Multimedia Tools and Applications (2024) 83:50225–50243
https://doi.org/10.1007/s11042-023-17464-6

Continuous release of temporal correlation location statistics with local differential privacy

Renhao Hu[1] · Hongjiao Li[1] · Jinguo Li[1] · Zhaohui Wang[1] · Baojin Wang[1]

Received: 24 May 2023 / Revised: 25 September 2023 / Accepted: 3 October 2023 /
Published online: 4 November 2023

Abstract

The continuous release of location statistics plays a significant role in various real-world applications, such as traffic management and customization of public services. However, existing literature primarily focuses on static scenarios or perturbing locations at a single timestamp, disregarding the consideration of temporal correlation in mobile users. This oversight leaves the data susceptible to privacy attacks, including inference attacks, resulting in extra privacy leakage. To address this challenge, we propose a Local Differential Privacy Budget Distribution and Streaming Data Releasing (LPBD) mechanism for real-world location datasets. Specifically, we investigate the problem of continuously releasing location statistics for infinite streams while protecting user privacy and quantify the impact of temporal correlation on privacy leakage. The LPBD is a novel w-event level privacy-preserving mechanism, which has the capability to provide an adequate privacy budget for each timestamp and effectively mitigate the privacy leakage problem resulting from temporal correlation. Experimental results demonstrate that LPBD enhances data availability with strong privacy guarantees compared to state-of-the-art baseline methods.

Keywords Streaming data · Continuous observation · Location statistics · Temporal correlation · Local differential privacy · w-event privacy

1 Introduction

Location statistics have been widely used in many data-driven applications, such as traffic volume control and customized public services [1–4]. However, the continuous collection and analysis of streaming data seriously violates each participant's privacy. These streaming data involve sensitive information about individuals, and privacy leakages will accumulate over time [5–7], leading to detrimental consequences. Differential Privacy (DP) [8] is a strict privacy architecture frequently applied in real-time surveillance systems to protect the privacy of all users. Local Differential Privacy (LDP), a distributed variation of DP [9, 10], ensures

✉ Hongjiao Li
hjli@shiep.edu.cn

[1] College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai 201306, China

each user's privacy without requiring the original data to leave the client or the use of trusted third-party servers. It is currently one of the most popular privacy-preserving paradigms and finds widespread use in the industry [11].

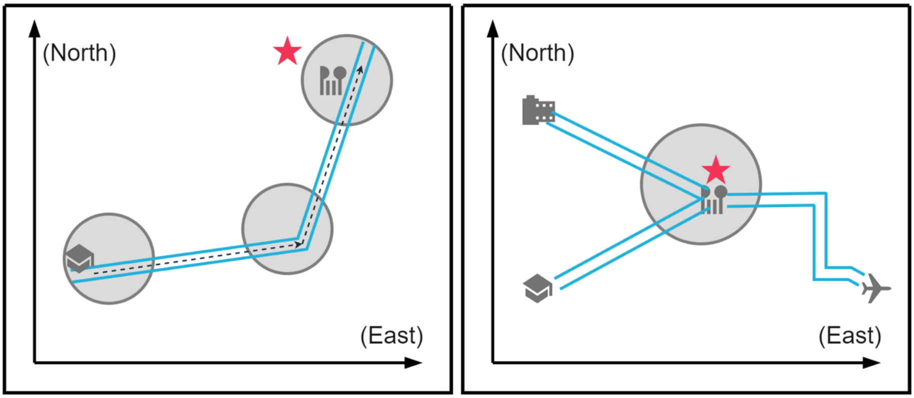
Differential privacy research for evolving (streaming) data analysis can be roughly categorized into three categories according to the level of privacy protection: event-level, user-level, and w-event level privacy. In earlier research [8, 12, 13], most studies focused on event-level privacy and user-level privacy. However, the former protects just one data point from the user's entire stream, whereas the latter is impractical for the majority of real-world scenarios because it requires protecting the user's presence throughout an infinite stream. To resolve this dilemma, the concept of w-event level [14–17] privacy is introduced. The aim is to guarantee that every privacy event within a sliding time window, consisting of w consecutive time intervals or timestamps, achieves the protection of differential privacy.

However, in streaming data collection scenarios, traditional location protection methods [4, 10, 18, 19] do not consider the temporal correlation of mobile users, making them vulnerable to various inference attacks and resulting in extra privacy leakages. The following example [20, 21] illustrates how inference attacks can lead to extra privacy leakage.

1. Consider a user who rides from school to the restaurant, as shown in Fig. 1(a) (where ”star” is). The perturbed position is released by selecting a point in each of three randomly selected circles. Although a single location appears to be protected at each timestamp, an adversary could potentially be able to accurately infer that the user is in a restaurant based on realistic road constraints or the user's movements, resulting in privacy leakage.
2. Consider the user's location "star" as depicted in Fig. 1(b). If the user can only be present in the four places shown in the graph at the current timestamp based on previous location estimates, then the perturbed location may reveal the true location. Thus, technically, the temporal correlation should have an effect on the radius of the circle.

Main contributions

To mitigate the challenge of extra privacy leakage induced by the temporal correlation between continuously generated stream data, we propose the LPBD mechanism, which complies with w-event level privacy constraints. The mechanism incorporates a budget distribution scheme capable of adapting to the presence or absence of temporal correlation between timestamps. In brief, LPBD is a straightforward and effective LDP mechanism designed to



(a) (b)

Fig. 1 Examples of privacy leakage caused by temporal correlations of user locations

address privacy preservation challenges arising from temporal correlation within user data. The main contributions of this paper are detailed below.

- We developed LPBD, a versatile LDP mechanism to assess dissimilarities between statistics, determining whether to release current or most recent statistics. This mechanism ensures adequate budget distribution for consecutive timestamps within each sliding time window, which is used for releasing statistics. The proposed distribution scheme is flexible and applicable to different datasets.
- We explored the challenge of releasing infinite streams of mobile user data in a local environment, comparing it with state-of-the-art baseline methods. Furthermore, we quantified the impact of temporal correlation on user data privacy leakage and provided theoretical proof of LPBD's ϵ-LDP compliance.
- We conducted a comprehensive series of experiments using real-world datasets to investigate the influence of two key parameters: the window length w and the total privacy budget value within sliding windows. Concurrently, we calculated the average error between the output value and the true value of each timestamp as an indicator to evaluate the usability and privacy loss of different methods. Experiments show that LPBD outperforms baseline methods in terms of data availability.

The remainder of the paper is structured as follows. Section 2 discusses related work. The preliminaries of this paper are given in Section 3. Section 4 quantifies the impact of privacy leakage. Section 5 details the LPBD mechanism and provides a privacy analysis of LPBD. Section 6 presents the experimental results. Finally, Section 7 draws the conclusions, discussing our future work.

2 Related work

The processing of streaming data and privacy-preserving mechanisms under DP have been extensively studied, initially focusing on two concepts: event-level DP, which hides individual events in streaming data, and user-level DP, which tries to hide all events of a single user over the temporal stream.

For instance, Dwork et al. [8] initiated the research and proposed an event-level DP mechanism based on binary tree technology for finite streams. Chan et al. [22], inherited the binary tree-based [8] method and extended the method to handle infinite streams. Nonetheless, event-level DP is usually not enough for privacy protection, while user-level DP is generally only implemented on finite streams. For example, Fan et al. [23] proposed FAST, a sampling and filtering framework that introduces noise for user-level DP over finite streams and adaptively samples data elements using the Kalman filter and Proportional Integral Derivative (PID) feedback error.

To resolve the dilemma, Kellaris et al. [14] proposed a w-event level privacy paradigm that tries to guarantee each privacy event is contained within a w-length sliding time window, protecting any sequence of events. Subsequently, they developed the Budget Distribution (BD) mechanism and the Budget Absorption (BA) mechanism based on a sliding window methodology. Furthermore, by integrating the w-event DP with FAST, Wang et al. [24] propose RescueDP, a multidimensional stream release mechanism that includes dynamic sampling, adaptive grouping, perturbation, and filtering to release multiple infinite streams with considerable data utility.

However, these DP mechanisms cannot be used directly in an LDP environment due to the assumption that untrusted servers cannot directly access raw data. The above mechanisms all protect user privacy under the assumption of a trusted server and cannot prevent internal attacks from inferring user privacy [21]. As for the streaming data, inspired by the construction of binary trees in [8], Erlingsson et al. [11] introduced a memory mechanism and proposed RAPPOR for releasing binary attributes with LDP to provide long-term LDP guarantees.

Besides, Joseph et al. [25] propose THRESH, an LDP mechanism for streaming data whose privacy budget is only consumed when users vote to decide whether a global update is required. Moreover, Ren et al. [16] extended the BD and BA [14] mechanisms into the LDP environment by proposing a mechanism for streaming data statistics under w-event LDP. Nevertheless, LDP mechanisms for streaming data are still in their infancy.

At the same time, none of the above series of works considered the impact of user data relevance on privacy protection. The subsequent work involves this aspect of the research. Cao et al. [5] investigated the potential privacy loss of traditional DP mechanisms under temporal correlation. Xiao et al. [20] investigated an adversary who knew the temporal correlations of specific users, and they developed the Planar Isotropic Mechanism (PIM) for continuous location-sharing scenarios. On the basis of [20], Fang et al. [21] incorporated the innovative Staircase mechanism into the location protection model. Hemkumar et al. [26] proposed the Privacy Budget Allocation (PBA) mechanism to address the temporal correlation problem in a centralized environment.

Building upon the aforementioned foundation, this paper explores the privacy protection challenges associated with temporal correlation streaming data continuously generated by users in the LDP scenario by adopting the w-event privacy paradigm. The goal is to address the existing gap in the literature and extend the understanding of this crucial aspect of privacy preservation.

3 Preliminaries

In this section, some definitions and related theorems will be briefly introduced.

3.1 Local differential privacy

LDP is a privacy paradigm built on the following neighborhood definitions:

Definition 1 (Neighborhood) *Two datasets $\mathcal{X}$ and $\mathcal{X}'$ are neighbors if they differ in at most one element. This means that one dataset is a subset of the other, and the larger dataset contains exactly one extra row.*

LDP guarantees that statistical query results for any neighboring datasets remain indistinguishable to protect sensitive personal data. Consequently, it prevents strong adversaries from making inferences about private information from search results. LDP proves especially valuable in distributed environments and does not require data collectors to be trusted. By perturbing their own data before sending it to the data collector, each user ensures that an attacker cannot compromise their privacy. As a result, raw data remains solely stored with the data generator (i.e., the user) and not shared with third parties. The formal definition is as follows:

Definition 2 (ϵ-LDP) *A randomized privacy mechanism $\mathcal{M}$ provides ϵ-LDP if for all pairs $\mathcal{X}, \mathcal{X}' \in$ Domain($\mathcal{M}$) and every possible output $\mathcal{O} \in$ Range($\mathcal{M}$):*

$$P_r[\mathcal{M}(\mathcal{X}) \in \mathcal{O}] \leq e^{\epsilon} \times P_r[\mathcal{M}(\mathcal{X}') \in \mathcal{O}]. \tag{1}$$

In the above definition, the constant ϵ is referred to as the privacy budget. It reflects the indistinguishability between real statistics and private statistics. Stronger privacy protection results from smaller ϵ, and vice versa.

The Random Response (RR) mechanism was introduced in [27], specifically, each user gives a correct or opposite answer to a sensitive question based on the outcome of a coin toss. As users gave random responses, the data collector was unable to determine the true answers of individuals, but usage statistics could still be extracted. The use of the RR mechanism in LDP mechanisms is widespread.

3.2 *w*-event privacy

For continuously releasing data streams, LDP has a variant termed w-event privacy [18, 19]. It provides provable privacy guarantees for any series of events occurring within any sliding window of length w. The following is the definition:

Definition 3 (w-neighboring) *Let $S_t = \{D_1, D_2, ..., D_t\}$ be a prefix stream of sequential data containing a dataset D_i with an arbitrary number of rows, each corresponding to a distinct user at each timestamp i. Two prefix streams, S_t and S_t' are considered to be w-neighboring for any positive integer w if the following conditions are satisfied:*

(1) *for each D_i and $D_i^{'}$, $i \in [1, t]$ and $D_i \neq D_i^{'}$, it holds that D_i and $D_i^{'}$ are neighboring, and;*

(2) *for each $i_1 \in [1, t]$, $i_2 \in [1, t]$, $i_1 < i_2$, and $i_1 \neq i_2$, it holds that $i_2 - i_1 + 1 \leq w$.*

Definition 4 (w-event privacy) *Let $\mathcal{M}$ represent a random mechanism, and O represent the domain of all possible $\mathcal{M}$ outputs. When S_t and $S_t{}'$ are w-$neighboring$, $\mathcal{M}$ satisfies w-event ϵ-local differential privacy (or, simply, w-event privacy) if it holds that:*

$$P_r[\mathcal{M}(\mathcal{S}_t) \in \mathcal{D}] \leq e^{\epsilon} \times P_r[\mathcal{M}(\mathcal{S}_t^{'}) \in \mathcal{D}]. \tag{2}$$

Definition 5 (Laplace mechanism) *For any query function f, suppose that each user u_i records a numerical attribute value $f(D_t)$ and that there are a total of n users. To add noise to $f(D_t)$, define the following random function:*

$$\mathcal{M}(D_t) = f(D_t) + Lap(\Delta f/\epsilon) \tag{3}$$

Where $f(D_t)$ is the value after noise addition and $Lap(\lambda)$ is a Laplace-distributed random variable with scale $\lambda = (\Delta f/\epsilon)$, i.e., noise. The difference between the maximum and minimum values in $f(D_t)$ is represented by Δf. The probability density function of the Laplace distribution is $P_r(x|\lambda) = \frac{1}{2\lambda}e^{\frac{-|x|}{\lambda}}$.

Sequential composition is an essential element under w-event privacy, which can be described as follows:

Theorem 1 (Sequential composition) *Consider $\mathcal{M}_i(v)$ is an ϵ_i-LDP mechanism with an input value of v, and $\mathcal{M}(v)$ is the sequential composition of $\mathcal{M}_1(v)$,..., $\mathcal{M}_k(v)$. Then a sequence of mechanism $\mathcal{M}(v)$ over a data stream $\mathcal{S}$ provides $\sum_{i=1}^{k} \epsilon_i$-LDP.*

4 Privacy leakage analysis under temporal correlation

In this section, the notation and prior knowledge are described in subsection 4.1. Subsequently, the mobility modeling process is introduced in subsection 4.2, which will be utilized in the analysis of privacy leakage discussed in subsection 4.3.

4.1 Prior knowledge

A collection of data containing n tuples, denoted by an index $[n] = \{1, 2, ..., n\}$, intends to release the output of a certain query function $s = f(\mathbf{x})$, where $\mathbf{x} = \{x_1, x_2, ..., x_n\}$. To preserve the privacy of all private attributes, it will return noisy answers $\mathbf{y} = \mathcal{M}(f(\mathbf{x}))$ by adding random noise extracted from certain distribution. Therefore, all possible outputs Ω construct the probability distribution $P_r(\mathcal{M}(f(\mathbf{x}) \in \Omega)$. A set of Θ is utilized to represent the extent to which the adversary possesses knowledge about data correlations. It is not possible to guarantee privacy against adversaries outside of Θ due to the lack of researchability under arbitrary distributions.

The degree of privacy leakage in correlated user data is influenced by the quantity of prior knowledge, as shown in various studies [28, 29]. When the adversary possesses knowledge of $\mathbf{x}_k$, the notation $\Psi_{i,k}$ is used to represent the adversary's attempt to infer information regarding the tuple x_i. The term "target" of the adversary is denoted as x_i, while $\mathbf{x}_k$ represents prior knowledge, $k \in [n] \backslash \{i\}$, where $[n] = \{1, 2, ..., n\}$. The more tuples in $\mathbf{x}_k$, the more prior knowledge the adversary has and the greater the threat to user private data.

4.2 Mobility model

In the context of streaming data releasing, it can be assumed that the adversary $\Psi_{i,k}$ possesses knowledge of the transition probabilities between the user's potential location. In this article's settings, the transition probability between the user's potential locations is modeled using a Markov Chain (MC) process , and they are referred to as $\theta \in \Theta$, where Θ represents the set distribution of all transition probabilities. A stochastic process that transitions from one state in the state space to another is called a MC process.

Definition 6 (Temporal correlations) *The temporal correlations between the data l_i^{t-1} and l_i^t of user u are described by transition matrices P_i, denoted as $P_r(l_i^{t-1}|l_i^t)$.*

The user's mobility is modeled as a first-order MC with temporal correlation sections. In this application, the first-order MC denotes a situation where the probability of a user changing their location state at a certain time depends solely on their previous position. The probability of transferring from one data point to another is described in MC by the transition matrix P_i, where the sum of transition probabilities for each row is fixed at 1. As an illustration, generate a transition matrix of dimension 2 as shown in Table 1. If a user u is at loc_1 (the current location), then the probability of coming from loc_2 (the previous location) is 0.3, expressed as $P_r[l_i^{t-1} = loc_2 | l_i^t = loc_1] = 0.3$.

Considering the temporal aspects, access probability, and transfer probability, the Markov model becomes a valuable tool for accurately predicting the movement of a user's location [5, 6].

Table 1 Example of a transition matrix

	loc_1	loc_2
loc_1	0.7	0.3
loc_2	0.1	0.9

4.3 Quantifying temporal correlation privacy leakage

The maximum ratio of two distributions (e.g., Laplace distribution) for all distinct values of $l_i^t, l_i^{t'}$, and all possible transition probability distributions is the Temporal Correlation Privacy Leakage (TPL) of privacy mechanism $\mathcal{M}_t$ with respect to Ψ_i^θ.

The privacy leakage of $\mathcal{M}_t$ with respect to T_i is called privacy leakage. Inspired by [5, 26, 30], TPL is defined as follows:

$$TPL_{\Psi_i^\theta}(T_i, \mathcal{M}_t) \overset{def}{=\!=} \sup_{l_i^t, l_i^{t'}, \theta} log \frac{P_r(p|l_i^t, S_t, \theta)}{P_r(p|l_i^{t'}, S_t, \theta)} \tag{4}$$

$$TPL_{\Psi_i^\theta}(\mathcal{M}_t) \overset{def}{=\!=} \max_{\forall T_i, i \in [n]} TPL_{\Psi_i^\theta}(T_i, \mathcal{M}_t) \tag{5}$$

In w-event privacy, the TPL of $\mathcal{M}_t$ with respect to any Ψ_i^θ where $i \in [n]$ is less than or equal to the total privacy budget ϵ.

$$\sup_{\forall T_i, i \in [n]} TPL_{\Psi_i^\theta}(\mathcal{M}_t) \leq \epsilon \tag{6}$$

Next, the (4) is transformed and simplified using the Bayesian theorem to obtain the impact of temporal correlation on privacy leakage in the continuous data-releasing setting, i.e.,

$$\begin{aligned} TPL_{\Psi_i^\theta}(\mathcal{M}_t) &= \sup_{o_1, l_i^t, l_i^{t'}, \theta} log \frac{P_r(p|l_i^t, S_t, \theta)}{P_r(p|l_i^{t'}, S_t, \theta)} \\ &= \sup_{o_1, ..., o_{t-1}, l_i^t, l_i^{t'}, \theta} log \frac{\sum_{l_i^{t-1}} P_r(o_1, ..., o_{t-1}|l_i^{t-1}, S_{t-1}) P_r(l_i^{t-1}|l_i^t)}{\sum_{l_i^{t-1}} \underbrace{P_r(o_1, ..., o_{t-1'}|l_i^{t-1'}, S'_{t-1})}_{(i) TPL(T_i, \mathcal{M}_{t-1})} \underbrace{P_r(l_i^{t-1}|l_i^t)}_{(ii) P_i}} \\ &+ \sup_{o_t, l_i^t, l_i^{t'}, \theta} log \frac{P_r(o_t|l_i^t, S_t)}{\underbrace{P_r(o_t|l_i^{t'}, S'_t)}_{(iii) PL_0(T_i, \mathcal{M}_t)}} \end{aligned} \tag{7}$$

The three annotated items in the (7) are subsequently discussed. The first item represents the privacy leakage at the previous timestamp $t-1$, the second item represents the temporal correlation as determined by the transition matrix P_i, and the last item is equal to the privacy leakage at time t with no temporal correlation.

As a result, note that if $t = 1$, then $TPL_{\Psi_i^\theta}(\mathcal{M}_t) = PL_0(\mathcal{M}_1)$; if t >1, then satisfy the following equation:

$$TPL_{\Psi_i^\theta}(\mathcal{M}_t) = TPL_{\Psi_i^\theta}(\mathcal{M}_{t-1}) + PL_0(\mathcal{M}_t) \tag{8}$$

Equation 8 also reveals that TPL is recursively computed and that privacy leakages may accumulate over time.

5 Proposed mechanism

This section elaborates on the details of the LPBD mechanism and provides proof of its satisfaction of ϵ-LDP.

5.1 Motivation

In general, w-event privacy mechanisms offer lower privacy guarantees compared to classical ϵ-LDP, particularly when user data is not independent across consecutive timestamps (i.e., correlated in time) [5, 28]. This limitation arises due to the inadequate privacy budget allocated at timestamps within a sliding window. Consequently, the existing state-of-the-art privacy budget distribution methods in w-event privacy, such as BD and BA mechanisms [14], are not suitable for scenarios with temporal correlation in data points.

In this paper, we accept the idea of sliding windows, improve the existing budget distribution scheme, and adapt it to the local environment. The proposed LPBD mechanism employs a well-designed budget distribution strategy to effectively reduce or even eliminate the impact of temporal correlation on privacy protection. Note that the total privacy budget ϵ consumed must be less than or equal to the sum of all budgets within a sliding window of length w. The length of the sliding window at any time i ranges from $t - w + 1$ to t.

5.2 Algorithm description

The generation of private statistics relies on real-world location statistics, while the LPBD mechanism simplifies this process into two distinct steps: the decision phase and the release phase. As shown in Fig. 2, location statistics for each entity are produced at each timestamp i. Subsequently, the real statistics are locally compared with the last private release. If they exhibit similarity, the most recent release version is returned according to the approximation strategy. Conversely, if they exhibit dissimilarity, the privacy budget is distributed based on the presence of temporal correlation, and this portion of the budget is utilized to perturb the real statistics before their release.

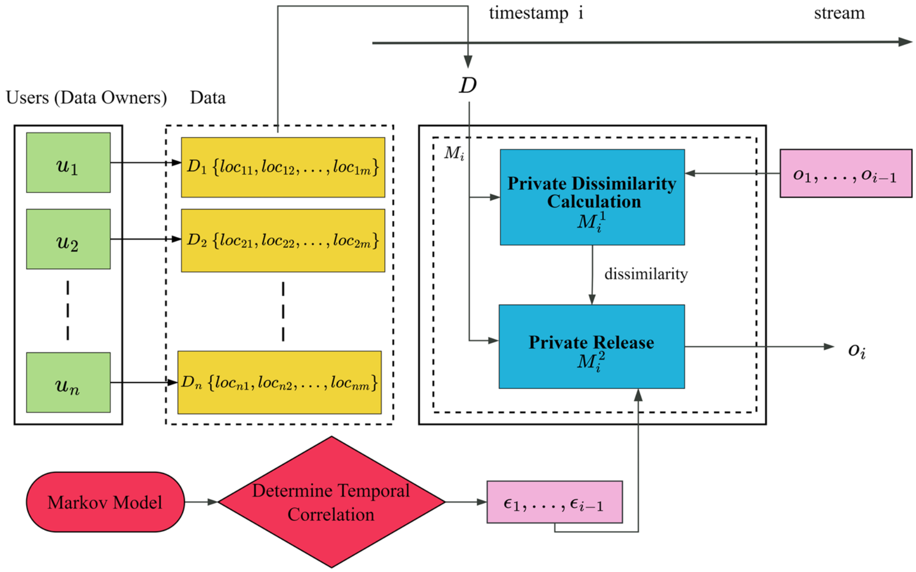

timestamp i
stream
D
Users (Data Owners)
Data
u_1
u_2
u_n
$D_1 \{loc_{11}, loc_{12}, \ldots, loc_{1m}\}$
$D_2 \{loc_{21}, loc_{22}, \ldots, loc_{2m}\}$
$D_n \{loc_{n1}, loc_{n2}, \ldots, loc_{nm}\}$
M_i
Private Dissimilarity Calculation M_i^1
dissimilarity
Private Release M_i^2
$o_1, \ldots, o_{i-1}$
o_i
Markov Model
Determine Temporal Correlation
$\epsilon_1, \ldots, \epsilon_{i-1}$

Fig. 2 The overall process structure of LPBD

Table 2 Notations summary

Notation	Description
c_i^r	Real location statistics for the current timestamp i
o_l^p	Released private location statistics for the last timestamp l
l	The last recent release timestamp l
ϵ	Total budget available for distribution
ϵ_i	Privacy budget assigned at timestamp i
$\sum \epsilon_i^1$	Privacy budget for decision phase
$\sum \epsilon_i^2$	Privacy budget for release phase

The LPBD mechanism M is comprised of a series of sub-mechanisms $\{M_1, M_2, ..., M_i, ..., M_t\}$, where each M_i takes the dataset $S_t[i] = D_i$ as input and releases private statistics o_i using the assigned privacy budget ϵ_i. As a result, M released a series of private statistical data, namely $\{o_1, o_2, ..., o_i, ..., o_t\}$. Mechanism M is divided into two distinct components: M_i^1 and M_i^2. These two phases, which operate sequentially, each take half of the overall privacy budget, i.e., $\epsilon^1 = \epsilon^2 = \epsilon/2$. Table 2 summarizes the notations used in this section.

The overall process of the LPBD mechanism releasing private statistics is outlined in Algorithm 1, which takes as input (i) total privacy budget ϵ, (ii) real location statistics c_i^r, (iii) last recent private release o_l^p, (iiii) similarity threshold Th, where Th is a defined threshold for deciding whether it is more beneficial to approximate the current statistics to the last private release than to disturb the current statistics.

Algorithm 1 Pseudocode of the overall process of LPBD.

```
Input: $c_i^T, o_l^p, Th, \epsilon$
Output: $private\ statistic\ o_i$
$\epsilon_i^d \leftarrow \epsilon/2$
//** The decision step **//
Calculate the similarity between $c_i^r$ and $o_l^p$ given the
 using privacy budget $\epsilon_i^d$
if $c_i^r$ and $o_l^p$ are similar then
    $o_i^p \leftarrow o_l^p$
else
//** The release step **//
    Consider distinct cases of temporal correlation
    $o_l^i \leftarrow$ Perturb with allocated budget
    return $o_i = c_i^r + noise$
end if
```

5.2.1 The decision phase

The decision phase M^1 is detailed in Algorithm 2. M^1 uniformly distributes the privacy budget from ϵ^1 to each timestamp within the current sliding window. At timestamp i, the sub-mechanism M_i^1 calculates the similarity result between the true statistic c_i^r and the last released private statistic o_l^p. The decision algorithm takes as input: (i) the current location statistic c_i^r, (ii) the latest recent release o_l^p, and (iii) the comparison threshold Th.

The user then uses the privacy budget ϵ_i^d to apply RR mechanism [27] to the similarity results, guaranteeing that the decision phase guarantees LDP. Unlike existing w-event privacy mechanisms, the proposed mechanism conducts local similarity testing on each user. At each timestamp, M_i^1 uses a budget of $\epsilon/2 * w$.

By comparing the statistical data of the current timestamp with the similarity of the previous private release, the decision algorithm determines whether a new private release is required. At each timestamp i, extract a dataset D_i with rows corresponding to specific users and columns corresponding to the total number of attributes for user u. The Mean Absolute Error (MAE) is used to calculate the similarity.

$$MAE = \frac{1}{u}\sum_{j=1}^{u} \|c_i^r[j] - c_l^P j]\| \tag{9}$$

where the vector c_i^r represents the real data obtained from D_i for timestamp i, and l is the private data published at the last released timestamp c_l^p. To make the decision phase M_i^1 satisfy differential privacy, noise needs to be added to the MAE.

Algorithm 2 The decision algorithm M^1.

```
Input: $c_i^r, o_l^p, Th, \epsilon^1$
Output: $Similarity\ result\ \xi_i^{'}$
$\epsilon_i^1 \leftarrow \epsilon/2 * w$
$Calculate\ the\ similarity : Th_i = c_i^r - o_l^p\ (use\ MAE)$
$\xi \leftarrow Th_i > Th$
$\xi_i^{'} = \begin{cases} 0,\ w.p.\ p(\text{w.p.}p = \text{with probability}\,p.) \\ 1,\ w.p.\ p \\ \xi_i,\ w.p.\ 1-2p \end{cases}$
```

5.2.2 The release phase

The release phase M^2 is detailed in Algorithm 3. First, M^2 separates the privacy budget ϵ^2 into the release privacy budget and the reserve privacy budget. Then, at timestamp i, M_i^1 transfers the similarity result to M_i^2. In one scenario, if M_i^2 decides not to release any new private statistics at the current timestamp i, the release privacy budget allotted at this timestamp will be stored and available for upcoming releases if necessary.

On the contrary, if the similarity exceeds the threshold Th, the real location statistics c_i^r are perturbed using the allocated privacy budget ϵ_i^p. Concurrently, the existence of temporal correlation at the current timestamp is determined. If there is a correlation between the current timestamp and the previous timestamp, M_i^2 will use part of the previously stored reserve budget to add to the current release phase.

To illustrate the budget distribution process clearly, Fig. 3 presents the budget allocation of the LPBD mechanism under continuous data release for 5 timestamps, with a sliding window length of w = 3. Suppose M releases new private outputs at timestamps 1, 2, 4, 5, where timestamp 3 releases an approximation of timestamp 2. In phase 1, M^1 allocates a fixed privacy budget at each timestamp. Then, in phase 2, M^2 allocates half of the allocated total privacy budget (i.e., $\sum \epsilon_i^2/2 = \epsilon/4$) in an exponentially decreasing manner within a sliding window of a set length.

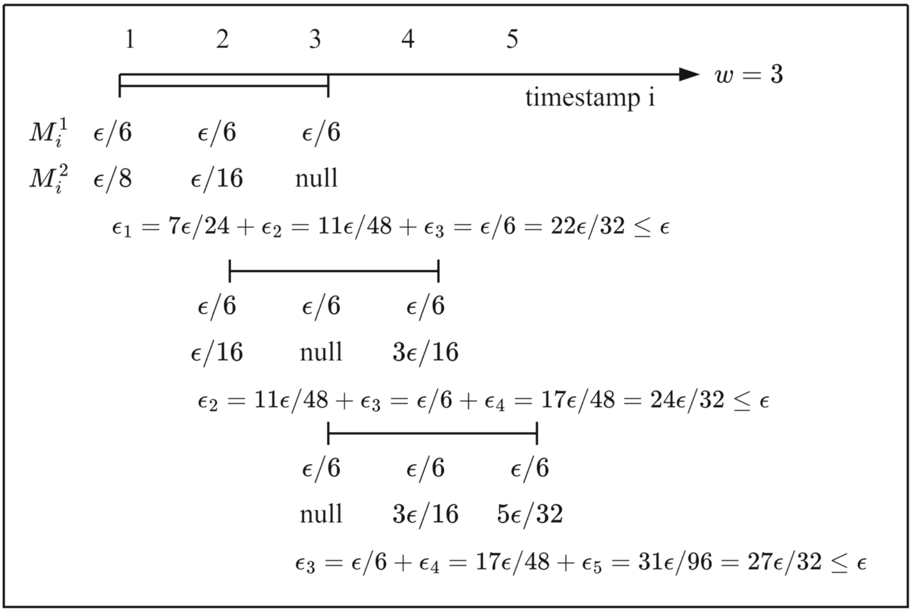

1 2 3 4 5
$w = 3$
timestamp i
M_i^1 $\epsilon/6$ $\epsilon/6$ $\epsilon/6$
M_i^2 $\epsilon/8$ $\epsilon/16$ null
$\epsilon_1 = 7\epsilon/24 + \epsilon_2 = 11\epsilon/48 + \epsilon_3 = \epsilon/6 = 22\epsilon/32 \leq \epsilon$
$\epsilon/6$ $\epsilon/6$ $\epsilon/6$
$\epsilon/16$ null $3\epsilon/16$
$\epsilon_2 = 11\epsilon/48 + \epsilon_3 = \epsilon/6 + \epsilon_4 = 17\epsilon/48 = 24\epsilon/32 \leq \epsilon$
$\epsilon/6$ $\epsilon/6$ $\epsilon/6$
null $3\epsilon/16$ $5\epsilon/32$
$\epsilon_3 = \epsilon/6 + \epsilon_4 = 17\epsilon/48 + \epsilon_5 = 31\epsilon/96 = 27\epsilon/32 \leq \epsilon$

Fig. 3 The distribution process of privacy budget within a sliding window of size w=3

Algorithm 3 The release algorithm M^2

```
Input: $w, c_i^r, \epsilon^2, \xi_i^{'}$
Output: $Release\ Noisy\ output\ o_i$
Calculate remaining budget : $\epsilon_i^2 = (\epsilon/4 - [\sum_{i=t-w+1}^{t-1} \epsilon_j])/2$
if $Similarity > TH$ then
    while $Correlation\ exists$ do
        $\epsilon_A = (\epsilon/4 - [\sum_{i=t-w+1}^{t-1} \epsilon_j^a])$
        Set $\epsilon_i^{re} = \epsilon_i^2$ $(Reserve\ from\ \epsilon_r e)$
        $\epsilon_i^2 = \epsilon_i^{re} + \epsilon_i^2$
        $return\ o_i = c_i^r + Lap(\lambda_i^2)$
    end while
    $return\ o_i = c_i^r + Lap(\lambda_i^2)$
else
    $return\ o_i = o_l$
end if
```

To put it simply, at timestamp 1, it allocates $\epsilon_1^2 = [(\epsilon/4 - 0)]/2 = \epsilon/8$. At timestamp 2, $\epsilon_2^2 = [\epsilon/4 - (0 + \epsilon/8)]/2 = \epsilon/16$. At timestamp 3, $\epsilon_3^2 = 0$ because there is no private output at timestamp 3. Since there is no temporal correlation between the above timestamps, no extra budget should be added. At timestamp 4, $\epsilon_4^2 = [\epsilon/4 - (0 + \epsilon/16)]/2 = 3\epsilon/32$ and add $3\epsilon/32$ part of the reserved budget to ϵ_4^2 due to the correlation between timestamps 4 and 3. Similarly, at timestamp 5, $\epsilon_5^2 = [\epsilon/4 - (0 + 3\epsilon/32)]/2 = 5\epsilon/64$ and adds extra budget $5\epsilon/64$. It should be noted that the cumulative privacy budgets in each of the sliding windows above are less than the total privacy budget ϵ.

5.3 Privacy analysis

In this subsection, we first prove that both the decision phase M^1 and the release phase M^2 satisfy w-event privacy. Subsequently, it is shown that the LPBD mechanism satisfies w-event privacy with the privacy budget ϵ.

Theorem 2 *The decision algorithm M^1 satisfies w-event privacy for $\sum \epsilon_i^1 = \epsilon/2$.*

Proof At timestamp i, let ξ_i be the statistical similarity comparison result of the current timestamp i with the previous release. Let $\xi^{'}$ represent any result of M_i^1 applied to ξ_i, then:

$$\frac{P_r(\xi_i^{'} = 0|\xi_i = 0)}{P_r(\xi_i^{'} = 0|\xi_i = 1)} = \frac{-2 \times p + 1 + p}{p} = \frac{1-p}{p} \tag{10}$$

Likewise available,

$$\frac{P_r(\xi_i^{'} = 1|\xi_i = 1)}{P_r(\xi_i^{'} = 1|\xi_i = 0)} = \frac{1-p}{p} \tag{11}$$

From the above formula, given the observed output $\xi_i^{'}$, the degree of indistinguishability between two potential inputs "0" and "1" is $ln(\frac{1-p}{p})$. Therefore, M_i^1 satisfies $ln(\frac{1-p}{p})$-differential privacy. According to Theorem 1, the privacy budget used by M^1 within a sliding window equals the total of the privacy budgets of M_i^1 such that $i - w + 1 \leq j \leq i$. Therefore, within the sliding window w, the privacy budget consumed by M^1 is as follows:

$$\epsilon/2 = \sum_{j=i-w+1}^{j=i} ln(\frac{1-p}{p}) = w \times ln(\frac{1-p}{p}) \tag{12}$$

where p is $\frac{1}{1+\epsilon^{\frac{\epsilon}{2\times w}}}$. As a result, M^1 satisfies w-event privacy with the privacy budget $\epsilon/2$. □

Theorem 3 *The release algorithm M^2 satisfies w-event privacy for $\sum \epsilon_i^2 = \epsilon/2$.*

Proof The private output of $q(D_i)$ is released by M_i^2, or it outputs null. In the given setting, the maximum change in the outcome of $q(D_i)$ due to the addition or removal of a row from D_i is limited to 1. Consequently, the sensitivity of q is limited to 1.

The M_i^2 introduces laplace noise with a scale of $\lambda_i^2 = 2/(\epsilon/4 - [\sum_{j=i-w+1}^{k-1} \epsilon_j])$. If there is a temporal correlation between the current and previous timestamps, M_i^2 borrows the extra budget from the deliberately reserved $\epsilon/4$. Therefore, M_i^2 is injected with noise of scale $\lambda_i^2 = 2/(\epsilon/4 - [\sum_{j=i-w+1}^{i-1} \epsilon_j]) + 2/(\epsilon/4 - [\sum_{j=i-w+1}^{i-1} \epsilon_j])$ if there is a correlation; otherwise, the noise is $\lambda_i^2 = 2/(\epsilon/4 - [\sum_{j=i-w+1}^{i-1} \epsilon_j])$. It is assumed that mechanism M consumes the entire extra budget ($\epsilon/4$) within a sliding window of length w. According to the definition of the Laplacian mechanism, M_i^2 is ϵ_i^2-local differentially private.

The next step is to demonstrate that LPBD holds within the sliding window for $\sum_{j=i-w+1}^{i-1} \epsilon_j \leq \epsilon$. From composition property Theorem 1, LPBD holds at j^{th} privacy budget is $\epsilon_j = \epsilon_j^1 + \epsilon_j^2$, then it equals to $\sum_{j=i-w+1}^{i} \epsilon_j = \sum_{j=i-w+1}^{i} \epsilon_j^1 + \sum_{j=i-w+1}^{i} \epsilon_j^2$. The total privacy budget for the sliding window is $\sum_{j=i-w+1}^{i} \epsilon_j = \epsilon/2 + \sum_{j=i-w+1}^{i} \epsilon_j^2$ because every ϵ_j^1 is set to $\epsilon/2 \cdot w$. Now, it needs to be proven that $\sum_{j=i-w+1}^{i} \epsilon_j^2 \leq \epsilon/2$. In our settings, $\sum_{j=i-w+1}^{i} \epsilon_j^2$ is $\sum_{j=i-w+1}^{i} (\epsilon/4 - [\sum_{j=i-w+1}^{i-1} \epsilon_j])/2 + \epsilon/4 - [\sum_{j=i-w+1}^{i-1} \epsilon_j])/2$. These two terms can be demonstrated by induction from inequality. We demonstrate that one of the terms is smaller than or equal to $\epsilon/4$ because both terms are equal. Consequently, it can be concluded that another term is also less than or equal to $\epsilon/4$. In the induction part that follows, the term is first simplified using geometric progression techniques, and then mathematical induction is used to prove that the term is less than or equal to $\epsilon/2$. Given this, the sub-mechanism M_i^2 always uses at most half of the privacy budget, or (ϵ/2). □

Theorem 4 *With a budget of ϵ, the LPBD mechanism satisfies w-event privacy.*

Proof Given a data stream S_t. The LPBD (i.e., mechanism M) applied to S_t, with sequential composition of M^1 and M^2. According to Theorem 1, the privacy budget consumed by M is the sum of the privacy budget consumed by phase M^1 and M^2. As shown in Theorems 2 and 3, M^1 and M^2 both satisfy w-event privacy with budget ϵ/2 respectively. Therefore, LPBD satisfies w-event privacy with a budget of ϵ. □

6 Experiments

In this section, a series of experiments is designed to evaluate the performance of the LPBD mechanism.

6.1 Experimental setup

The conducted experiments are all based on public benchmark real-world datasets. These datasets are used to assess the performance of LPBD and different baseline methods under various privacy parameters, including different total privacy budgets ϵ and sliding window lengths w.

Real-world Datasets The performance evaluation of LPBD was conducted using the following real-world location datasets:

1. The $T\text{-}Drive$ dataset [31, 32] contains real-time taxi trajectories in Beijing. It comprises one-week trajectories of 10,357 taxis, with $T = 672$ timestamps (each at the 15-minute level). We obtained $N = 10{,}357$ data streams for each taxi and extended the stream to four weeks to facilitate experiments with larger sliding window lengths w.
2. The $ShangHai$ dataset [33] is a commonly used public trajectory dataset of approximately 5,000 buses and taxis in Shanghai. It was collected by the Hong Kong University of Science and Technology on February 20, 2007, with a data sample interval of approximately 120 seconds.

The location datasets comprise various data sequences, each containing details such as timestamps, vehicle numbers, location latitudes and longitudes, etc. The two datasets differ in terms of data volume, data collection interval, data map area, and other attributes. Consequently, these two datasets were selected to evaluate the efficiency and stability of the LPBD mechanism across various scenarios.

Performance Metrics In our experimental setup, the LPBD mechanism aims to provide privacy protection for each user data point at each timestamp within a sliding window of size w. We evaluate the utility of LPBD and baseline methods using the MAE (as defined in (9)), the Mean Relative Error (MRE), and the Root Mean Square Error (RMSE). These metrics assess the discrepancy (or error) between a privacy statistic and its true value. Furthermore, all these metrics exhibit good mathematical properties. MAE demonstrates better robustness to outliers, while MRE is more sensitive to the raw value of the true position statistic. On the other hand, RMSE is more sensitive to higher error values due to the squaring of errors. The formal descriptions of MRE and RMSE are as follows:

$$MRE = \frac{1}{u}\sum_{j=1}^{u}\frac{\|c_i^r[j] - c_l^P[j]\|}{max\ \{c_i^r[j], c_l^P[j]\}} \tag{13}$$

$$RMSE = \sqrt{\frac{1}{u}\sum_{j=1}^{u}(\|c_i^r[j] - c_l^P[j]\|)^2} \tag{14}$$

The experimental environment All experiments (including the main experimental LPBD and baseline methods) are conducted on a Windows 10 machine with 16 GB RAM and an AMD Core 8 CPU @3.2 GHz. Additionally, experimental plots are drawn using Origin software. The system models for the experiments were developed using the Python and Java programming languages.

6.2 Baseline *w*-event LDP methods

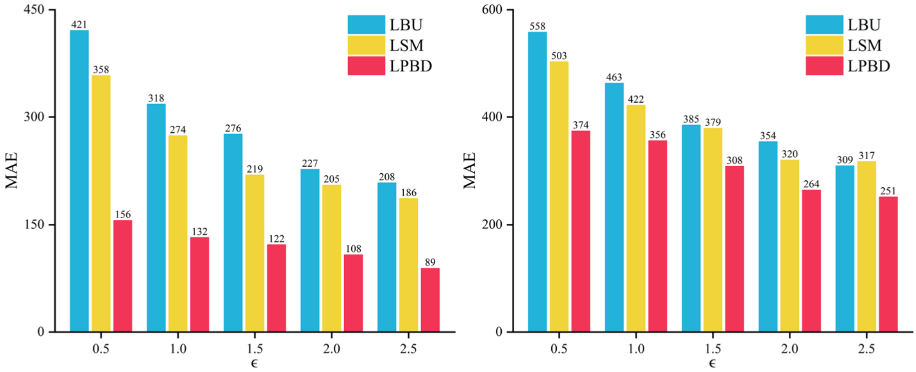
To the best of our knowledge, existing LDP research lacks studies similar to LPBD that specifically address the temporal correlation of user data in local environments. In order to facilitate experimental comparison, we have designed two baseline methods, namely the LDP Budget Uniform Method (LBU) and the LDP Sampling Method (LSM), to evaluate the performance of LPBD.

One straightforward method, LBU, is to uniformly distribute the privacy budget ϵ across all w timestamps in the sliding windows. At each timestamp, each user reports perturbed values using a fixed budget of ϵ/w. Evidently, for any window of size w, the sum of ϵ_i within it is equal to ϵ, thereby satisfying the conditions of Theorem 1 and ensuring w-event privacy. However, a limitation of LBU is that for large values of w, the budget allocated at each timestamp becomes very small, resulting in a large increase in noise scale.

In another baseline method named LSM, each user invests the entire budget ϵ in a single (sampling) timestamp within a window while using an approximate strategy to reserve a budget for the next w - 1 timestamps. However, if the statistics deviate significantly from previous releases, the resulting estimates may be subject to considerable errors. Consequently, for streams with minimal changes, LSMs can demonstrate effective performance by conserving the privacy budget. On the other hand, in streams with substantial changes, the error of those skipped timestamps may become excessive.

6.3 Experiment design and results

Figures 4 to 7 depict the results of the four sets of experiments, along with MAE and MRE results for the three mechanisms. The vertical axis represents the error values, while the hor-



(a) (b)

Fig. 4 MAE vs. ϵ while fixing w = 20 (a) T-Drive (b) ShangHai datasets

izontal axis represents the different query parameters. Tables 3 and 4 present the relationship between RMSE results and privacy parameters.

Utility vs. Privacy budget Figures 4 and 5 illustrate the release accuracy of all compared w-event LDP methods on various real-world datasets with different privacy budget ϵ. The experiments involve varying the budget ϵ from 0.5 to 2.5 in increments of 0.5 while keeping a constant sliding window length w of 20. These experiments were repeated 100 times, and the mean value of the measure was calculated as the result.

From the experimental results, a noticeable trend is observed: as the privacy budget ϵ increases, the data utility improves significantly (i.e., the error becomes smaller), aligning with the principles of differential privacy and the privacy analysis of LPBD, which demonstrates the trade-off between data utility and privacy. Compared to the LPBD mechanism, the baseline method exhibits higher MAE and MRE values. This discrepancy arises due to the fixed privacy budget utilized by the LBA and LSM at each timestamp, disregarding the temporal correlation among location data points. In contrast, the LPBD mechanism allocates an adequate privacy budget even when the location data points are temporally correlated across consecutive timestamps. It is worth noting that MRE is highly responsive to the raw values of the real location statistics, providing further confirmation that LPBD effectively minimizes the disparity between real and private statistics, leading to improved utility.

Table 3 presents the RMSE values of different methods under varying budget values as an auxiliary metric. Comparing Fig. 4, it can be seen that, under the same conditions, the

Table 3 Comparison of RMSE with different privacy budget ϵ

Methods	$T\text{-}drive$[1]				$ShangHai$[2]			
	$\epsilon = 1$	$\epsilon = 1.5$	$\epsilon = 2$	$\epsilon = 2.5$	$\epsilon = 1$	$\epsilon = 1.5$	$\epsilon = 2$	$\epsilon = 2.5$
LBU	1.127	0.933	0.812	0.737	2.013	1.674	1.512	1.347
LSM	0.981	0.784	0.681	0.575	1.841	1.585	1.391	1.275
LPBD	0.465	0.419	0.368	0.297	1.533	1.342	1.146	0.989

Note: The RMSE value is expressed with four significant figures, using scientific notation (10^3).
[1] This represents the $T\text{-}drive$ dataset.
[2] This represents the $ShangHai$ dataset

Table 4 Comparison of RMSE with different window length w

Methods	*T-drive*[1]				*ShangHai*[2]			
	$w = 10$	$w = 20$	$w = 30$	$w = 40$	$w = 10$	$w = 20$	$w = 30$	$w = 40$
LBU	5.216	7.355	9.871	11.76	4.154	7.448	11.33	13.04
LSM	5.018	6.725	8.719	10.12	3.887	6.621	9.563	12.06
LPBD	2.741	3.573	4.218	5.210	3.421	5.664	8.049	9.560

Note: The RMSE value is expressed with four significant figures, using scientific notation (10^3).

[1] This represents the *T-drive* dataset.

[2] This represents the *ShangHai* dataset

RMSE value is slightly greater than the MAE value, respectively. This disparity is due to RMSE first accumulating the squares of the errors and then extracting the square root, which magnifies the influence of larger errors. In contrast, MAE reflects the real error without squaring. Therefore, a smaller RMSE means improved performance because it represents a relatively smaller maximum error.

Utility vs. Sliding window Figures 6 and 7 illustrate the release accuracy of all compared w-event LDP methods with different sliding window length w on all datasets. The experiments involve varying w from 10 to 50 in increments of 10 while keeping a constant privacy budget ϵ set to 1.0. These experiments were also repeated 100 times, and the resulting metric values were averaged.

The analysis of the experimental results reveals a general trend: MAE and RMSE results of all methods increase with w. This effect is attributed to the fact that the window length directly impacts the privacy budget allocated to each timestamp. However, with a continuous increase in the window length w, the privacy utility may not exhibit a straightforward decrease and can potentially remain stable or even improve. The occurrence of this phenomenon depends on the intrinsic characteristics of datasets.

Moreover, LPBD consistently demonstrates lower errors compared to LBA and LSM. This improvement can be attributed to LPBD's capability to allocate a more appropriate and adequate amount of privacy budget to temporal correlation timestamps within the sliding window. In contrast, the baseline method tends to minimize the ratio of privacy budget

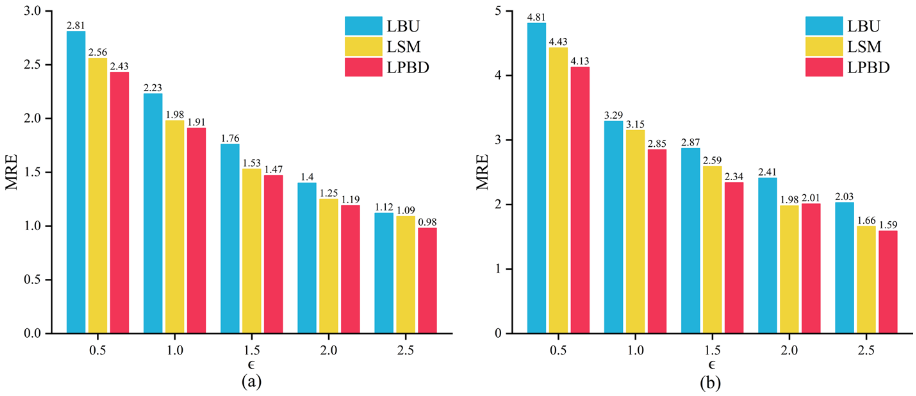


Fig. 5 MRE vs. ϵ while fixing $w = 20$ (a) T-Drive (b) ShangHai datasets

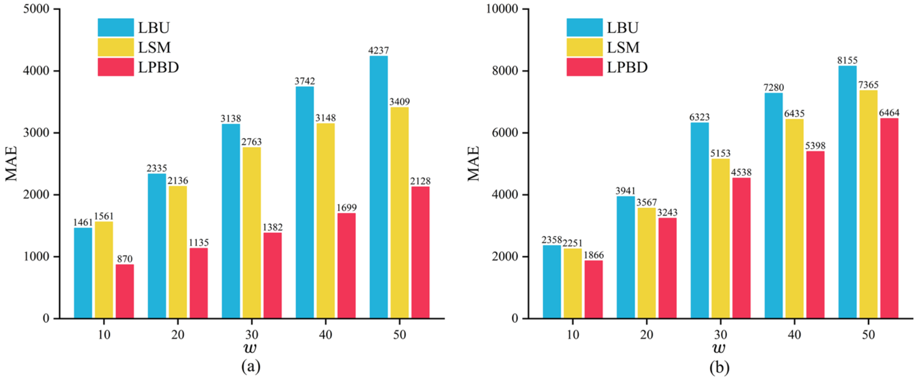


Fig. 6 MAE vs. w while fixing $\epsilon = 1.0$ (a) T-Drive (b) ShangHai datasets

allocation within the sliding window, which results in insufficient protection for temporal correlation location data points.

Table 4 presents the MRE values of different methods with varying w, providing additional validation of the utility improvement achieved by LPBD.

7 Conclusion

This paper proposed a novel privacy-preserving mechanism named LPBD that preserves the privacy of continuously generated streaming location data provided by various users in a local environment. LPBD is allowed to process temporally correlated data through well-designed privacy budget distribution strategies. The experiments have been conducted, and results show that LPBD ensures high accuracy, strong stability, and low error improvement over the baseline methods. The mechanism's performance over the well-known datasets and comparison with baseline methods showed that LPBD is more secure, efficient, and optimal. The future aim of this work is to reduce the consumption of the privacy budget

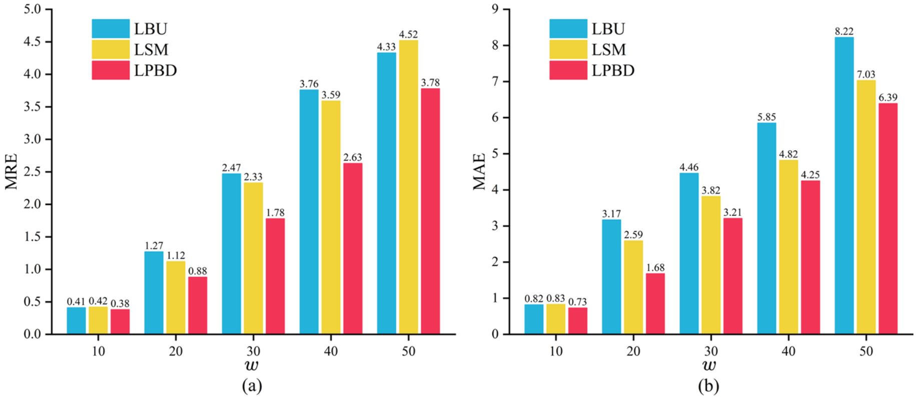


Fig. 7 MRE vs. w while fixing $\epsilon = 1.0$ (a) T-Drive (b) ShangHai datasets

while ensuring privacy protection strength and to expand the scheme to other areas, such as solving the privacy protection scenario of various entities accessing multimedia data.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 61702321)

Data Availability The datasets used to support the findings of this study are available from the corresponding author upon request.

Declarations

Consent to participate This study only base on the theoretical basic research. It is not involving humans.

Consent for publication All authors give consent for the publication of identifable details, which can include photograph(s) and/or videos and/or case history and/or details within the text ("Material") to be published in the Journal of Multimedia Tools and Applications.

Confcts of interest/Competing interests The authors declare that they have no conflict of interest.

References

1. Zhang X, Hamm J, Reiter MK, Zhang Y (2019) Statistical privacy for streaming traffic. In: Proceedings of the 26th ISOC Symposium on network and distributed system security
2. Wang Q, Zhang Y, Lu X, Wang Z, Qin Z, Ren K (2016) Real-time and spatio-temporal crowd-sourced social network data publishing with differential privacy. IEEE Trans Dependable Secure Comput 15(4):591–606
3. Wang T, Hu Z (2021) Real-time stream statistics via local differential privacy in mobile crowdsensing. In: Mobile multimedia communications: 14th EAI international conference, mobimedia 2021, virtual event, proceedings. Springer, pp 432–445
4. Cunningham T, Cormode G, Ferhatosmanoglu H, Srivastava D (2021) Real-world trajectory sharing with local differential privacy. arXiv preprint arXiv:2108.02084
5. Cao Y, Yoshikawa M, Xiao Y, Xiong L (2018) Quantifying differential privacy in continuous data release under temporal correlations. IEEE Trans Knowl Data Eng 31(7):1281–1295
6. Cao Y, Xiong L, Yoshikawa M, Xiao Y, Zhang S (2018) Contpl: controlling temporal privacy leakage in differentially private continuous data release. In: Proceedings of the VLDB endowment. International conference on very large data bases, vol 11. NIH Public Access, p 2090
7. Cao X, Cao Y, Yoshikawa M, Nakamura A (2022) Boosting utility of differentially private streaming data release under temporal correlations. In: 2022 IEEE International conference on big data (big data). IEEE, pp 6605–6607
8. Dwork C, Naor M, Pitassi T, Rothblum GN (2010) Differential privacy under continual observation. In: Proceedings of the forty-second ACM symposium on theory of computing, pp 715–724
9. Xiong X, Liu S, Li D, Cai Z, Niu X (2020) Real-time and private spatio-temporal data aggregation with local differential privacy. Journal of Information Security and Applications 55:102633
10. Wang H, Hong H, Xiong L, Qin Z, Hong Y (2022) L-srr: local differential privacy for location-based services with staircase randomized response. In: Proceedings of the 2022 ACM SIGSAC Conference on computer and communications security, pp 2809–2823
11. Erlingsson Ú, Pihur V, Korolova A (2014) Rappor: randomized aggregatable privacy-preserving ordinal response. In: Proceedings of the 2014 ACM SIGSAC Conference on computer and communications security, pp 1054–1067
12. Chen Y, Machanavajjhala A, Hay M, Miklau G (2017) Pegasus: data-adaptive differentially private stream processing. In: Proceedings of the 2017 ACM SIGSAC Conference on computer and communications security, pp 1375–1388
13. Fan L, Xiong L (2012) Real-time aggregate monitoring with differential privacy. In: Proceedings of the 21st ACM International conference on information and knowledge management, pp 2169–2173
14. Kellaris G, Papadopoulos S, Xiao X, Papadias D (2014) Differentially private event sequences over infinite streams. Proceedings of the VLDB endowment 7(12):1155–1166

15. Wang Z, Liu W, Pang X, Ren J, Liu Z, Chen Y (2020) Towards pattern-aware privacy-preserving real-time data collection. In: IEEE INFOCOM 2020-IEEE Conference on computer communications. IEEE, pp 109–118
16. Ren X, Shi L, Yu W, Yang S, Zhao C, Xu Z (2022) Ldp-ids: local differential privacy for infinite data streams. In: Proceedings of the 2022 International conference on management of data, pp 1064–1077
17. He Y, Wang F, Deng X, Ni J, Feng J, Liu S (2022) Ordinal data stream collection with condensed local differential privacy. In: 2022 IEEE 24th Int conf on high performance computing & communications; 8th Int conf on data science & systems; 20th Int conf on smart city; 8th Int conf on dependability in sensor, cloud & big data systems & application (HPCC/DSS/SmartCity/DependSys). IEEE, pp 562–569
18. Errounda FZ, Liu Y (2021) Collective location statistics release with local differential privacy. Futur Gener Comput Syst 124:174–186
19. Errounda FZ, Liu Y (2018) Continuous location statistics sharing algorithm with local differential privacy. In: 2018 IEEE International conference on big data (big data), pp 5147–5152
20. Xiao Y, Xiong L (2015) Protecting locations with differential privacy under temporal correlations. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, pp 1298–1309
21. Fang R, Han J, Yu J, Yao X, Peng H, Lu J (2021) Differentially private location preservation with staircase mechanism under temporal correlations. In: Proceedings of the 16th EAI International conference. Springer, pp 75–92
22. Chan T-HH, Shi E, Song D (2011) Private and continual release of statistics. ACM Transactions on Information and System Security (TISSEC) 14(3):1–24
23. Fan L, Xiong L (2013) An adaptive approach to real-time aggregate monitoring with differential privacy. IEEE Trans Knowl Data Eng 26(9):2094–2106
24. Wang Q, Zhang Y, Lu X, Wang Z, Qin Z, Ren K (2016) Rescuedp: real-time spatio-temporal crowd-sourced data publishing with differential privacy. In: IEEE INFOCOM 2016-The 35th Annual IEEE International conference on computer communications. IEEE, pp 1–9
25. Joseph M, Roth A, Ullman J, Waggoner B (2018) Local differential privacy for evolving data. Advances in Neural Information Processing Systems 31
26. Hemkumar D, Ravichandra S, Somayajulu DV (2021) Impact of data correlation on privacy budget allocation in continuous publication of location statistics. Peer-to-Peer Networking and Applications 14:1650–1665
27. Dwork C, Roth A et al (2014) The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9(3–4):211–407
28. Li Y, Ren X, Yang S, Yang X (2019) Impact of prior knowledge and data correlation on privacy leakage: a unified analysis. IEEE Trans Inf Forensics Secur 14(9):2342–2357
29. Chen J, Ma H, Zhao D, Liu L (2017) Correlated differential privacy protection for mobile crowdsensing. IEEE Transactions on Big Data 7(4):784–795
30. Rafiei M, Elkoumy G, Aalst WM (2022) Quantifying temporal privacy leakage in continuous event data publishing. In: Cooperative information systems: 28th International conference, CoopIS 2022, proceedings. Springer, pp 75–94
31. Yuan J, Zheng Y, Xie X, Sun G (2011) T-drive: enhancing driving directions with taxi drivers' intelligence. IEEE Trans Knowl Data Eng 25(1):220–232
32. Schäler C, Hütter T, Schäler M (2022) Benchmarking the utility of w-event differential privacy mechanisms–when baselines become mighty competitors
33. Cocchia A (2014) Smart and digital city: a systematic literature review. Smart city: how to create public and economic value with high technology in urban space, 13–43

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