t-1)}, p_2^{(t-1)}, \cdots, p_m^{(t-1)}$. That is the initial state probability distribution. Suppose there are n possible positions for a user's trajectory, namely $l_1, l_2, \ldots, l_n$. The state transition probability from position l_i to position l_j is denoted as $P\left(l_i \rightarrow l_j\right)$, then matrix P is formed, which is called state transition probability matrix.

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \tag{6}$$

Then, the state transition probability matrix is used to calculate the possible position at time t as $L^{(t)} = l_1^{(t)}, l_2^{(t)}, \cdots, l_m^{(t)}$, and its probability value is $P^{(t)} = p_1^{(t)}, p_2^{(t)}, \cdots, p_m^{(t)}$, where $P^{(t)} = P^{(t-1)}P$, is the attack probability of the possible position at time t.

Assume that an attacker's attack starts at the initial position of the trajectory and continues in the direction of the trajectory. The property of Markov process is used to calculate the attack probability of nodes in the prefix tree, and the sensitivity is adjusted by calculating the probability, so as to adjust the allocated privacy budget. The privacy budget adjustment algorithm based on Markov is shown in Algorithm 3.

In Algorithm 3, the access probability of each trajectory is firstly calculated, and then the access probability of each position in the trajectory is calculated as the initial probability state distribution (lines 1–8 of Algorithm 3). Then, the state transition matrix is calculated according to the proportion of check-in times in the data set, so as to obtain the attack probability at time t (lines 9–11 of Algorithm 3). Finally, sensitivity and privacy level are adjusted linearly according to the attack probability, so as to adjust the privacy budget (lines 12–15 of Algorithm 3).

Algorithm 3: Privacy Budget Adjustment Algorithm Based on Markov

```
Input: Trajectory set TB after allocating privacy budget
Output: Trajectory set TC after adjusting privacy budget
Begin
1: for every trajectory $T_i$ in TB do;
2:     $S_{T_i} \leftarrow \frac{\sum_{j=1}^{M} S_{lj}}{M_i}$;
3:     $P_{T_i} \leftarrow \frac{S_{T_i}}{S_{T_1}+\cdots+S_{T_N}}$;
4:     for every position $l_j$ in $T_i$ do;
5:         $p_{lj} \leftarrow P_{T_i} \cdot \frac{S_{lj}}{\sum_{j=1}^{M} S_{lj}}$;
6:         $P^{(t-1)} \leftarrow p_{lj}$;
7:     end for
8: end for
9: $p_{ij} \leftarrow$ TB;
10: $P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}$;
11: $P^t = P^{(t-1)} . P$;
12: for every position $l_j$ in TB do
13:     $S'_{l_j} \leftarrow S_{l_j} + 10 \times P^t$;
14:     $pl'_{l_j} \leftarrow S'_{l_j}$;
15:     $\varepsilon'_{l_j} \leftarrow pl'_{l_j}$;
16: end for
17: return TC
end
```

4.4. Location Recommendation under Differential Privacy Protection

Through the previous three sections, the privacy budget assigned by the user for each location is available. Then, the Laplace mechanism is used to add the corresponding noise to the sensitivity of the position to change the privacy level of the position in this paper.

As the location privacy level changes, it is difficult for an attacker to discover a user's true preference for the location.

After the location privacy level is changed, the interest score of user u on location l is calculated by Equation (7). Where, $S'_{u,l}$ and w'_{score} represent the position sensitivity and position score weight after adding noise, respectively:

$$IG_{u,l} = S'_{u,l} \times w'_{score} \tag{7}$$

Position l is most sensitive when pl is minimal. However, when the location score is calculated, the weight of the location will increase as the privacy level of the location increases. Therefore, Equation (8) is used in this paper to calculate the score weight of the position.

$$w_{score} = \frac{pl_{n-r+1}}{\sum_r^n pl_r} \tag{8}$$

Since location sensitivity is used as location score directly, the score difference between locations will be too large, affecting the accuracy of the results. Therefore, $IG_{u,l}$ is normalized to obtain the normalized location score $IGN_{u,l}$, and then the scoring matrix $Matrix_{IGN}$ of users and locations is constructed, $IGN_{u,l}$ is shown as follows:

$$IGN_{u,l} = \frac{IG_{u,l} - min(IG_{u,l})}{max(IG_{u,l}) - IG_{u,l}} \tag{9}$$

After obtaining score matrix $Matrix_{IGN}$, the Pearson correlation coefficient is used to calculate users' similarity $sim(u,v)$, and user similarity matrix $Matrix_{sim}$ is constructed, where $sim(u,v)$ represented the similarity between user u and user v.

$$sim(u,v) = \frac{\sum_{l \in l(u,v)} \left(IGN_{u,l} - \overline{IGN_{u,l}}\right)\left(IGN_{v,l} - \overline{IGN_{u,l}}\right)}{\sqrt{\sum_{l \in u} \left(IGN_{u,l} - \overline{IGN_{u,l}}\right)^2}\sqrt{\sum_{l \in v} \left(IGN_{v,l} - \overline{IGN_{u,l}}\right)^2}} \tag{10}$$

where, $l(u,v)$ represents the common check-in location set of user u and user v, and $\overline{IGN_{u,l}}$ represents the average location score of user u. Finally, according to the user similarity matrix $Matrix_{sim}$, n users with the highest similarity to the target user are regarded as similar users. In addition, the locations of similar users are set and the locations not visited by target users are arranged in descending order of score, and the first n locations are recommended to target users. The location recommendation algorithm is as follows.

Assume that an attacker's attack starts at the initial position of the trajectory and continues in the direction of the trajectory. The property of the Markov process is used to calculate the attack probability of nodes in the prefix tree, and the sensitivity is adjusted by calculating the probability, so as to adjust the allocated privacy budget. The privacy budget adjustment algorithm based on Markov is shown in Algorithm 4.

In Algorithm 4, noise is first added to the sensitivity and privacy level of the location (lines 1–3 of Algorithm 4). Then score weight and interest score are calculated and normalized (lines 5–6 of Algorithm 4). Lines 7–9 of Algorithm 4 calculate the similarity between users and take the first n users with the highest similarity. Finally, the position with the highest interest score among the first n locations that are not visited by similar users is selected for recommendation (lines 10–15 of Algorithm 4).

Algorithm 4: Location Recommendation Algorithm

```
Input: Trajectory set TC after adjusting privacy budget
Output: Location recommendation set LR
Begin
for every position $l_i$ in TC do;
    $S'_{u,l} \leftarrow S_{u,l} + \mathrm{Lap}\left(\frac{\Delta f}{\varepsilon_i}\right)^d$;
    $pl'_r \leftarrow S'_{u,l}$;
    $w_{score}' \leftarrow \frac{pl_{n-r+1}}{\sum_r^n pl_r}$;
    $IG_{u,l} \leftarrow S'_{u,l} \times w'_{score}$;
    $IGN_{u,l} \leftarrow \frac{IG_{u,l}-min(IG_{u,l})}{max(IG_{u,l})-IG_{u,l}}$;\\ Normalize for $IG_{u,l}$
    for every user $v_j$ in TC do
        $sim\left(u,v_j\right) \leftarrow \frac{\sum_{l\in l(u,v_j)}\left(IGN_{u,l}-\overline{IGN_{u,l}}\right)\left(IGN_{v_j,l}-\overline{IGN_{u,l}}\right)}{\sqrt{\sum_{l\in u}\left(IGN_{u,l}-\overline{IGN_{u,l}}\right)^2}\sqrt{\sum_{l\in v_j}\left(IGN_{v_j,l}-\overline{IGN_{u,l}}\right)^2}}$;
            Arrange $sim\left(u,v_j\right)$ in descending order, take the top-$n$ users in $v_j$;
        for the $l_k$ in top-$n$ users that are not accessed by the target user do
            Arrange $IGN_{v_j,l_k}$ in descending order, take the top-$n$ locations in $l_k$
        end for
    end for
end for
return top-$n$ $l_k$
end
```

5. Performance Evaluations

5.1. Experimental Environment

In this section, the scheme in this paper is simulated and analyzed. The simulation platform is realized by using PYTHON language. The computer platform used for the experiment is an Intel Core I5-6300HQ computer with 8 GB memory and Windows 10 64-bit computer. In this experiment, the real public location data set Gowalla [33] is used. In order to obtain better experimental results, check-in records of 2000 active users in one year are selected. The Gowalla's data format is shown in Table 4, which contains the user's unique identification, check-in time, and location information.

Table 4. Partial data from Gowalla dataset.

User ID	Check-In Time	Latitude of Check-In Location	Longitude of Check-In Location	Location ID
0	2010-10-19T23:55:27Z	30.2359091167	−97.7951395833	22847
0	2010-10-18T22:17:43Z	30.2691029532	−97.7493953705	420315
0	2010-10-17T23:42:03Z	30.2557309927	−97.7633857727	316637
0	2010-10-17T19:26:05Z	30.2634181234	−97.7575966669	16516
0	2010-10-16T18:50:42Z	30.2742918584	−97.7405226231	5535878
0	2010-10-12T23:58:03Z	30.261599404	−97.7585805953	15372

5.2. Feasibility Analysis

Firstly, the feasibility of the scheme is analyzed. As shown in Figure 6, five recommendation positions with $n = 5$ are generated to visually display the effect of the recommendation algorithm. Where, the blue line segment represents the user's trajectory, and the yellow marks represent the recommended locations of the user. It can be seen from Figure 6 that the recommended positions are similar to the positions through which the user trajectory passes, and there is no overlap with the trajectory. Therefore, the SPDP proposed in this paper is feasible.

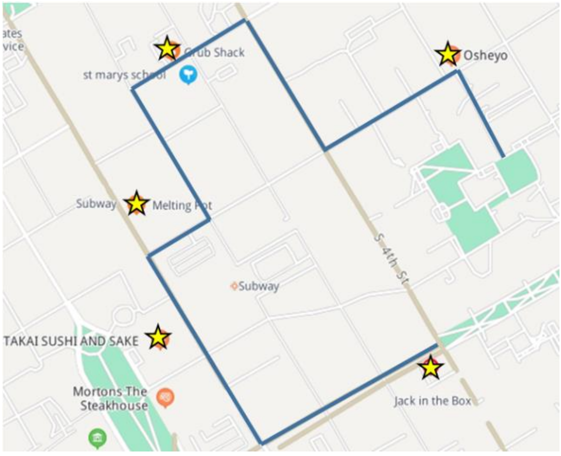

Figure 6. One example of the recommendation result.

5.3. Evaluation Indicators

In order to evaluate the application effect of the scheme proposed in this paper in the location recommendation service, the evaluation indexes commonly used in the recommendation system were selected in the experiment: *Precision*, *Recall* and F-Score [34]. Then, in order to analyze the efficiency, the variation of the algorithm operation time is shown when the number of location recommendations is different. Finally, in order to evaluate the privacy protection degree after adding noise, the ratio of location sensitivity is statistically compared, which means the ratio of the number of locations with different privacy levels to the total number of locations.

Precision and *Recall* are defined as Equation (11) and Equation (12), respectively, where U represents the user set, LR represents the length of recommendation list, $R(u)$ represents the recommended location set of user u, and $T(u)$ represents the interest location set of user u in the test set.

$$Precision = \frac{1}{|U|}\sum_{u \in U}\frac{|R(u) \cap T(u)|}{LR} \tag{11}$$

$$Recall = \frac{1}{|U|}\sum_{u \in U}\frac{|R(u) \cap T(u)|}{T(u)} \tag{12}$$

F-Score represents overall recommendation quality by weighting *Precision* and *Recall*. The comprehensive recommendation effect of the proposed scheme can be evaluated by comparing F-Score. The higher the F-Score is, the higher the recommendation quality is. The definition is shown in Equation (13):

$$\mathrm{F-Score} \;= \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

The location sensitivity ratio is defined in Equation (14), where $|pl_r|$ represents the total number of user's locations with privacy level of pl_r, $|l_u|$ represents the total number of user's locations, and $|U|$ represents the total number of users.

$$R = \frac{\sum_{u \in U}\frac{|pl_r|}{|l_u|}}{|U|} \tag{14}$$

5.4. Experimental Results

In order to prove the effectiveness of SPDP in this paper, it is compared with BOSD in reference [18] and UD of uniform distribution in reference [17]. In this experiment, a total of three levels of location privacy are set. The initial sensitivity for this paper is set to 10. When the sensitivity is less than 10, the location privacy level is 3, that is, the location is not sensitive. Then, the threshold interval is initially set to 40. When the sensitivity is between

10 and 50, the location privacy level is 2, meaning that the sensitivity of the location is normal. When the sensitivity exceeds 50, the maximum privacy level of the location is 1, that is, the location is a sensitive location. Top-*n* (*n* = 5) is used to obtain the set of candidate locations with the highest similarity.

In terms of *Precision* and *Recall*, the recommendation quality of the three methods is inferior to that before adding noise. It can be seen from Figures 7 and 8 that adding Laplace noise will reduce the effect of location recommendation. This is because the noise changes the statistical characteristics of the original trajectory data set and produces certain errors, thus affecting the result of location recommendation. However, the *Precision* and *Recall* of the SPDP are still better than BOSD scheme and UD scheme, reaching 22.4% and 22.7%, respectively. This is because this paper considers the influence of semantic location and further adjusts the allocation of location privacy budget through Markov chain, which makes the recommendation effect of this scheme better.

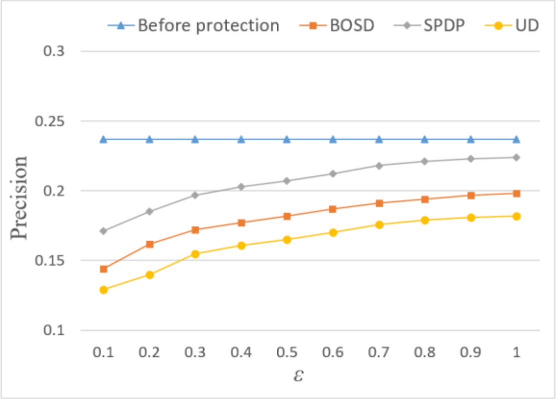

Figure 7. Influence of privacy budget ε on Precision.

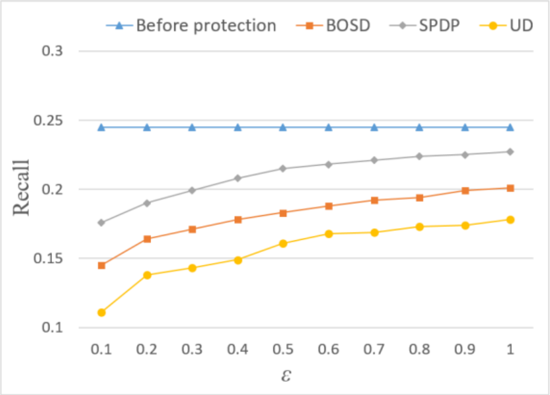

Figure 8. Influence of privacy budget ε on Recall.

The influence of ε on F-Score is shown in Figure 9. The SPDP proposed by this paper is superior to BOSD and UD in terms of comprehensive recommendation quality, with F-Score reaching 22.5%, but there is still some recommendation quality loss. In this paper, the privacy budget allocation method based on location sensitivity keeps the frequency characteristics of users' original trajectory access, and considers the influence of semantic location and users' preference for location. It can better maintain the similarity between positions and reduce the similarity error caused by noise addition. However, due to the sparse check-in data of users, the experimental results will be affected to some extent.

Since the location privacy level is determined by setting a threshold based on the sensitivity of the location, the final result is also affected by the setting of the threshold range. Therefore, the relationship between the comprehensive recommendation quality and the threshold range is compared when the total location privacy levels are 3 and the privacy budget is 0.5. The threshold interval increases from 20, as shown in Figure 10.

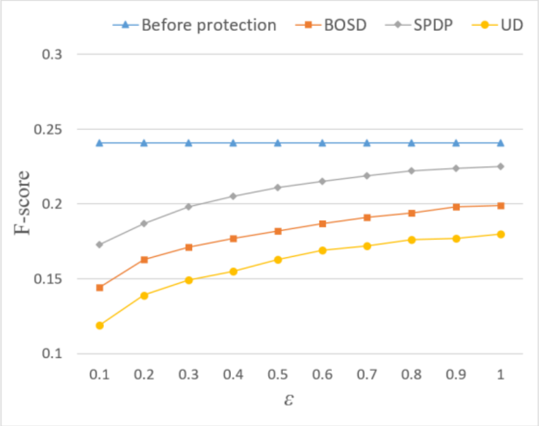

Figure 9. Influence of privacy budget ε on F-Score.

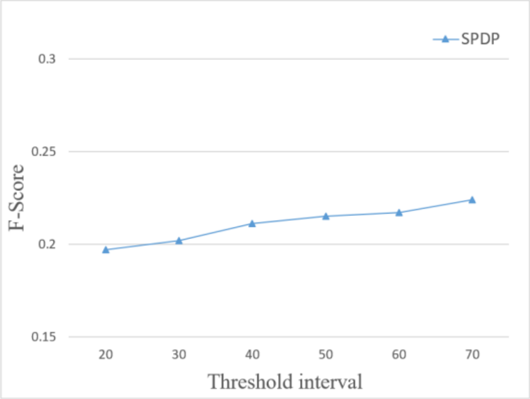

Figure 10. Influence of threshold interval on F-Score.

It can be seen from Figure 10 that, as the threshold increases, the threshold range between each privacy level increases, and the overall recommendation quality also improves. This is because, as the threshold range increases, the number of high privacy locations gradually decreases, so the noise added to them also decreases. However, a decrease in the number of locations with higher privacy levels means a decrease in the intensity of privacy protection. Therefore, it is necessary to set a reasonable threshold range to achieve a certain balance between the quality of location recommendation and privacy protection.

Figure 11 shows the influence of the number of recommended locations on the operation time. The privacy budget is set to 0.5. As can be seen from Figure 11, the operation time is proportional to the number of recommended locations, that is, the more the number of recommended locations, the more operation time. This is because the more the number of recommended locations, the more time it takes to calculate the sensitivity of location points, and the more time it takes to calculate the privacy level of each location point. So, the higher the number of recommended locations, the longer the operation time.

Figure 12 illustrates the impact of privacy budgets on location sensitivity. As shown in Figure 12, with the increase of privacy budget, the added noise gradually decreases, and the proportion of the sensitive position and normal position gradually decreases, while the proportion of the insensitive position gradually increases. This is because the scheme based on sensitivity partition in this paper allocates the budget according to the privacy level and adds more noise to the position with high sensitivity. As the number of sensitive locations increases, it is difficult for the attacker to distinguish the real sensitive locations, thus reducing the probability of identification. Therefore, better privacy protection for sensitive positions on the trajectory can be provided by this scheme.

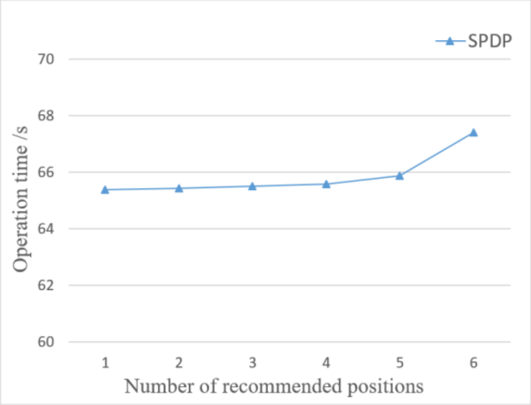

Figure 11. Influence of number of recommended positions on operation time.

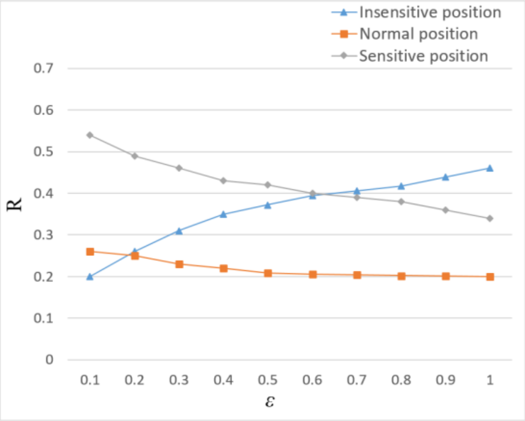

Figure 12. Location sensitivity ratio R under different privacy budget ε.

5.5. Theoretical Analysis

Definition 10. ***Sequential Composition*** [7]. *Given a query algorithm* $M : D \to R^d$ *that supports a random mechanism, if for any data set D and its adjacent data set* D'*, algorithm M satisfies Formula (1) for any output O, then the random algorithm M satisfies ε-differential privacy.*

Sequential composition means that given database D and n random algorithms $\{A_1, A_2, \cdots, A_n\}$, if each algorithm A_i acting on data set D satisfies $\varepsilon_i - DP$, then the sequential sequence group on D satisfies $\left(\sum_{i=1}^{n} \varepsilon_i\right) - DP$. Sequential composition indicates that when multiple algorithm sequences act on a data set at the same time, the final privacy budget is the sum of each algorithm's privacy budget.

Definition 11. ***Parallel Composition*** [7]. *Divide a database D into n disjoint sets* $\{D_1, D_2, \cdots, D_n\}$*, and apply a random algorithm* $\{A_1, A_2, \cdots, A_n\}$*, respectively, on each set, and* A_i *satisfies* $\varepsilon_i - DP$*. Then, the parallel sequence combination on D satisfies* $(max\varepsilon_i) - DP$*. The parallel composition indicates that if multiple algorithms operate on disjoint subsets of a data set, the final privacy budget is the maximum of each algorithm's privacy budget.*

Theorem 1. *Given the total privacy budget ε, Algorithm 3 ensures ε-differential privacy.*

Proof of Theorem 1. In the process of building the noise prefix tree, the noise prefix tree TC is constructed with an easy to understand query model. Consider the height of a prefix tree. It is known that all nodes on the same layer of the prefix tree contain a disjoint set of trajectories. According to Definition 11, the total privacy budget required by each layer is limited by the worst case, that is, $\bar{\varepsilon} = \frac{\varepsilon}{h}$ [17]. Allocating privacy budgets at different levels follows Definition 10. Since there are at most h levels, the total privacy budget required to construct a noisy prefix tree is $\leq h \times \bar{\varepsilon} = \varepsilon$. □

6. Conclusions

In order to protect the trajectory data security of data release, a semantics- and prediction-based differential privacy protection scheme for trajectory data is proposed in this paper. In this scheme, trajectory sequences are stored by a prefix tree structure, and the privacy level of the location is divided by check-in statistics combined with the influence of semantic location. Then, the privacy budget is allocated according to the privacy level, and further adjusted through the Markov chain. The appropriate differential privacy noise is added to the user's check-in position sensitivity, and the position sensitivity level is changed to achieve the effect of privacy protection. By analyzing the experimental results of real location data sets, the proposed scheme can protect the trajectory privacy of users and reduce the impact of differential privacy noise on the quality of service effectively.

The scheme proposed in this paper is based on the centralized differential privacy, which requires that the third-party service providers are completely trusted and will not actively steal or passively leak users' private information. However, in practical applications, it is impossible to find an absolutely secure third-party service provider. Therefore, the local differential privacy model will be introduced to better resist attacks on third-party servers in future research work, so as to achieve better privacy protection effects.

Author Contributions: Conceptualization, J.Z. and Y.L.; methodology, J.Z.; software, Y.L.; validation, L.L. and X.Y.; writing—original draft preparation, Y.L.; writing—review and editing, J.Z. and Q.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No.61902069), Natural Science Foundation of Fujian Province of China (2021J011068).

Institutional Review Board Statement: The study did not require ethical approval and, therefore, chose to exclude this statement.

Informed Consent Statement: The study did not involve humans, therefore, chose to exclude this statement.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: We declare that we have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Aloufi, A.; Hu, P.; Liu, H.; Chow, S.; Choo, K. Universal Location Referencing and Homomorphic Evaluation of Geospatial Query. *Comput. Secur.* **2020**, *102*, 102–137. [CrossRef]
2. Wook, K.; Kennedy, E.; Seon, K.; Dohn, C.; Beakcheol, J. A Survey of differential privacy-based techniques and their applicability to location-Based services. *Comput. Secur.* **2021**, *111*, 102464.
3. Zhang, J.; Xu, L.; Tsai, P. Community structure-based trilateral stackelberg game model for privacy protection. *Appl. Math. Model.* **2020**, *86*, 20–35. [CrossRef]
4. Errounda, F.; Liu, Y. Collective location statistics release with local differential privacy. *Future Gener. Comput. Syst.* **2021**, *124*, 174–186. [CrossRef]
5. Min, M.; Wang, W.; Xiao, L.; Xiao, Y.; Han, Z. Reinforcement Learning-Based Sensitive Semantic Location Privacy Protection for VANETs. *China Commun.* **2021**, *18*, 244–260. [CrossRef]
6. Guo, J.; Yang, M.; Wan, B. A Practical Privacy-Preserving Publishing Mechanism Based on Personalized k-Anonymity and Temporal Differential Privacy for Wearable IoT Applications. *Symmetry* **2021**, *13*, 1043. [CrossRef]
7. Feng, D.; Zhang, M.; Ye, Y. Research on location trajectory publishing technology based on differential Privacy Model. *J. Electron. Inf. Technol.* **2020**, *42*, 74–88.

8. Liu, Q.; Yu, J.; Han, J.; Yao, X. Differentially private and utility-aware publication of trajectory data. *Expert Syst. Appl.* **2021**, *180*, 115–120. [CrossRef]
9. Hemkumar, D.; Ravichandra, S.; Somayajulu, D.V.L.N. Impact of data correlation on privacy budget allocation in continuous publication of location statistics. *Peer Peer Netw. Appl.* **2021**, *14*, 1650–1665. [CrossRef]
10. Chen, S.; Fu, A.; Yu, S.; Ke, H.; Su, M. DP-QIC: A differential privacy scheme based on quasi-identifier classification for big data publication. *Soft Comput.* **2021**, *25*, 7325–7339. [CrossRef]
11. Bao, T.; Xu, L.; Zhu, L.; Wang, L.; Li, R.; Li, T. Privacy-Preserving Collaborative Filtering Algorithm Based on Local Differential Privacy. *China Commun.* **2021**, *18*, 42–60. [CrossRef]
12. Rahimi, S.M.; Far, B.; Wang, X. Contextual location recommendation for location-based social networks by learning user intentions and contextual triggers. *GeoInformatica* **2021**, *26*, 1–28. [CrossRef]
13. Li, W.; Liu, X.; Yan, C. STS: Spatial–Temporal–Semantic Personalized Location Recommendation. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 538. [CrossRef]
14. Dwork, C. Differential privacy. In Proceedings of the 33rd International Colloquium on Automata, Languages and Programming, Venice, Italy, 10–14 July 2006; pp. 1–12.
15. Dwork, C.; McSherry, F.; Nissim, K.; Smith, A. Calibrating noise to sensitivity in private data analysis. In Proceedings of the 3rd Theory of Cryptography Conference (TCC), New York, NY, USA, 4–7 March 2006; pp. 363–385.
16. McSherry, F.; Talwar, K. Mechanism design via differential privacy. In Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS), Providence, RI, USA, 20–23 October 2007; pp. 94–103.
17. Chen, R.; Fung, B.C.; Desai, B.C.; Sossou, N.M. Differentially private transit data publication: A case study on the montreal transportation system. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; ACM: New York, NY, USA, 2012; pp. 213–221.
18. Yin, C.; Ju, X.; Yin, Z.; Wang, J. Location recommendation privacy protection method based on location sensitivity division. *J. Wirel. Commun. Netw.* **2019**, *2019*, 266. [CrossRef]
19. Shao, D.; Jiang, K.; Kister, T.; Bressan, S.; Tan, K.-L. Publishing Trajectory with Differential Privacy: A Priori vs. A Posteriori Sampling Mechanisms. In Proceedings of the Database and Expert Systems Applications, Bilbao, Spain, 30 August–3 September 2010; Springer: Berlin/Heidelberg, Germany, 2013; pp. 357–365.
20. He, X.; Cormode, G.; Machanavajjhala, A.; Procopiuc, C.M.; Srivastava, D. DPT: Differentially private trajectory synthesis using hierarchical reference systems. *Proc. VLDB Endow.* **2015**, *8*, 1154–1165. [CrossRef]
21. Li, M.; Qin, Z.; Wang, C. Sensitive Semantics-Aware Personality Cloaking on Road Network Environment. *Int. J. Secur. Its Appl.* **2014**, *8*, 133–146. [CrossRef]
22. Haldar, N.A.H.; Li, J.; Ali, M.E.; Cai, T.; Chen, Y.; Sellis, T.; Reynolds, M. Top-k Socio-Spatial Co-engaged Location Selection for Social Users. *IEEE Trans. Knowl. Data Eng.* **2020**, 1–16. [CrossRef]
23. Adomavicius, G.; Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 734–749. [CrossRef]
24. Lian, D.; Ge, Y.; Zhang, F.; Yuan, N.J.; Xie, X.; Zhou, T.; Rui, Y. Content-Aware Collaborative Filtering for Location Recommendation Based on Human Mobility Data. In Proceedings of the 2015 IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 14–17 November 2015; pp. 261–270.
25. Lian, D.; Zhao, C.; Xie, X.; Sun, G.; Chen, E.; Rui, Y. Geo MF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24 August 2014; pp. 831–840.
26. Lian, D.; Kai, Z.; Yong, G.; Cao, L.; Chen, E.; Xie, X. Geo MF++: Scalable Location Recommendation via Joint Geographical Modeling and Matrix Factorization. *ACM Trans. Inf. Syst. TOIS* **2018**, *36*, 1–29. [CrossRef]
27. Ding, R.; Chen, Z. Rec Net: A deep neural network for personalized POI recommendation in location-based social networks. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 1631–1648. [CrossRef]
28. Guo, J.; Zhang, W.; Fan, W.; Li, W. Combining Geographical and Social Influences with Deep Learning for Personalized Point-of-Interest Recommendation. *J. Manag. Inf. Syst.* **2018**, *35*, 1121–1153. [CrossRef]
29. Xu, M.; Han, J. Next Location Recommendation Based on Semantic-Behavior Prediction. In Proceedings of the 2020 5th International Conference on Big Data and Computing, Chengdu, China, 28–30 May 2020; pp. 65–73.
30. Mitra, S.; Banerjee, S.; Naskar, M.K. Remodelling correlation: A fault resilient technique of correlation sensitive stochastic designs. *Array* **2022**, *15*, 100219. [CrossRef]
31. Myllyaho, L.; Nurminen, J.K.; Mikkonen, T. Node co-activations as a means of error detection Towards fault-tolerant neural networks. *Array* **2022**, *15*, 100201. [CrossRef]
32. Hua, S.; Qi, L. Convergence of a second order Markov chain. *Appl. Math. Comput.* **2014**, *241*, 183–192. [CrossRef]
33. Cho, E.; Myers, S.A.; Leskovec, J. Friendship and mobility: User movement in location-based social networks. In Proceedings of the 17th ACM SIGKDD International Conference on KNOWLEDGE DISCOVERY and Data Mining, San Diego, CA, USA, 21–24 August 2011; ACM Press: New York, NY, USA, 2011; pp. 1082–1090.
34. Jiao, X.; Xiao, Y.; Zheng, W.; Xu, L.; Wu, H. Exploring Spatial and Mobility Pattern's Effects for Collaborative Point-of-Interest Recommendation. *IEEE Access* **2019**, *7*, 158917–158930. [CrossRef]