EURASIP Journal on
Information Security

RESEARCH Open Access

Trajectory-aware privacy-preserving method with local differential privacy in crowdsourcing

Yingcong Hong[1,2], Junyi Li[1,2*], Yaping Lin[1], Qiao Hu[1] and Xiehua Li[1]

Abstract

In spatial crowdsourcing services, the trajectories of the workers are sent to a central server to provide more personalized services. However, for the honest-but-curious servers, it also poses a challenge in terms of potential privacy leakage of the workers. Local differential privacy (LDP) is currently the latest technique to protect data privacy. However, most of LDP-based schemes have limitations in providing good utility due to extensive noise in perturbing trajectories. In this work, to balance the privacy and utility, we propose a novel pattern-aware privacy protection method called trajectory-aware privacy-preserving with local differential privacy (TALDP). The key idea is that, rather than applying the same degree of perturbation to all location points, we employ adaptive privacy budget allocation, assigning varied privacy budgets to individual location points, thereby mitigating the perturbation's impact and enhancing overall utility. Meanwhile, to ensure the privacy, we give the different perturbing points to different privacy budgets according to their important degree for the patterns of the trajectories. In particular, we use Karman filter method to select the important location points and decide their privacy budgets. We conduct extensive experiments on three real datasets. The results show that our approach improves the utility over many other current methods while still provide good the privacy protection.

Keywords Crowdsourcing, Differential privacy, Information security

1 Introduction

With the proliferation of mobile smart devices, such as smartphones and smart wristbands, and the rapid development of crowdsourcing applications, the relationship between crowdsourced data generated by citizens and the concept of a smart city is becoming increasingly relevant. The availability of data is crucial for the functionality of a smart city [1]. In the crowdsourcing service model, task requesters disseminate tasks with specific target locations or routes through a centralized third-party server, which then allocates these tasks to the respective workers. This process generates a substantial amount of spatiotemporal data associated with the patterns of workers' behavior. While the collection and analysis of this data provides an effective way for servers to assign tasks to workers [2], it also poses privacy threats by potentially exposing sensitive workers' data, such as their locations and movement trajectories. Recent studies have demonstrated that even the disclosure of a small amount of mobile path or location data can result in attacks on anonymous users [3].

In order to mitigate such risks, differential privacy provides a robust protection mechanism [4]. Currently, differential privacy is a practical remedy for preserving the privacy of location and trajectory information [5]. However, in the design of scenarios involving differential privacy [6–8], it requires an authorized and trustworthy data center to collect users' mobile path data. Yet, according to current investigations, 78% of users are still reluctant to let applications collect their mobile path data, fearing that such data collection poses a significant threat to their privacy [9, 10]. As a result, industrial applications

*Correspondence:
Junyi Li
junyilee@hnu.edu.cn
[1] College of Computer Science and Electronic Engineering, Hunan University, Changsha 410012, China
[2] Hunan Provincial Key Laboratory of Blockchain Infrastructure and Application, Hunan University, Changsha 410012, China

have extensively utilized local differential privacy (LDP) [11] (such as Google [12], Apple (https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf), and Microsoft [13]) making it a more appropriate choice for gathering private trajectory data since it does not rely on a trusted third-party server, enables users to locally modify data, and transmits the modified data to an untrusted third-party server. However, some studies have shown that the effectiveness of LDP in protecting privacy is hampered by its emphasis on utility, which comes at the cost of privacy protection [14, 15], particularly, the existing LDP schemes add noise to all points in the trajectory which reduces the utility for the application. To deal with this issue, Haydari [16] adopted a method that involved random sampling of path points and applying noise to the selected points. This approach enabled the construction of random paths between road segments through the utilization of the exponential differential privacy mechanism. However, it relies on a trusted external party to gather and alter the data, and the privacy budget cannot be adaptively adjusted for different path points. In the differential privacy model proposed by Wang.H [17], local differential privacy is primarily used, and the server allocates privacy budgets to achieve efficient adaptive privacy budget allocation for local differential privacy. It focuses on perturbing individual location points by analyzing their distribution within the area, neglecting the correlation between the various location points along the trajectory. On the other hand, Z.Wang [18] took into account the interdependencies between location points in the process of differential privacy protection, reducing the impact of noise on data patterns by minimizing unnecessary perturbations. It has been demonstrated that for time-series data collection, the addition of differential privacy noise alters the original pattern of the data, which subsequently leads to a reduction in utility. Although his research is based on time-series data collection, it aligns well with our trajectory data protection needs.

In this paper, inspired by [18], we focus on protecting the workers' job location privacy and their trajectory information, and propose a segmented differential privacy with noise method. Aiming to preserve the usability of trajectory data while protecting the privacy of user trajectories, we strive to maintain the trajectory patterns even after the introduction of differential privacy noise. To achieve this goal, there are several challenges: (1) how to determine the interdependencies between location points in a trajectory, specifically the importance of each location point. In [18], they suggest using piecewise linear approximation (PLA) in conjunction with a pattern-aware sampling technique to decide whether to sample and disturb the current data point. However, determining the interdependencies between location points in a trajectory poses a significant challenge. (2) How to determine the privacy budget allocation for location points involves deciding how much privacy protection should be allocated to each point. Assigning an equal privacy budget to all location points is not conducive to the usability of the trajectory. Therefore, this becomes another challenge that we need to address. (3) Striking a balance between privacy and the utility of pattern. Preserving patterns in sequences entails sacrificing a certain degree of privacy protection, where larger perturbations often discard more pattern sequences to achieve better privacy protection. Retaining useful pattern information within the preserved pattern sequences while meeting the required privacy protection demands poses a challenge.

To address the aforementioned challenges, this paper proposes a trajectory-aware privacy preserving method with local differential privacy in a new crowdsourcing service model, called trajectory-aware local differential privacy (TALDP). Specifically, regarding challenge (1), we determine the interdependencies between location points by employing a Kalman filter-based trajectory prediction method. When the predicted positions of two consecutive points closely align with the actual positions, we believe that the correlation between points is relatively weak, indicating a minor impact on the trajectory. Conversely, if the predicted positions deviate significantly from the actual positions, we consider the correlation between the points to be strong, indicating a greater impact on the trajectory. In challenge (2), we adaptively allocate privacy budgets based on the importance of location points in the trajectory. As opposed to allocating the same budget for privacy to every location, our approach adaptively allocates a privacy budget based on how much each point affects the trajectory while preserving the pattern of the original trajectory as much as possible. We quantify the dependency between location points through importance assessment, using it as a criterion to adaptively allocate privacy budgets to points of different importance. Our principal contributions are listed below:

- We have designed a novel method for trajectory pattern-aware differential privacy protection, aiming to safeguard the trajectory privacy of workers in crowdsourcing services while ensuring the utility of perturbed trajectory information. Specifically, we have devised a new differential privacy protection scheme tailored to the way servers collect trajectory information from worker users in crowdsourcing services. This technique improves the usefulness of partially perturbed trajectory details while still safeguarding worker trajectory information privacy.

- We propose a method for allocating privacy budget for path points. By using Kalman filtering and inferring the next location point based on the uploaded location points, we compare the prediction error with the actual location point to determine its importance. Adaptive privacy budgets of different sizes are then allocated to perturb location points according to their importance, meeting the requirements of location point privacy protection.
- We run experiments on three real datasets, and the results show that our system preserves significant trajectory patterns while performing better than current mechanisms.

The rest of this paper is organized as follows. In Sect. 2, we present problem formulation in this domain. Then, we show some preliminaries in Sect. 3 and detail our framework in Sect. 4. Finally, we wrap up this paper and give an outlook in Sect. 5 and related work in Sect. 6.

2 Problem formulation

In this section, we first present the description of the system model in Sect. 2.1, and then introduce some background knowledge in Sect. 2.2.

2.1 System model

In a crowdsourcing service, there are typically three parties involved: the server, the task requester, and the workers. We make the assumption that the server is trustworthy yet inquisitive, which means it will faithfully carry out all of the responsibilities assigned to it but might be tempted to look through private user information. Here, we focus on the relationship between the workers and the server because sensitive user data may be compromised through the sensitive data generated during the workers' task processes.

In Fig. 1, every worker generates data while working, locally perturbs the raw data, and at predetermined timestamps transmits the perturbed data to the server. In order to complete all tasks requested by the requester, the server can assign tasks to all workers and provide public services. To improve the efficiency of task completion by workers, the server can provide personalized services by analyzing the data uploaded by the workers.

Current work mainly focuses on protecting the privacy of worker trajectories, for instance, employing probability distributions with superior effectiveness, leveraging exponential mechanisms, and adopting alternative definitions of differential privacy methods. But how to protect the privacy data of each worker while improving data utility to achieve personalized services has not been well addressed.

2.2 Preliminaries

We introduce the background information in this section. Table 1 contains a list of this paper's primary annotations.

Given two datasets D and D', They are defined as neighboring datasets if there is only one record that differs between the two datasets. The function ϕ is added noise to serve the purpose of achieving *ϵ-differential privacy*. Privacy budget ϵ and privacy protection exhibit an inverse relationship, whereby a smaller ϵ cause stronger privacy protection. The amount of noise is measured by *sensitivity* $\Delta\phi$. A widespread noisy method is *Laplace mechanism Lap*$(\frac{\Delta\phi}{\epsilon})$.

Definition 1 **(ϵ-Differential Privacy [19]).** For two neighboring datasets D and D', a privacy mechanism $\mathcal{F}$ satisfies ϵ-differential privacy if, and for every output $\Omega \subset range(\mathcal{F})$, it has:

$$Pr[\mathcal{F}(D) = \Omega] \leq Pr[\mathcal{F}(D') = \Omega] \times e^{\epsilon} \tag{1}$$

Definition 2 **(Sensitivity [19]).** Given query function ϕ: D $\rightarrow$ $\mathbb{R}^n$ on dataset $\mathbb{R}$ with n attributes, and any two neighboring datasets D and D', the sensitivity $\Delta\phi$ is:

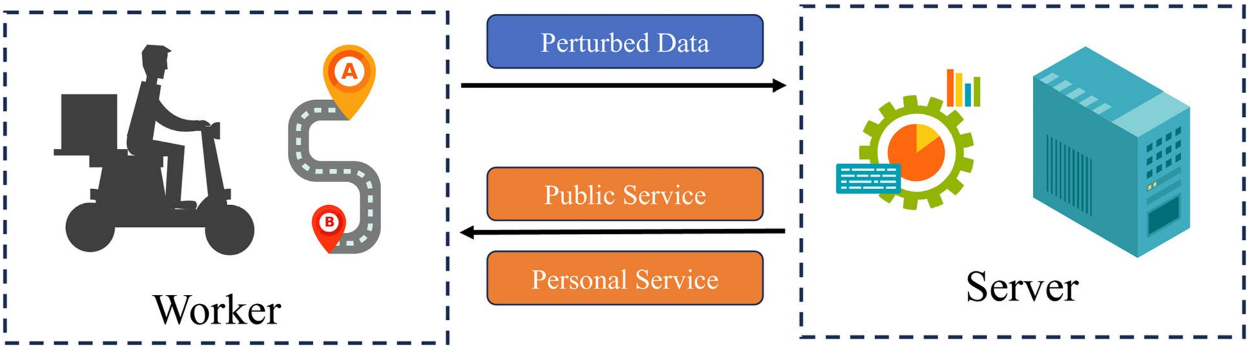


Fig. 1 The model of worker's data collection

Table 1 Main notations

Symbol	Definition
D,D'	Original dataset and neighboring dataset
ϕ,$\Delta\phi$	Query function and sensitivity
$\mathcal{F}$,ϵ	Privacy mechanism and privacy budget
Ω	All possible outputs of the privacy mechanism
ω	The size of metric-base ω-event privacy
F_k	State transition matrix
$\hat{x}_{k\|k}, \hat{x}_{k\|k-1}$	Estimated value and predicted value
B_k	Disturbance transfer matrix
Q_k	Process noise covariance matrix
H_k	Observation matrix
z_k	Current observed actual
R_k	Covariance matrix of observation noise

$$\Delta\phi = \max_{D,D'} ||\phi(D) - \phi(D')||_p \tag{2}$$

In general, $p = 1$,which corresponds to the L_1 norm.

Definition 3 **(Laplace Mechanism** [19]**).** The result of query function ϕ over a dataset D is $\phi(D) = (X_1, X_2, \ldots, X_n)$. The privacy mechanism $\mathcal{F}$ satisfies ϵ-differential privacy if $\mathcal{F}$ is defined as follow:

$$\mathcal{F}(D) = \phi(D) + \left[Lap_1\left(\frac{\Delta\phi}{\epsilon}\right), \ldots, Lap_n\left(\frac{\Delta\phi}{\epsilon}\right)\right] \tag{3}$$

Definition 4 **(ω-neighboring).** For a non-zero positive integer ω, database D and D' of length l are ω-neighboring if: for each $i \in l$, $D[i] \neq D'[i]$, it holds that D and D' are neighboring, and for each $i_1 < i_2$ and $i_1 + i_2 + 1 \leq \omega$, $D[i_1] \neq D'[i_1]$ and $D[i_2] \neq D'[i_2]$.

Definition 5 **(Metric-based ω-Event ϵ−Differential Privacy).** D is a dataset, and $\mathcal{F}$ be a privacy mechanism, metric-based ω-event ϵ−differential privacy is defined as follows: all ω-neighboring dataset D, D' and each possible output $\Omega \subset Range(\mathcal{F})$, it has:

$$Pr[\mathcal{F}(D) = \Omega] \leq Pr[\mathcal{F}(D') = \Omega] \times e^{\epsilon d(D,D')} \tag{4}$$

where $d(D, D')$ is the Euclidean distance between D and D'. The mechanism adhering to metric-based ω-event ϵ−differential privacy offers the minimal utility loss while safeguarding the privacy of general time-series data across consecutive ω timestamps.

3 Trajectory-aware local differential privacy

In this section, we introduce TALDP, in Fig. 2, a privacy protection mechanism for worker trajectory pattern perception in crowdsourcing services. In the material that follows, we give a brief introduction to TALDP before going into greater detail on its main elements and architecture.

3.1 Overview of TALDP

Our goal is to perturb the trajectory information while preserving useful pattern information in the trajectory. Therefore, our main idea is to adaptively perturb useful patterns based on their importance levels. TALDP mainly comprises three mechanisms: trajectory pattern prediction, trajectory pattern importance evaluation, importance-aware perturbation.

3.1.1 *Trajectory pattern prediction*

In this module, we employ Kalman filtering to predict the possible future trajectories due to its high accuracy

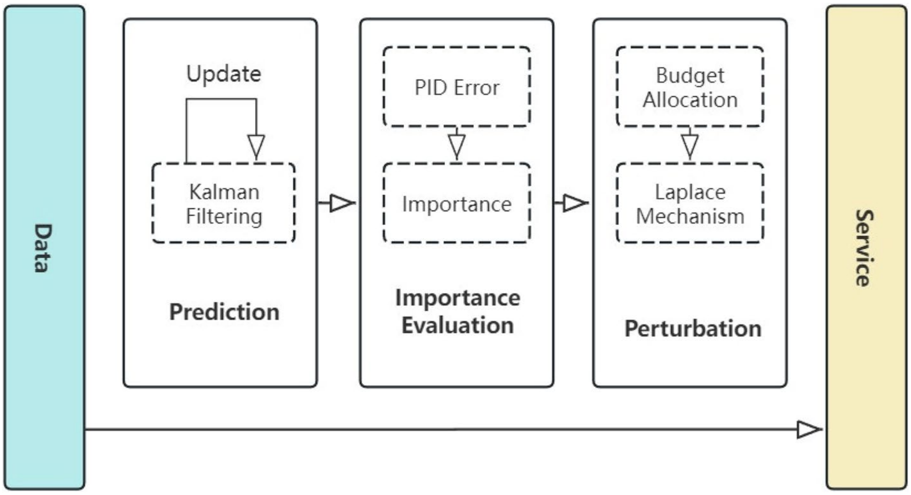


Fig. 2 The overview of the TALDP

and correctable characteristics in trajectory prediction and target tracking. Kalman filtering dynamically estimates the next state at each iteration, considering both historical states and observation data. Kalman filters offer real-time capability, accuracy, robustness, and cost-effectiveness in trajectory prediction, making them one of the preferred methods in numerous application areas. Therefore, we adopt Kalman filtering as the method used for trajectory prediction.

3.1.2 Trajectory pattern importance evaluation

This module is designed to quantify the importance of each position point. When the degree of change in position point data is equivalent, the rapid changes in short-term position point data possess a stronger influence on the overall pattern of the trajectory compared to the slow changes in long-term position point data. This phenomenon is reflected in the trajectory prediction module by observing whether the error between the predicted values and the actual values is experiencing sharp fluctuations. Therefore, We use the PID algorithm to dynamically evaluate data, measuring both the error and importance of position points.

3.1.3 Importance-aware perturbation

This module aims to leverage the importance of each position point obtained, determined in the previous module, to describe the budget that should be allocated to that position point. Based on this budget value, Laplace noise is added to the position point. We ensure that the importance-aware budget allocation is sensitive to trajectory patterns and minimizes the loss of patterns resulting from the addition of noise.

Finally, the perturbed data is transmitted to the server to complete the data upload task. The servers can utilize the approximate results for analyzing the interests and hobbies of workers, aiming to achieve more effective task allocation. We shall provide a thorough introduction to each component module in the text that follows.

3.2 Trajectory pattern prediction

This section describes a method for perceiving trajectory patterns based on the Kalman filter that is consistent with the visual experience. This method is used to assess the extent to which the current trajectory position point influences the current trajectory pattern. The prediction phase and the update step are the two stages that make up the Kalman filter process.

3.2.1 Prediction step

In the prediction step, the current state is linearly projected to obtain the next state prediction based on the system's dynamics model. Additionally, the uncertainty of the system state is estimated using the process noise model. The predicted value of the state vector and the estimated covariance matrix can be calculated using the following formulas:

$$\begin{aligned} \hat{x}_{k|k-1} &= F_k\hat{x}_{k-1|k-1} + B_k u_k \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^T + Q_k \end{aligned} \tag{5}$$

where F_k is the state transition matrix that describes how the current state evolves from the previous state, $\hat{x}_{k-1|k-1}$ is the estimated value of the state vector at the previous time step, B_k denotes the disturbance transfer matrix, u_k denotes the system state noise of the motion model, P_k is the posterior estimation error covariance matrix, and Q_k is the process noise covariance matrix that describes the uncertainty in the system's dynamic model.

The predicted value of the state vector is updated using the state transition matrix and control input, while the estimated covariance matrix is updated based on the system's dynamic model and process noise. Through the prediction step, the next time step's state prediction value and covariance matrix can be obtained, preparing for the subsequent update step.

3.2.2 Update step

In the update step, the predicted state is compared with the actual observed value based on the observation model. The difference between the prediction and observation is calculated, and the weights of the state prediction are adjusted accordingly to make it closer to the actual observation. The uncertainty of the observation is also estimated using the observation noise model. The updated values of the state vector and covariance matrix can be calculated using the following formulas:

$$\begin{aligned} K_k &= P_{k|k-1}H_k^T(H_k P_{k|k-1}H_k^T + R_k)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_K(z_k - Hk\hat{x}_{k|k-1}) \\ P_{k|k} &= (I - K_k H_k)P_{k|k-1} \end{aligned} \tag{6}$$

where H_k is the observation matrix that describes the relationship between observations and the state vector, z_k is the current observed actual, and R_k is the covariance matrix of observation noise, which stands for the observation model's uncertainty.

Update step adjusts the state estimate based on the new observation and updates the covariance matrix to reflect the revised uncertainty. By calculating the Kalman gain, information from the observation is integrated into the state estimate, improving the accuracy of the estimation of the true system state. The updated values of the state vector and covariance matrix are then used as the initial values for the prediction step at the next time step.

By iteratively performing the prediction and update steps, the Kalman filter progressively estimates the optimal state of the system. It not only considers the current observation but also utilizes the statistical properties of the historical state information and observation data, resulting in accurate and adaptable state estimation. In trajectory prediction, the Kalman filter continuously estimates the state of each position point in the trajectory, allowing for the prediction of the next position at a given time. This approach, which is based on historical data and dynamic estimation, provides high accuracy and robustness in trajectory prediction tasks.

3.3 Trajectory pattern importance evaluation

Due to the varying impact of different location points on trajectory patterns, in this section, we use data dynamics to quantify their significance.

Through the trajectory prediction using the Kalman filter in the previous section, we can determine the impact of each location point on the uploaded trajectory. This is because the predicted results of the Kalman filter reflect to some extent the current trend of the trajectory pattern and provide the possible next location point based on the current trajectory pattern. When the error between the predicted location point and the actual location point is within a certain range, we consider that the changes in the location point are in a certain pattern, which we refer to as a long-term pattern. In this case, the influence of the location point on the trajectory pattern is relatively small. However, when there is a significant deviation between the predicted location point and the actual location point, indicating a rapid change in the location, we refer to this as a short-term pattern. In this case, the influence of the location point on the trajectory pattern is relatively large.

We determine the importance level of each location point by using the PID error. First, we define feedback as the difference between the expected and actual values:

$$F_i = |x_i - \hat{x}_i| \tag{7}$$

where x_i is the actual location point and $\hat{x}_i$ is the predicted location point.

Hence, the importance, which is represented by the PID error, can be computed as follows:

$$\gamma[i] = K_p F_i + K_i \sum_{n=0}^{i} F_n + K_d (F_i - F_{i-1}) \tag{8}$$

The PID control comprehensively considers the current error, historical error, and the rate of change of error, by adjusting the proportional coefficients of each component, thereby changing the emphasis on the focus of consideration. Through this approach, we can better assess the influence of the current location point on the trajectory pattern, thus more accurately evaluating its level of importance.

3.4 Trajectory-aware perturbation

In this section, we employ an trajectory-aware randomization mechanism that adaptively allocates privacy budget and injects noise into various locations to minimize the leakage of trajectory patterns. Privacy budget is allocated based on importance, and each location point in the trajectory receives an addition of Laplace noise. The experimental results demonstrate that the trajectory pattern is successfully protected by this technique.

Algorithm 1 *Trajectory-aware perturbation*

```
Input: The importance of location points $\gamma$, private budget $\epsilon$,
    true location point x.
Output: Perturbed value $x^*$
    for $i = 1 \rightarrow +\infty$ do
        while $x \neq NULL$ do
            $\epsilon^{'} = \epsilon - \sum_{j=i-\omega+1}^{i-1} \epsilon[j]$
            $p = 1 - exp(-\gamma[i])$
            $b = \frac{max||x_i, x_{i-1}||_1}{\epsilon}$
            $x^* = x + \frac{1}{2b} exp(-\frac{|x|}{b})$
        end while
    end for
```

3.4.1 *Importance budget allocation*

For trajectory sequence data, the metric-based ω-event privacy requirement states that the cumulative budget for any sliding window of ω location points should not exceed ϵ. This means that each location point in the trajectory can be allocated a certain amount of budget.

In TALDP, our goal is to maintain the pattern of trajectories as much as possible. The position locations that significantly affect the trajectory patterns should undergo only minimal alterations. Therefore, the location points with higher importance should receive a larger privacy budget. To elaborate, the following is the definition of the percentage function that allocates the remaining budget to the current sampling point:

$$p = 1 - exp(-\gamma[i]) \tag{9}$$

The range of p from 0 to 1 is guaranteed by the exponential function. The privacy budget allocated to the current location point is calculate as $\epsilon[i] = p * \epsilon'$, where ϵ' is the remaining budget in the ω-event window $\epsilon' = \epsilon - \sum_{j=i-\omega+1}^{i-1} \epsilon[j]$.

As mentioned in [18], when there are consecutive significant points (location points), it can rapidly deplete the privacy budget, resulting in insufficient allocation of budget to subsequent highly important significant points (location points). This is similar to our approach, with the difference being that [18] focuses on data sampling points in a data stream, whereas we concentrate on the process of workers' location uploading for crowdsourcing services. We employ the Kalman filtering method to predict the location points in the trajectory. Kalman filtering provides accurate predictions of future location points, regardless of whether the uploaded location points are sparse or dense and whether they have a significant impact on the trajectory pattern. By using Kalman filtering to compare the predicted and actual position points, we can draw relatively accurate conclusions, without facing the issue of rapid privacy budget consumption. Therefore, we can determine the privacy budget for a particular location point using this method.

3.4.2 Laplace perturbation

Laplace perturbation is a commonly used method for adding noise in differential privacy. It utilizes random numbers generated from the Laplace distribution to perturb the original data, thereby achieving differential privacy protection. Its definition is as follows:

$$Laplace(x|b) = \frac{1}{2b} exp\left(-\frac{|x|}{b}\right) \tag{10}$$

where the b represents the scale parameter, and their values determine the shape and peak degree of the Laplace distribution. And b has the following relationship with ϵ:

$$b = \frac{\triangle\phi}{\epsilon} \tag{11}$$

where $\triangle\phi$ represents the sensitivity in differential privacy, which is also the range of oscillation after adding noise. We represent the sensitivity by the maximum distance between adjacent location points in the trajectory:

$$\triangle\phi = max||x_i, x_{i-1}||_1 \tag{12}$$

To give the distribution of Laplace noise a more comprehensible picture, we showed its probability density function; as shown in Fig. 3, Laplace noise is related to the sensitivity of the data and has been proven to have ability to provide strong privacy protection.

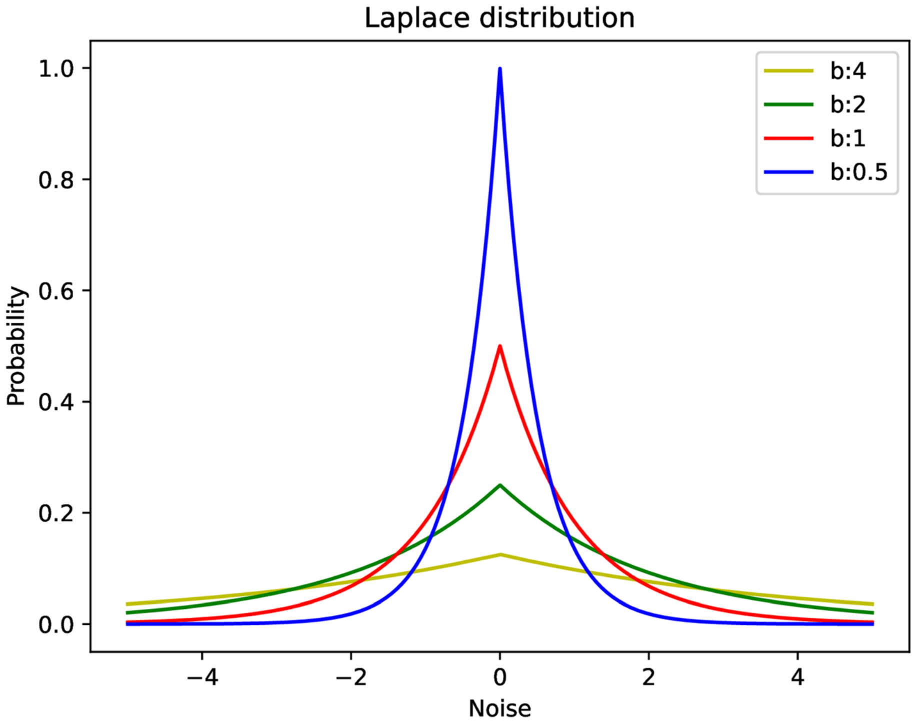


Fig. 3 The probability density function of Laplace

Since the position points are two-dimensional data, and Laplace noise generates one-dimensional values, we have transformed the obtained Laplace noise accordingly. We conceptualize the Laplace noise result as defining a circle centered on the original position point. Consequently, the original position point is perturbed to a random location within that circular range. This approach achieves the intended purpose of adding noise.

When Laplace noise, adapted to the level of importance, is added to important data, it becomes easier to distinguish between data points when a pattern change occurs. This makes trajectory data more useful, while also providing privacy protection by reducing distinguishability between data points when no pattern change occurs.

3.5 Theoretical analysis

We demonstrate that the TALDP method maintains differential privacy in this subsection.

Theorem 1 *The importance-aware perturbation satisfies ϵ-LDP.*

Proof For two location point x, x', where x denotes the latitude and longitude coordinates of the location points and x' another location point adjacent to x, abbreviated as x and x', assuming that the current remaining privacy budget is $\epsilon[i]$. As per the Laplace perturbation definition, we can obtain:

$$\frac{Pr[x^*|x]}{Pr[x^*|x']} = \frac{\frac{1}{2b}e^{\left(-\frac{|x|}{b}\right)}}{\frac{1}{2b}e^{\left(-\frac{|x'|}{b}\right)}} = e^{-\frac{\epsilon[i]}{\Delta\phi}(||x^*-x||-||x^*-x'||)} \tag{13}$$

It can always be true that $||x^* - x|| - ||x^* - x'|| \leq ||x - x'||$. Thus Eq. 13 can be updated:

$$e^{-\frac{\epsilon[i]}{\Delta\phi}(||x^*-x||-||x^*-x'||)} \leq e^{-\frac{\epsilon[i]}{\Delta\phi}(||x,x'||)} \tag{14}$$

where $-\frac{||x,x'||}{\Delta\phi}$ is constant. Thus, our importance-aware perturbation satisfies ϵ-LDP.

Theorem 2 *TALDP satisfies metric-based ω-event ϵ-differential privacy.*

Proof In Proof 1, we show that TALDP satisfies ϵ-LDP. As for the ω-event ϵ-differential privacy, any perturbed point with an allocated budget of $k\epsilon/\omega$, it recycles the privacy budget in the previous $k-1$ points. Thus, its previous $k-1$ points and following $k-1$ points all have zero allocated privacy budget. As a result, the sum of budgets of any ω successive points satisfies $0 \leq \sum_{k=i+1}^{i+\omega} \epsilon[k] \leq \epsilon$. We recover the privacy budget for $k \leq i$ in the process of allocating the budget and adaptively allocate the privacy budget for the next location points in the remaining privacy budget.

Thus, TALDP satisfies metric-based ω-event privacy.

4 Experiments

In this section, we assess the performance of the suggested TALDP using four real-time series datasets. We approach our experiments from three different angles: evaluating the utility of statistical estimation of trajectory position points, assessing the utility of trajectory pattern analysis, and evaluating the impact of importance-aware randomization on pattern preservation.

4.1 Experimental datasets

We use four real-world location datasets for our investigations.

- Gowalla dataset [20] is a classic dataset used for location recommendation and social network analysis. It collects location check-in data from users of the social networking application Gowalla. Gowalla was a once-popular location-sharing app where users could mark their locations on a map and share their activities and travel experiences with other users.
- Geolife dataset (http://snap.stanford.edu/data/loc-gowalla.html) collects 17,621 trajectories from 182 users during the period of April 2007 to August 2012. Each GPS trajectory in this dataset is represented as a sequence of time-stamped points, containing latitude and longitude information, with various sampling rates. Ninety-one percent of the trajectories are densely sampled. The dataset captures a wide range of users' outdoor movements, including daily routines and leisure activities.
- ShangHai dataset [21] refers to the approximately 100,000 aggregate GPS tracks collected from taxis operating in Shanghai in 2007. This dataset comprises extensive time-series data, documenting real-time location information of each taxi across various areas of Shanghai, along with associated timestamps and speed information.

4.2 Comparison

In this section, We compared TALDP with three methods focusing on position point privacy protection and two methods focusing on trajectory pattern privacy protection, as shown below:

- GRR [22]: A commonly used and fundamental LDP scheme involves adding random perturbations based on pre-defined probabilities.
- PLDP [23]: An optimized LDP framework has been developed to achieve high availability, and the model proposed in this framework is suitable for other types of data, although it was specifically designed for spatial data.
- L-SRR [17]: For various location-based LDP frameworks, high availability is required to privately collect and analyze user locations. To address this challenge, a new randomization mechanism called "staircase random response" is proposed, which has theoretical guarantees for both privacy protection and availability. Experimental results demonstrate the feasibility of the proposed mechanism.
- PLPC [24]: A differential privacy protection scheme proposed for vector-based and frequently accessed locations in trajectory databases.
- LDPTPM [25]: A trajectory privacy protection method is proposed, which focuses on protecting the regions of interest for users rather than their entire region.

4.3 Experimental parameters

In performance comparison, we assess TALDP's performance from three perspectives: utility of trajectory's position points and utility of trajectory pattern analysis and the effect of TALDP. Specifically, we use the average L1-distance method as a measure of utility for location point statistics analysis and estimate the error as a measure of utility for trajectory pattern analysis.

4.3.1 L1-distance (Manhattan distance)

A metric used in mathematics and computer science to measure the absolute difference between two points in a coordinate system. It calculates the sum of the absolute differences between the corresponding coordinates of the points along each dimension,which is defined as follows:

$$L_1 = \frac{1}{n}\sum_{n=1}^{n} |x_i - \hat{x}_i| + |y_i - \hat{y}_i| \tag{15}$$

where, for $i \in n$, x_i and y_i are the real coordinate values, and $\hat{x}_i$ and $\hat{y}_i$ are the perturbed data of the actual position corresponding to index i.

4.3.2 Dynamic time warping (DTW)

The DTW distance is commonly used to measure the similarity of patterns in time series matching and is also applicable for trajectory pattern matching. The DTW distance is calculated using a dynamic programming approach, and the calculation formula is as follows:

$$\begin{aligned} DTW(i,j) &= dist(T_i, \hat{T}_i) + DTW' \\ DTW' &= \begin{cases} DTW(i-1, j-1) \\ DTW(i, j-1) \\ DTW(i-1, j) \end{cases} \end{aligned} \tag{16}$$

where T_i and $\hat{T}_i$ are the real trajectory positions and privacy-protect trajectory positions.

In our assessment, each evaluation indicator is normalized for comparison. We set $K_p = 0.1$,$K_i = 0.15$,$K_d = 0.1$ for importance characterization as default. The privacy budget ranges are divided into two categories based on the comparison objects: 0.1 to 1.0 and 1.0 to 8.0. The reason for selecting different differential privacy budgets is that when the budget is smaller, the added perturbation value is larger. At this point, the availability of the location cannot accurately reflect the true data availability. Conversely, when the budget is larger, the calculated trajectory similarity is generally lower, which also cannot accurately reflect the true data availability. All scenarios are budgeted to meet ω-event ϵ differential privacy for fairness.

4.4 Performance comparison

In this section, we will first analyze the effective of Kalman filter for trajectory prediction. Next, we show the performance comparison on the utility of trajectory position point statistical analysis, assessing the utility of trajectory pattern analysis.

4.4.1 Effective of Kalman filter

In this section, we have experimentally demonstrated the practical applicability of Kalman filter in predicting trajectory patterns. We first compare the variability between real and predicted trajectories in terms of latitude and longitude. In Fig. 4, Kalman filter predictions result in trajectories that exhibit a similar pattern to the original trajectory. This indicates that Kalman filter is able to accurately capture and predict the dynamics of the trajectories through its unique algorithmic logic. In Fig. 5, we present the errors between the Kalman filter prediction results and the original trajectory, along with the error variation, utilizing the same data set as in Fig. 4. We observe that when the trajectory pattern remains stable, the error variation tends to remain at a smoother level. Conversely, when the trajectory pattern undergoes significant changes, the error variation increases accordingly. This aligns with our notion that the importance of the current location point is determined by the degree of error change. Specifically, a smaller error change

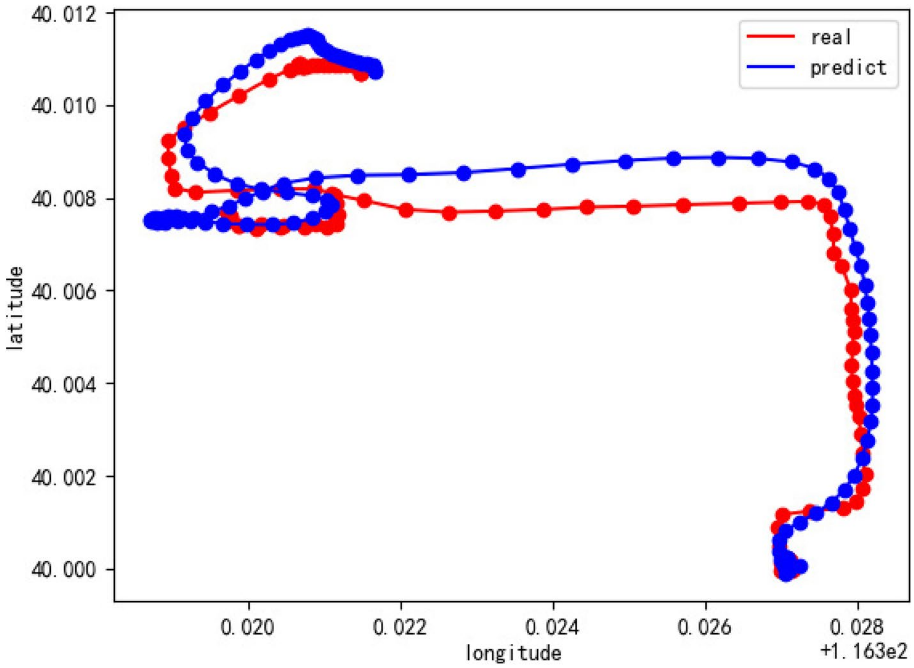


Fig. 4 Comparison of real and predicted trajectories

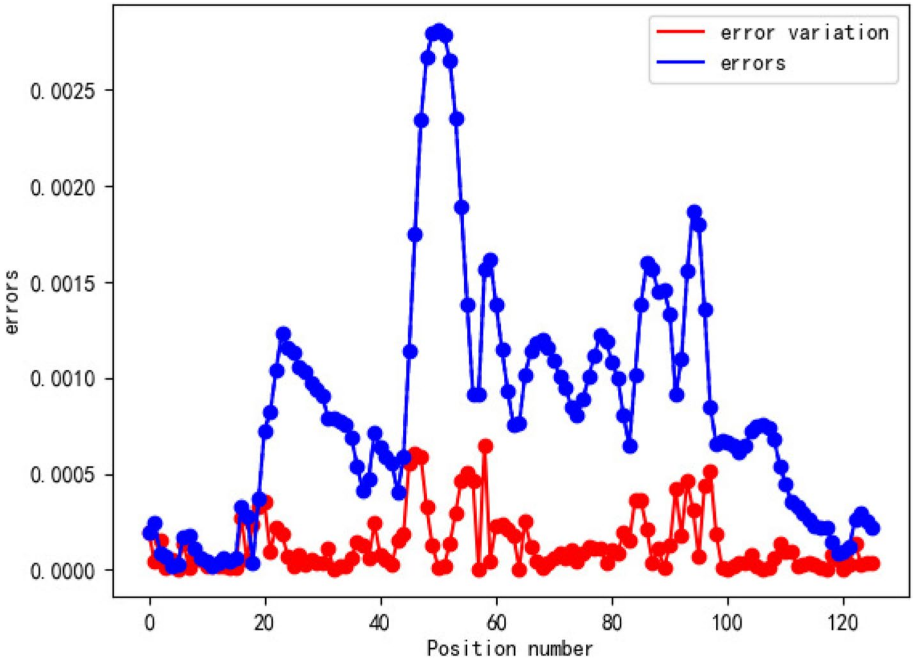
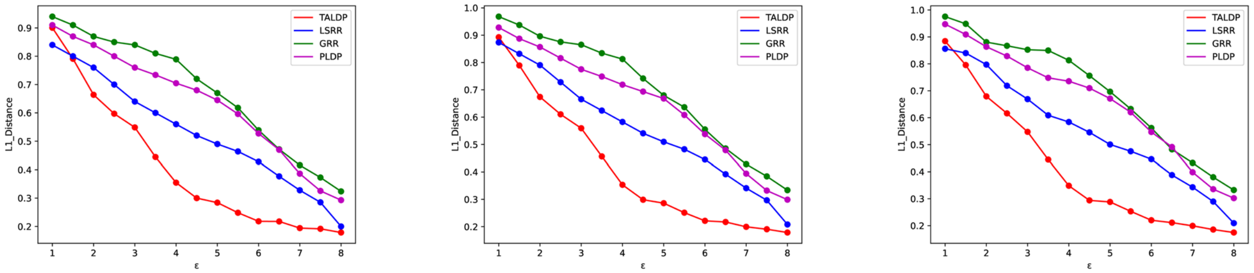


Fig. 5 Errors of real and predicted trajectories

indicates a smoother trajectory pattern and thus a lower importance, while a larger error change signifies a more significant change in the trajectory pattern and thus a higher importance. Therefore, we utilize the Kalman filter to forecast trajectory variations and dynamically allocate the privacy budget by comparing the predicted outcomes with the actual results. This approach aims to enhance the utility of the data.



(i) GeoLife (ii) Gowalla (iii) Shanghai

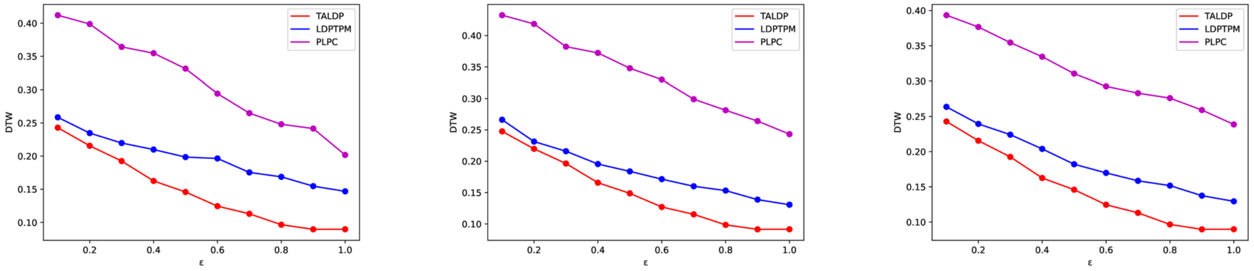
Fig. 6 Experimental results of the statistical analysis

4.4.2 Utility of trajectory's position point

In this section, we compare TALDP with three existing trajectory protection schemes that focus on location distribution. Figure 6 shows the average L1 distance of all four mechanisms decreases as the privacy budget increases. In differential privacy, as the privacy budget increases, the noise perturbation added at each location point decreases, leading to a smaller sum of the calculated offsets for each point. We can see that TALDP exhibits a transition from a rapid to a gradual change in the average L1 distance as the privacy budget increases. This is because TALDP allocates the remaining privacy budget based on the importance level of each location point, distributing it by a certain multiple. As the total privacy budget grows exponentially, the privacy budget allocated to certain location points also increases proportionally, leading to a rapid decline in the L1 distance. This reflects the high utility of TALDP even with a small privacy budget. In the graph, we can see that our scheme has similar utility compared to other schemes at $\epsilon = 1$ and $\epsilon = 8$, but it demonstrates better utility performance in the range of $\epsilon \in [1, 8]$, which also reflects the significant improvement in the utility of trajectory data achieved by TALDP.

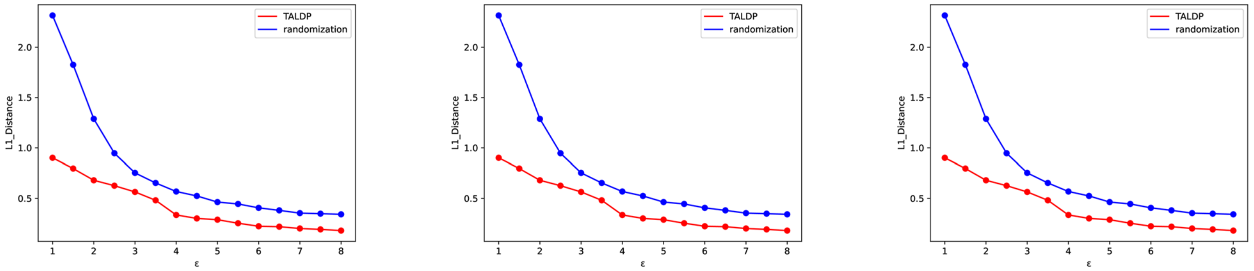
4.4.3 Utility of trajectory's pattern

In this section, we compare our approach with two existing methods that focus on improving the utility of trajectory data. The comparison metric is the degree of pattern matching between the perturbed and original trajectory data. It is worth noting that DTW only represents whether the patterns between trajectories are similar. We normalize the DTW results obtained by TALDP and two existing methods for trajectory protection and compare them as shown in the graph. In Fig. 7, all three methods show a decreasing trend with the increase of ϵ, which matches the fact that higher perturbation is added with a smaller privacy budget. The faster rate of change in TALDP is consistent with our previous explanation. It can be seen that TALDP always presents a better utility performance during the process of ϵ change. It can be observed that LDPTPM and TALDP exhibit similar levels of usability; however, our method still maintains certain advantages during the ϵ variation process. This is because LDPTPM is based on data-driven LDP, allocating budget based on the state of each location point. In contrast, TALDP considers the impact of location points on the entire trajectory, thus holding a certain advantage in terms of trajectory pattern usability.



(i) Geolife (ii) Gowalla (iii) Shanghai

Fig. 7 Experimental results of the statistical analysis



(i) Geolife (ii) Gowalla (iii) Shanghai

Fig. 8 Utility comparison with different randomization methods

4.4.4 *The effect of TALDP*

In this section, we compare the effectiveness of adaptive privacy budget allocation. We compare the adaptive privacy budget allocation method with traditional random privacy budget allocation. From the Fig. 8, it can be observed that compared to the substantial degradation in utility of random privacy budget allocation at low privacy budgets, adaptive privacy budget allocation still maintains good utility. Allocating an equal level of privacy budget to each location point in a trajectory would result in each point experiencing the same degree of perturbation, thereby compromising the usability of trajectory data. In contrast, our method takes into account the trend of trajectory changes, aiming to preserve the trajectory patterns as much as possible, thus achieving better usability. Furthermore, as the privacy budget increases, the adaptive privacy budget allocation method exhibits better utility and stability compared to random privacy budget allocation.This demonstrates the advantages of the adaptive privacy budget allocation method over random allocation.

In our experiments, we compare the different schemes by changing the budget to observe the changes of the utility. In the experiments, Figs. 6, 7, and 8 give the comparing results. In the figures, the smaller ordinate value means the better utility. The figures show that, for the same budget (meaning same privacy quality), our scheme has the best utility.

5 Related work

The services for space crowdsourcing, or SC, have advanced significantly. Additionally, this has caused the problem of location privacy to become more complicated. There are lots of well-liked ways to keep location privacy private. We review pertinent work in this section.

5.1 Space crowdsourcing

Spatial crowdsourcing service is a service that employs crowdsourcing methods for the collection, processing, and analysis of geographical spatial data. In spatial crowdsourcing services, individuals or organizations can use online platforms to post tasks such as map annotation, geographical information collection, and route planning, which are then completed by a group of volunteers or professionals through online participation [26]. These tasks typically involve collecting, annotating, verifying, or analyzing data related to geographical locations, features, and terrains [27]. The advantage of spatial crowdsourcing services lies in their ability to utilize large-scale volunteer networks to rapidly and efficiently complete tasks related to the collection and processing of geographic information, while also reducing costs. Through crowdsourcing, global volunteer resources can be mobilized to collect and analyze large-scale geographic spatial data, particularly for tasks requiring extensive manpower and time, such as map updates, geographic information verification, and satellite image interpretation [28]. Furthermore, spatial crowdsourcing services provide individuals with opportunities to participate in social welfare, academic research, and commercial projects, while also promoting the open sharing and widespread application of geographic information data. However, spatial crowdsourcing services also face challenges in terms of data quality, privacy protection, and task management, requiring corresponding measures to ensure the accuracy, security, and credibility of the data.

5.2 Differential privacy

Several traditional local-privacy protection patterns have been proposed to address the issue of privacy leakage, such as k-anonymity [29, 30]. These techniques, however, are quite susceptible to attacks leveraging previous knowledge. Higher security is attained by another kind of privacy protection technique that uses encryption [31]. However, the computational cost and communication overhead can be very high. In [32], Dwork developed the concept of differential privacy (DP) first. Many different privacy protection frameworks have been suggested for various spatial crowdsourcing perception tasks (such as data pushing [33], incentive mechanism [34], task allocation [35]). However, in cases where users do not trust the server, it is possible that the DP model is not appropriate for real-world location uses.

For privacy data collection, local-differential privacy (LDP) allows each user to upload data after perturbing it locally, which offers more security than centralized DP approaches. Additionally, Andrés et al. loosened the geo-indistinguishability protection of areas inside a radius. Loosened by Chen et al. [23], LDP enable users to choose customized privacy budgets for private location gathering. However, they cannot guarantee strict LDP. Zhang et al. [25] proposed the SPDM-TSR method, which is based on spatiotemporal constraint for mining interesting regions, focusing on protecting the areas of interest to users rather than the entire region. Wang, Han et al. [17] proposed the L-SRR method, which privately collects and analyzes user locations with high utility. However, They focus on the distribution of location points while ignoring the impact of location points on trajectories, leading to the loss of trajectory information. In this paper, our solution has achieved better results in terms of the availability of trajectory information while protecting location point privacy. Testing on actual datasets has also shown the feasibility of this approach.

6 Conclusions

In this paper, we focus on the real-time trajectory data collection of servers in crowdsourcing services during worker work and propose a pattern recognition-based trajectory privacy protection mechanism called TALDP. This mechanism aims to protect the privacy and security of worker trajectories while preserving the original patterns in the trajectory data. We propose an importance-aware scheme that evaluates the importance level of location points in the trajectory using Kalman filtering and adapts the perturbation level based on the importance level, rather than providing the same level of privacy protection for each location point. Numerous tests conducted on real-world datasets confirm that TALDP is a useful tool for enhancing pattern usability.

Authors' contributions

Yingcong Hong contributed to writing—original draft preparation, review, and editing. The author read and approved the final manuscript. Junyi Li and Yaping Lin are responsible for the refinement of the idea and supervision of this literature analysis. Qiao Hu and Xiehua Li have contributed in reviewing the article.

Funding

This work is supported by the National Natural Science Foundation of China Under Grant 62272158.

Availability of data and materials

The compiled and refined data and results of this manuscript should be available upon request by the authors.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors gave their consent for publication.

Competing interests

The authors declare no competing interests.

Received: 28 April 2024 Accepted: 14 August 2024
Published online: 02 September 2024

References

1. A. Fornaroli, D. Gatica-Perez, Urban crowdsourcing platforms across the world: A systematic review. Digit. Gov. Res. Pract. **4**(3), 19 (2023). https://doi.org/10.1145/3603256
2. G. Marzano, J. Lizut, L.O. Siguencia, Crowdsourcing solutions for supporting urban mobility. Procedia Comput. Sci. **149**, 542–547 (2019). https://doi.org/10.1016/j.procs.2019.01.174
3. Y.A. de Montjoye, C. Hidalgo, M. Verleysen et al., Unique in the crowd: The privacy bounds of human mobility. Sci. Rep. **3**, 1376 (2013). https://doi.org/10.1038/srep01376
4. S.-S. Ho, S. Ruan, in *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS (SPRINGL '11)*. Differential privacy for location pattern mining (Association for Computing Machinery, New York, 2011), pp. 17–24. https://doi.org/10.1145/2071880.2071884
5. M.E. Gursoy, L. Liu, S. Truex, L. Yu, W. Wei, in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*. Utility-aware synthesis of differentially private and attack-resilient location traces (Association for Computing Machinery, New York, 2018), pp. 196–211. https://doi.org/10.1145/3243734.3243741
6. X. He, G. Cormode, A. Machanavajjhala, C.M. Procopiuc, D. Srivastava, DPT: Differentially private trajectory synthesis using hierarchical reference systems. Proc. VLDB Endow. **8**(11), 1154–1165 (2015). https://doi.org/10.14778/2809974.2809978
7. M. Mohammady, S. Xie, Y. Hong, M. Zhang, L. Wang, M. Pourzandi, M. Debbabi, in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. R2DP: A universal and automated approach to optimizing the randomization mechanisms of differential privacy for utility metrics with no known optimal distributions (Association for Computing Machinery, New York, 2020). pp. 677-696. https://doi.org/10.1145/3372297.3417259
8. H. Wang, S. Xie, Y. Hong, VideoDP: A flexible platform for video analytics with differential privacy. Proc. Priv. Enhancing Technol. **2020**, 277–296 (2020)
9. K. Chatzikokolakis, E. ElSalamouny, C. Palamidessi, A. Pazii, Methods for location privacy: A comparative overview, now. (2017). https://doi.org/10.1561/3300000017
10. H.H. Arcolezi, J.-F. Couchot, B. Al Bouna, X. Xiao, Improving the utility of locally differentially private protocols for longitudinal and multidimensional frequency estimates. Digit. Commun. Netw. **10**(2), 369–379 (2024). https://doi.org/10.1016/j.dcan.2022.07.003
11. G. Cormode, S. Jha, T. Kulkarni, N. Li, D. Srivastava, T. Wang, in *Proceedings of the 2018 International Conference on Management of Data (SIGMOD '18)*. Privacy at scale: Local differential privacy in practice (Association for Computing Machinery, New York, 2018), pp. 1655–1658. https://doi.org/10.1145/3183713.3197390
12. Ú. Erlingsson, V. Pihur, A. Korolova, in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. RAPPOR: Randomized aggregatable privacy-preserving ordinal response (Association for Computing Machinery, New York, 2014), pp. 1054–1067. https://doi.org/10.1145/2660267.2660348
13. B. Ding, J. Kulkarni, S. Yekhanin, in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Collecting telemetry data privately (Curran Associates Inc., Red Hook, 2017), pp. 3574–3583
14. J.W. Kim, B. Jang, Workload-aware indoor positioning data collection via local differential privacy. IEEE Commun. Lett. **23**, 1352–1356 (2019)
15. X. Zhao, Y. Li, Y. Yuan, X. Bi, G. Wang, LDPart: Effective location-record data publication via local differential privacy. IEEE Access **7**, 31435–31445 (2019)
16. A. Haydari, C.-N. Chuah, M. Zhang, J. Macfarlane, S. Peisert, in *Proceedings of the 38th Annual Computer Security Applications Conference (ACSAC '22)*. Differentially private map matching for mobility

trajectories (Association for Computing Machinery, New York, 2022), pp. 293–303. https://doi.org/10.1145/3564625.3567974

17. H. Wang, H. Hong, L. Xiong, Z. Qin, Y. Hong, in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22)*. L-SRR: Local differential privacy for location-based services with staircase randomized response (Association for Computing Machinery, New York, 2022), pp. 2809–2823. https://doi.org/10.1145/3548606.3560636
18. Z. Wang, W. Liu, X. Pang, J. Ren, Z. Liu, Y. Chen, in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*. Towards pattern-aware privacy-preserving real-time data collection (Toronto, 2020), pp. 109–118. https://doi.org/10.1109/INFOCOM41043.2020.9155290
19. C. Dwork, J. Lei, in *Proceedings of the forty-first annual ACM symposium on Theory of computing (STOC '09)*. Differential privacy and robust statistics (Association for Computing Machinery, New York, 2009), pp. 371–380. https://doi.org/10.1145/1536414.1536466
20. Y. Zheng, L. Zhang, X. Xie, W.-Y. Ma, in *Proceedings of the 18th international conference on World wide web (WWW '09)*. Mining interesting locations and travel sequences from GPS trajectories (Association for Computing Machinery, New York, 2009), pp. 791–800. https://doi.org/10.1145/1526709.1526816
21. Hong Kong Univ. Sci. Technol., Smart City Res. Group, Shanghai, China, Feb. 2007.
22. S. Wang, L. Huang, P. Wang, H. Deng, H. Xu, W. Yang, in *Wireless Algorithms, Systems, and Applications. WASA 2016*, ed by Q. Yang, W. Yu, Y Challal. Private weighted histogram aggregation in crowdsourcing. Lecture Notes in Computer Science, vol 9798 (Springer, Cham, 2016)
23. R. Chen, H. Li, A.K. Qin, S.P. Kasiviswanathan, H. Jin, in *2016 IEEE 32nd International Conference on Data Engineering (ICDE)*. Private spatial data aggregation in the local setting (Helsinki, 2016), pp. 289–300. https://doi.org/10.1109/ICDE.2016.7498248
24. Chen P , Gu J , Zhu D ,et al. A Dynamic Time Warping based Algorithm for Trajectory Matching in LBS[J]. International Journal of Database Theory & Application. **6**, (2013)
25. W. Zhang, Z. Xie, A. M. Vera Venkata Sai, Q. Zia, Z. He, G. Yin, A Local Differential Privacy Trajectory Protection Method Based on Temporal and Spatial Restrictions for Staying Detection, in Tsinghua Science and Technology. **29**(2), 617–633 (2024). https://doi.org/10.26599/TST.2023.9010072.
26. M.E. Andrés, N.E. Bordenabe, K. Chatzikokolakis, C. Palamidessi, in *Proceedings of the 2013 ACM SIGSAC conference on Computer and communications security (CCS '13)*. Geo-indistinguishability: Differential privacy for location-based systems (Association for Computing Machinery, New York, 2013), pp. 901–914. https://doi.org/10.1145/2508859.2516735
27. A. Degbelo, C. Granell, S. Trilles, D. Bhattacharya, S. Casteleyn, C. Kray, Opening up smart cities: citizen-centric challenges and opportunities from GIScience. ISPRS Int. J. Geo Inf. **5**, 16 (2016)
28. Crowdsourced Data Mining for Urban Activity: Review of Data Sources, Applications, and Methods[J]. J Urban Plan Dev. **146**(2), 4020007.1–4020007.15 (2020). https://doi.org/10.1061/(ASCE)UP.1943-5444.0000566
29. P. Samarati, Protecting respondents' identities in microdata release. IEEE Trans. Knowl. Data Eng. **13**(6), 1010–1027 (2001). https://doi.org/10.1109/69.971193
30. L. Sweeney, k-Anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **10**, 557–570 (2002)
31. P.M. Asuquo, H.S. Cruickshank, J.G. Morley, C.P. Anyigor Ogah, A. Lei, W. Hathal, S. Bao, Z. Sun, Security and privacy in location-based services for vehicular and mobile communications: An overview, challenges, and countermeasures. IEEE Internet Things J. **5**, 4778–4802 (2018)
32. Dwork, C. Differential Privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds) Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science. **4052**, (Springer, Berlin Heidelberg, 2006). https://doi.org/10.1007/11787006_1
33. Z. Wang, X. Pang, Y. Chen, H. Shao, Q. Wang, L. Wu, H. Chen, H. Qi, Privacy-preserving crowd-sourced statistical data publishing with an untrusted server. IEEE Trans. Mob. Comput. **18**(6), 1356–1367 (2019). https://doi.org/10.1109/TMC.2018.2861765
34. Z. Wang, X. Pang, J. Hu, W. Liu, Q. Wang, Y. Li, H. Chen, When mobile crowdsensing meets privacy. Comm. Mag. **57**(9), 72–78 (2019). https://doi.org/10.1109/MCOM.001.1800674
35. Z. Wang, J. Hu, R. Lv, J. Wei, Q. Wang, D. Yang, H. Qi, Personalized privacy-preserving task allocation for mobile crowdsensing. IEEE Trans. Mob. Comput. **18**, 1330–1341 (2019)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Yingcong Hong received obtained a Bachelor's degree in Software Engineering from Hunan University. Currently pursuing a Master's degree in Software Engineering at the School of Information Science and Engineering, Hunan University.His research interests include privacy protection and blockchain.

Junyi Li received the B.S., M.S., and Ph.D. degrees in computer application from Hunan University, in 1993, 2001, and 2008, respectively. Since 2005, he has been an Associate Professor with Hunan University. From 2009 to 2010, he was a Visiting Researcher with Lakehead University, Canada.His research interests include big data analysis, software engineering, privacy protection, and blockchain.

Yaping Lin Professor and Ph.D. supervisor. Obtained a bachelor's degree from Hunan University in 1982, a master's degree from the National University of Defense Technology in 1985, and a doctoral degree from Hunan University in 2000. From 2004 to 2005, conducted visiting research at the University of Texas at Arlington in the United States. Engaged in teaching and research in computer science and technology and software engineering for many years. His main research interests include computer networks, cloud security, and machine learning.

Qiao Hu became a Member (M) of IEEE in 2017. She received her B.S. degree and M.S. degree from Central South University, China, in 2000 and 2003, respectively. Subsequently, she received her Ph.D. degrees from Shanghai Jiaotong University, China, in 2007, in Communication and Information System. She is currently an Assistant Professor in the College of Computer Science and Electronic Engineering, Hunan University, China.Her research interests include cryptography, cloud computing, system security.

Xiehua Li received her B.S. degree and M.S. degree from Central South University, China, in 2000 and 2003, respectively. Subsequently, she received her Ph.D. degrees from Shanghai Jiaotong University, China, in 2007, in Communication and Information System. She is currently an Assistant Professor in the College of Computer Science and Electronic Engineering, Hunan University, China.Her research interests include cryptography, cloud computing, and system security.