Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, and Chi-Yin Chow, *Senior Member, IEEE*

Abstract—The sequential pattern in the human movement is one of the most important aspects for location recommendations in geosocial networks. Existing location recommenders have to access users' raw check-in data to mine their sequential patterns that raises serious location privacy breaches. In this paper, we propose a new **P**rivacy-preserving **LO**cation **RE**commendation framework (PLORE) to address this privacy challenge. First, we employ the nth-order additive Markov chain to exploit users' sequential patterns for location recommendations. Further, we contrive the probabilistic differential privacy mechanism to reach a good trade-off between high recommendation accuracy and strict location privacy protection. Finally, we conduct extensive experiments to evaluate the performance of PLORE using three large-scale real-world data sets. Extensive experimental results show that PLORE provides efficient and highly accurate location recommendations, and guarantees strict privacy protection for user check-in data in geosocial networks.

Index Terms—Location recommendations, sequential patterns, additive Markov chain, probabilistic differential privacy, noise injection.

1 Introduction

LATEST-generation mobile devices allow users to participate in geosocial networks such as Foursquare, Gowalla and Brightkite. Geosocial networks enable users to share their experiences in visiting specific locations of interest, e.g., restaurants, museums, and stores. Upon physically visiting a location, a user performs a virtual check-in operation in geosocial networks via her handheld device to share her current location with friends. To help users explore new places and shape life, geosocial networks also recommend users with interesting locations based on their preferences learned from their check-in histories. Based on the fact that the human movement exhibits sequential patterns [1], such sequential patterns become increasingly important in location recommendations. For example, the recent studies [2], [3], [4], [5], [6], [7], [8] achieve location recommendations through Markov chain models. A Markov chain first extracts sequential patterns from check-in location sequences of users and then exploits them to derive the transition probability from one location or sequence to a target location, i.e., the probability of a user visiting a target location given her historical check-in location sequence.

The functionality of location recommendations requires a large amount of user check-in locations. However, disclosing visited locations of users to untrusted recommender systems raises serious location privacy breaches, since the history of visited locations can reveal sensitive details about an individual's health status, political views, or lifestyle [9]. To provide formal location privacy guarantees against adversaries with background knowledge, the two pioneer works [8], [10] leverage ε-differential privacy [11] to inject carefully chosen random noise into aggregate statistics of check-in locations from users and only release these noisy aggregate statistics to location recommender systems. The privacy budget ε is *inversely proportional* to the strength of privacy protection and the magnitude of the injected noise is *directly proportional* to **the sensitivity of the underlying aggregate algorithm, i.e., the maximum variety of the aggregate statistics resulting from the addition or removal of a single user's check-in record that includes her all check-in locations**. The ε-differential privacy protects each user by guaranteeing that a single user's check-in record cannot significantly affect the released noisy aggregate statistics; in other words, an adversary cannot learn with significant probability whether a certain user is included into the released information or not.

Specifically, the study [10] adds noise into the aggregate counts of check-ins for each location from all users and simply recommends users with the same popular locations having large noisy counts. Nonetheless, it does *not consider personalization* that is one of the most essential requirements of recommender systems. More sophistically, our previous research [8] injects noise into the aggregate counts of each subsequence occurring in check-in location sequences of all users. Then, it utilizes the classical higher-order Markov chain with noisy counts of subsequences to predict the personalized probability of a user visiting a target location given her historical check-in location sequence. However, due to the prohibitively expensive computational cost of the classical higher-order Markov chain, the work [8] performs location recommendations at *a coarse level of granularity*, i.e., recommending users with coarse-grained areas, e.g., city blocks or zipcode regions, instead of a specific location of interest, e.g., restaurants or museums that would be more attractive to users.

This paper aims to study the *personalized* and *fine-grained* location recommendations using sequential patterns with differential privacy protection. There are mainly two interdependent key challenges: *high recommendation accuracy* and *strict location privacy*. (1) **High recommendation accuracy.** On the one hand, the recommendation accuracy of the non-personalized method [10] is very low, since it returns the same locations to all users. On the other hand, the coarse-grained method [8] also generates inaccurate recommendations for users, because it returns large areas instead of specific locations. Thus, most current works [2], [3], [4], [5], [6] employ the first-order Markov chain by assuming that a newly possible visiting location of a user only relies on her

• *J.-D. Zhang and C.-Y. Chow are with Department of Computer Science, City University of Hong Kong, Hong Kong.*
E-mail: jzhang26@cityu.edu.hk, chiychow@cityu.edu.hk.

latest visited location. Nevertheless, in reality the new location depends on not only the latest visited location but also the earlier visited locations in her check-in history. As a result, the first-order Markov chain often suffers from low recommendation accuracy. (2) **Strict location privacy.** The bigger challenge is to rigorously preserve the location privacy of users without significantly decreasing the recommendation accuracy due to two facts. (a) As a single user may visit many locations, the sensitivity (i.e., the maximum variety of aggregate statistics caused by adding or removing a single user's record) is considerably high and hence large noise is required to inject into the aggregate statistics to provide differential privacy protection. (b) Since users only check in a very small fraction of total locations in a geosocial network, the check-in data of users to locations are highly sparse. Thus, the statistics (e.g., counts of subsequences used by the Markov chain) aggregated from the sparse check-in data consist of a lot of very small values. As a result, the *large noise* usually dominates the *small true aggregate statistics*, which severely deteriorates the recommendation accuracy. Although there are some methods [12], [13] that can publish statistics (e.g., frequent itemsets) on highly sensitive data with differential privacy, they still require a relatively large amount of noise for the small (i.e., non-frequent) statistics from sparse check-in data. Thus, these methods are not appropriate in our settings.

In this paper, we propose a new probabilistic differential **P**rivacy-preserving **LO**cation **RE**commendation framework for geosocial networks to address the two key challenges on *high accuracy* and *strict privacy*, called **PLORE**. (1) **High recommendation accuracy.** To improve the recommendation accuracy, PLORE applies the nth-order additive Markov chain proposed in our recent work [7]. With the aggregate counts of two-gram subsequences extracted from historical check-in location sequences of all users, the additive Markov chain considers all visited locations in the check-in history of a user to derive her visiting probability on new locations, instead of only using her latest visited location adopted by the first-order Markov chain. Therefore, our additive Markov chain can overcome the limitation of the first-order Markov chain and is a much more accurate location recommendation model using users' check-in sequences, as shown in our experiments. (2) **Strict location privacy.** Although we could apply the differential privacy in location recommendations at the high cost of accuracy, as in the non-personalized method [10] or coarse-grained method [8], we use the probabilistic differential privacy approach to achieve the trade-off between *high recommendation accuracy* and *strict location privacy* for personalized fine-grained location recommendations. As aforementioned, the high sensitivity demands a prohibitive amount of noise compared to the true aggregate statistics with small values, which renders the published noisy statistics worthless for location recommendation models, e.g., the additive Markov chain. To strive for a good trade-off between privacy and accuracy, we relax ε-differential privacy into (ε, δ)-probabilistic differential privacy that achieves ε-differential privacy with at least $1 - \delta$ probability, i.e., δ represents the maximum probability of a breach of ε-differential privacy. To accomplish (ε, δ)-probabilistic differential privacy with high probability, PLORE models a distribution on the variety of aggregate statistics caused by the addition or removal of a single user's all check-in locations, determines a lower bound that is larger than the variety with probability at least $1 - \delta$ based on the estimated distribution, and exploits the *lower bound of variety* instead of the sensitivity (i.e., the *maximum variety*) to reduce the noise injected into aggregate statistics. We will show that PLORE guarantees (ε, δ)-probabilistic differential privacy with a low value of δ and a low magnitude of noise that indicates negligible loss of privacy and accuracy.

This study is a significant extension to our previous work [7] by embedding the privacy protection ability in the additive Markov chain developed in [7]. The main contributions of this study can be summarized as follows.

- We investigate the personalized and fine-grained location recommendations which are significantly distinct from and much more challenging than the non-personalized [10] or coarse-grained [8] location recommendations, so the mechanism for differential privacy protection in [8], [10], [12], [13] is not applicable in this paper. Thus, we propose a new probabilistic differential privacy-preserving location recommendation framework. To the best of our knowledge, this work is the first attempt to study the personalized fine-grained location recommendations with the differential privacy protection.
- We find a lower bound of the variety of aggregate statistics from adding or removing a single user's record to diminish the noise added into the aggregate statistics. Accordingly, we achieve strict (ε, δ)-probabilistic differential privacy with the negligible loss of accuracy. We also provide theoretical proofs for the privacy protection of our proposed probabilistic differential privacy approach. (Section 5)
- We conduct extensive experiments to evaluate the performance of PLORE on *accuracy*, *privacy* and *efficiency* using three large-scale real-world data sets collected from Foursquare, Gowalla and Brightkite. Extensive experimental results show that PLORE achieves high recommendation efficiency and accuracy, and strict location privacy. (Sections 6 and 7)

The remainder of this paper is organized as follows. Section 2 highlights related work. Section 3 defines the research problems and introduces the overview of PLORE. We propose the additive Markov chain in Section 4 and the probabilistic differential privacy mechanism in Section 5. In Sections 6 and 7, we present our experiment settings and analyze the performance of PLORE, respectively. Finally, we conclude this paper in Section 8.

2 Related Work

Collaborative filtering is often extended to recommend locations for users by integrating check-in data, social links between users, or geographical coordinates of locations.

Location recommendations. With the rapid growth of geosocial networks like Foursquare, Gowalla and Brightkite, location recommendations became an essential functionality. There are five main categories of location recommendation techniques: *basic collaborative filtering techniques*, *social techniques*, *geographical techniques*, *sequential techniques* and *privacy-preserving techniques*. (1) *Basic collaborative filtering techniques.* Most current studies provide location recommendations by using the basic collaborative filtering techniques on users' check-in data [14]. The collaborative filtering techniques also have been extensively extended to integrate with other information, e.g., social links between users and geographical coordinates of locations. (2) *Social techniques.* Because social friends are more likely to share common interests, social link information in geosocial networks has been widely utilized to measure the similarity between users for the use of collaborative filtering techniques [15], [16], [17], [18], [19], [20]. (3) *Geographical techniques.* Since the geographical

proximity between locations significantly affects the check-in behaviors of users on the locations, the studies [16], [17], [18], [19], [20], [21], [22] assume that if a location is closer to the locations visited by a user or the current location of a user, it is more likely to be visited by the same user. (4) *Sequential techniques.* Based on the fact that the human movement exhibits sequential patterns [1], the sequential patterns are increasingly exploited in location recommendations. Most works extract sequential patterns from check-in location sequences as a transition probability matrix and generate location recommendations using the Markov chain on the transition probability matrix. For example, the works [2], [3], [4], [5], [6] apply the first-order Markov chain and suffer from inaccurate location recommendations due to the strong assumption that a newly possible visiting location of a user only relies on her latest visited location. The classical nth-order Markov chain is limited to the coarse-grained location recommendations due to its relatively high complexity [8]. Our recent papers [7], [23] develop the valid and efficient nth-order additive Markov that considers the dependency of the newly possible visiting location of a user on her all visited locations. (5) *Privacy-preserving techniques.* The functionality of location recommendations needs to access a large amount of check-in histories of users to location that raises serious location privacy breaches, since the history of visited locations can reveal sensitive details about an individual's health status, political affiliations or alternative lifestyle [9]. To the best of our knowledge, when using the historical check-in data to make location recommendations, only two pioneer existing works [8], [10] preserve location privacy of the users participating in the check-in data at the high cost of recommendation accuracy. Specifically, the study [10] simply recommends users with the same popular locations without personalization while our previous research [8] performs location recommendations at a coarse level of granularity, i.e., recommending users with coarse-grained areas, e.g., city blocks or zipcode regions. However, personalization is one of the most essential requirements of recommender systems and it may be more attractive to recommend users with a specific location of interest, e.g., restaurants or museums. To this end, this paper studies the personalized and fine-grained location recommendations using sequential patterns with location privacy protection.

Location privacy. Several existing techniques address location privacy threats. (1) *Cryptography.* The cryptographic approaches are appropriate when one needs to privately retrieve a specific item from a data set [24]. However, location recommendations require a broader set of operations to recommend top-k locations. Furthermore, cryptographic methods are very expensive computationally. (2) *K-anonymity.* The K-anonymity techniques ensure that each published location must be indistinguishable among other $K-1$ locations [25]. Similarly, in a published trajectory database LK-privacy requires each sequence with a maximum length of L to be shared by at least K records. For example, the work [26] applies LK-privacy to anonymize trajectory data of passengers for flow analysis. This type of transformation is fast and simple, but it is vulnerable to background knowledge attacks [9], [27], [28]. This is particularly a problem in geosocial networks, since check-in histories can be used to derive the identities behind reported locations. (3) *Differential privacy.* Differential privacy [11] provides mathematically rigorous privacy guarantees against adversaries with background knowledge by only allowing aggregate queries on data and adding noise to each answer to achieve protection. An adversary cannot learn with significant probability whether a certain individual is included in the data set or not, but it is still possible to utilize the noisy aggregate statistics about the data to conduct data mining. Differential privacy has been used in publication on spatial data [28] or sequence data [27], [29], search log sanitization [30], data partitioning [31], frequent itemset mining on transaction data [32], hierarchical histograms for answering range queries [33], classification models [34], social recommendations [35], and recent location recommendations [8], [10], [36]. In these works, the real counts are usually large and the sensitivity is low. For example, the work [29] publishes the transit data of passengers from the Montreal transportation system. The system only includes 68 metro stations and 944 bus stations and passengers regularly transit among stations. Thus, the transit counts of passengers among stations are relatively large. The recent study [36] adds noise into the check-in counts of individual locations rather than the check-in transitions between locations; the individual counts are much larger than the check-in transition counts. However, these existing methods cannot be applied to the release of the user check-in transitions studied in this paper, since the transition counts are often small and the sensitivity is high. (4) *Geo-indistinguishability.* Some works [9], [37], [38] exploit ε-geo-indistinguishability or its variants to directly obfuscate geographical coordinates of locations instead of adding noise into aggregate statistics of locations. The ε-geo-indistinguishability transforms a real location into another location drawn at random from a geographical space based on the differential privacy. These methods are not suitable for location recommendations, in which the exact locations need to be returned for users. (5) *(ε, δ)-differential privacy.* A well-known variant of ε-differential privacy is (ε, δ)-differential privacy (or called (ε, δ)-indistinguishability) that has a small amount of privacy loss due to a slight relaxation of the constraint on the output distributions of two neighboring data sets. The two researches [39], [40] apply the (ε, δ)-differential privacy to mine frequent location patterns from trajectory data. (6) *Probabilistic differential privacy.* Although some studies [12], [13] can deal with the problem of publishing large statistics on highly sensitive data, they still require too large noise for some applications with a lot of small statistics, e.g., the user check-in transition counts. Thus, differential privacy could be too strong to be practically achievable for the applications with high sensitivity and small counts. For this purpose, it is usually relaxed to probabilistic differential privacy that permits a low breach probability of differential privacy. Note that (ε, δ)-probabilistic differential privacy is not a variant of (ε, δ)-differential privacy; it has shown that (ε, δ)-probabilistic differential privacy is stronger than (ε, δ)-differential privacy or (ε, δ)-indistinguishability [41]. In our work, to deal with the high sensitivity and small count problems in the context of personalized and fine-grained location recommendations, we employ probabilistic differential privacy to preserve location privacy of users.

3 Problem Statement and Overview

We introduce the preliminaries in Section 3.1, and present the research problems and overview of PLORE in Section 3.2. TABLE 1 lists the key symbols in this paper.

3.1 Preliminaries

3.1.1 Additive Markov Chain

Definition 1. Check-in location sequence and transition. A check-in location sequence of user u is denoted by $S_u = \langle l_n \rightarrow$

TABLE 1
Key notations defined in this paper

Notation	Description
L	Set of all locations in a geosocial network
l	Some location and $l \in L$
u	Some user
n	Number of locations visited by a user or the nth-order
S_u	Sequence of locations visited by user u: $S_u = \langle l_n \rightarrow \cdots \rightarrow l_2 \rightarrow l_1 \rangle$, where l_i is the i**th-latest** visited location
$l_i \rightarrow l_j$	Two-gram subsequence or transition from l_i to l_j
$\Pr(l\|u)$	Transition probability of u visiting l after S_u
$g(l, l_i, i)$	Contribution of location $l_i \in S_u$ to $\Pr(l\|u)$
$n_{\max}$	Maximum constraint of transitions
$\mathcal{D}$	Data set of historical check-in sequences from all users
$\mathcal{C}$	Counts of each transition $l_i \rightarrow l_j$ occurring in $\mathcal{D}$
$\mathcal{B}$	Weighted sum of some entries in $\mathcal{C}$
$\mathcal{A}$	Random algorithm
$\mathcal{A}(\mathcal{D})$	Output of $\mathcal{A}$ over $\mathcal{D}$
X	Noise random variable
Y	Variety random variable
Z	Auxiliary random variable ($Z > Y$)
ε	Privacy budget ($\varepsilon > 0$); the smaller, the stricter.
δ	Breach probability of ε-differential privacy ($0 \leq \delta < 1$)
$\Delta(\delta)$	Lower bound variety of Y meeting Equation (11)

$\cdots \rightarrow l_2 \rightarrow l_1 \rangle$ such that user u visits locations from l_n to l_1 orderly, in which each two-gram subsequence $l_i \rightarrow l_j$ is also called a transition representing u visiting l_i directly before l_j.

Note that l_i in S_u represents the i**th-latest** visited location by user u for the sake of presentation. The check-in sequence of a user can be easily obtained by ordering her all check-in locations based on check-in time.

Definition 2. General additive Markov chain. Given user u's check-in location sequence $S_u = \langle l_n \rightarrow \cdots \rightarrow l_2 \rightarrow l_1 \rangle$, the nth-order additive Markov chain generally defines the transition probability of user u visiting a target location l after S_u by

$$\Pr(l|u) = \sum_{i=1}^{n} g(l, l_i, i), \tag{1}$$

where $g(l, l_i, i)$ is the additive contribution of the location l_i to the transition probability $\Pr(l|u)$.

In Definition 2, the essential task is to determine the additive contribution of each visited location to the transition probability for a specific application.

3.1.2 Probabilistic Differential Privacy

Definition 3. Check-in record or record for short. A check-in record means a single user's complete check-in history, i.e., including her all check-in locations.

Definition 4. Neighboring data sets. Two data sets $\mathcal{D}_1$ and $\mathcal{D}_2$ are neighboring, if $\mathcal{D}_1$ and $\mathcal{D}_2$ differ in only one record, written as $|\mathcal{D}_1 - \mathcal{D}_2| = 1$. That is, the record is present in only one of the two data sets.

Definition 5. ε-differential privacy [11]. A randomized algorithm $\mathcal{A}$ over data sets satisfies ε-differential privacy, if for any neighboring data sets $\mathcal{D}_1$ and $\mathcal{D}_2$ and any output $o \in Range(\mathcal{A})$

$$\left| \ln \left(\frac{\Pr[\mathcal{A}(\mathcal{D}_1) = o]}{\Pr[\mathcal{A}(\mathcal{D}_2) = o]} \right) \right| \leq \varepsilon, \tag{2}$$

where the probability is taken over $\mathcal{A}$'s randomness.

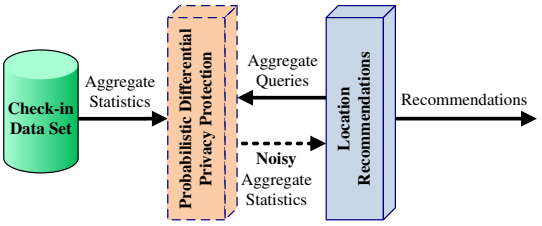


Fig. 1. The overview of PLORE

Differential privacy formally guarantees that no individual record can significantly affect the output of $\mathcal{A}$, i.e., the output distribution is nearly the same whether that record is present in the data set or not. Consequently, an adversary cannot significantly learn more information than her background knowledge using the released output of $\mathcal{A}$. In other words, for a record owner any privacy breach will not be a result of participating in the data set.

Definition 6. (ε, δ)-probabilistic differential privacy [41]. A randomized algorithm $\mathcal{A}$ satisfies (ε, δ)-probabilistic differential privacy, if for all data sets $\mathcal{D}$ there exists the output space $\Omega \subseteq Range(\mathcal{A})$ such that

$$\Pr[\mathcal{A}(\mathcal{D}) \in \Omega] \geq 1 - \delta,$$

and for any neighboring data sets $\mathcal{D}_1$ and $\mathcal{D}_2$ and any $o \in \Omega$

$$\left| \ln \left(\frac{\Pr[\mathcal{A}(\mathcal{D}_1) = o]}{\Pr[\mathcal{A}(\mathcal{D}_2) = o]} \right) \right| \leq \varepsilon.$$

This definition guarantees that (ε, δ)-probabilistic differential privacy achieves ε-differential privacy with at least $1 - \delta$ probability, namely, δ represents the maximum probability of a breach of ε-differential privacy. Note that (ε, δ)-probabilistic differential privacy is a stronger notion than (ε, δ)-indistinguishability [41].

3.2 Problem Definitions and Overview

The two correlated problems are defined as follows.

Problem 1: Location recommendations. Given the data set $\mathcal{D}$ of check-in sequences from all users and a certain user u's check-in sequence $S_u = \langle l_n \rightarrow \cdots \rightarrow l_2 \rightarrow l_1 \rangle$, the goal is to predict the transition probability $\Pr(l|u)$ of user u visiting any target location $l \in L$ after S_u, based on the aggregate statistics extracted from the check-in data set $\mathcal{D}$, and then return the top-k *new locations* with the highest probability $\Pr(l|u)$ for u. (Section 4)

The essential differences between location predictions and location recommendations are that: (1) Location predictions aim to learn the regular moving patterns of users from trajectory data and then predict the next location of users given their current location; the predicted location may have been visited by the users before and may not be relevant to the personalized preference of users, e.g., home or office; (2) Location recommendations focus on learning the personalized preference of users from their check-in data and then recommending new locations for the users; the recommended location has never been visited by the users before and should be relevant to the user preference as much as possible.

Problem 2: Probabilistic differential privacy protection. To provide (ε, δ)-probabilistic differential privacy for users participating in the check-in data set $\mathcal{D}$, we inject carefully chosen random noise into the aggregate statistics mined from the check-in data set $\mathcal{D}$ and only release the noisy aggregate statistics for the use in location recommendations. (Section 5)

User	Sequence
u_1	$l_b \rightarrow l_c \rightarrow l_a$
u_2	$l_a \rightarrow l_b \rightarrow l_c$
u_3	$l_c \rightarrow l_b$
u_4	$l_b \rightarrow l_c \rightarrow l_d$

(a) Check-in data set $\mathcal{D}$

$\Longrightarrow$

	l_a	l_b	l_c	l_d
l_a	0	1	0	0
l_b	0	0	3	0
l_c	1	1	0	1
l_d	0	0	0	0

(b) Transition counts $\mathcal{C}$

Fig. 2. An example of aggregating transition count matrix $\mathcal{C}$ from check-in data set $\mathcal{D}$

Overview of PLORE. Fig. 1 depicts the overview of PLORE for the correlations between the two modules of *location recommendations* and *probabilistic differential privacy protection*. The module of probabilistic differential privacy protection (residing with the owner of check-in data) extracts aggregate statistics from the check-in data set, injects random noise into the aggregate statistics based on the proposed differentially private method in Section 5, receives aggregate queries from the module of location recommendations, and answers it with the noisy aggregate statistics. The location recommendations module then collects the noisy aggregate statistics and makes recommendations for users based on the developed additive Markov chain in Section 4.

In PLORE, users are not allowed to directly send queries in order to save the privacy budget. The location recommendations module will send only one query to the differential privacy module for each user based on her own check-in sequence to generate the recommendations for the user.

4 Location Recommendation Model

In this section, we assume the location recommendation module can directly access to the check-in data set in order to extract aggregate statistics. Section 4.1 prepares aggregate statistics for the additive Markov chain to predict transition probabilities in Section 4.2.

4.1 Aggregating Statistics from Check-in Data Set

Aggregating transition counts $\mathcal{C}$. To prepare sequential patterns for the Markov chain model to derive the transition probability from one location or sequence to a target location, we model the sequential patterns as the aggregate counts of transitions that are extracted from a set of check-in location sequences of all users. Let $\mathcal{C}(l_i \rightarrow l_j)$ be the aggregate count of transition $l_i \rightarrow l_j$. It is easy to acquire a transition count matrix $\mathcal{C}$ for each pair $l_i, l_j \in L$ by scanning the data set $\mathcal{D}$ of check-in location sequences from all users.

Example. Fig. 2 depicts the process of aggregating transition count matrix $\mathcal{C}$ from check-in data set $\mathcal{D}$. For example, since the two-gram subsequence $l_b \rightarrow l_c$ has occurred three times in $\mathcal{D}$, the count of transition $l_b \rightarrow l_c$ is 3, i.e., $\mathcal{C}(l_b \rightarrow l_c) = 3$.

It is worth emphasizing that: (1) The developed location recommendation model only uses these very basic aggregate statistics, i.e., counts of two-gram subsequences or transitions instead of the raw check-in data, which is the first requirement to protect personal location privacy. (2) The ε-differential privacy has been widely used to protect the standard aggregate counts that are extracted from dense data [12], [13], [27], [28]. However, in the context of location recommendations, these transition counts are aggregated from highly sparse check-in data since users only check in a little fraction of locations in a geosocial network; they consist of a large amount of small values, e.g., 0 or 1 as shown in Fig. 2(b). Thus, a little noise still dominates the true signal in the transition counts. (3) The n-gram ($n > 2$) models are not applicable in this work, because most counts of two-grams are zero, not to mention the counts of n-grams. We will deeply discuss how to protect these transition counts with small values from adversaries based on (ε, δ)-probabilistic differential privacy in Section 5.

4.2 Predicting Transition Probabilities

Given user u's check-in location sequence $S_u = \langle l_n \rightarrow \cdots \rightarrow l_2 \rightarrow l_1 \rangle$, we aim to predict the transition probability $\Pr(l|u)$ of user u visiting any target location $l \in L$ after S_u. We will use the transition count matrix $\mathcal{C}$ that is aggregated from check-in data set $\mathcal{D}$ in Section 4.1.

First-order Markov chain as a baseline. The current works [2], [3], [4], [5], [6] derive the transition probability by employing the first-order Markov chain,

$$\Pr(l|u) = \frac{\mathcal{C}(l_1 \rightarrow l)}{\sum_{l \in L} \mathcal{C}(l_1 \rightarrow l)} \propto \mathcal{C}(l_1 \rightarrow l), \forall l \in L. \quad (3)$$

The first-order Markov chain assumes that the probability of user u visiting a target location l only relies on her latest visited location l_1 in S_u. Nevertheless, in reality the transition probability may depend on all her visited locations $l_n, \ldots, l_2, l_1$ in the sequence S_u.

The proposed nth-order additive Markov chain. In our recent study [7], we are inspired to utilize the nth-order Markov chain to improve accuracy. Unfortunately, it is prohibitively expensive to apply the classical nth-order Markov chain, because its number of states $O(|L|^{n+1})$ increases exponentially with the order n, where $|L|$ is the total number of locations in a geosocial network and is usually very large. To this end, we contrive an efficient nth-order additive Markov chain based on Definition 2. The developed additive Markov chain only exploits the transition count matrix $\mathcal{C}$ and significantly reduces the number of states from $O(|L|^{n+1})$ to $O(|L|^2)$.

Determining additive contribution in additive Markov chain. To develop a concrete additive Markov chain based on Definition 2, the essential task is to determine the additive contribution $g(l, l_i, i)$ of the location l_i to the transition probability $\Pr(l|u)$ for the specific location recommendation application. As done in the first-order Markov chain in Equation (3), the additive contribution $g(l, l_i, i)$ is also computed based on the transition count $\mathcal{C}(l_i \rightarrow l)$,

$$g(l, l_i, i) \propto 2^{-\alpha i} \cdot \mathcal{C}(l_i \rightarrow l), \forall l \in L, \quad (4)$$

where $2^{-\alpha i}$ represents the sequence decay weight with the decay rate parameter $\alpha \geq 0$ and the larger α is, the higher is the decay rate. In Equation (4), the transition count $\mathcal{C}(l_i \rightarrow l)$ is weighed by the sequence decay weight $2^{-\alpha i}$ by leaning towards recently visited locations in S_u (i.e., l_i with small i), since the locations with recent check-in timestamps usually have stronger influence on a newly possible visiting location than the locations with old timestamps [1].

Implementation of additive Markov chain. Eventually, given user u's check-in location sequence $S_u = \langle l_n \rightarrow \cdots \rightarrow l_2 \rightarrow l_1 \rangle$, we can derive the transition probability of user u visiting any target location $l \in L$ after S_u, based on the general additive Markov chain defined in Equation (1) and the additive contribution given by Equation (4):

$$\Pr(l|u) \propto \sum_{i=1}^{n} 2^{-\alpha i} \cdot \mathcal{C}(l_i \rightarrow l), \forall l \in L. \quad (5)$$

In terms of Equation (5), we can recommend the new locations with top-k $\Pr(l|u)$ for u.

5 Probabilistic Differential Privacy

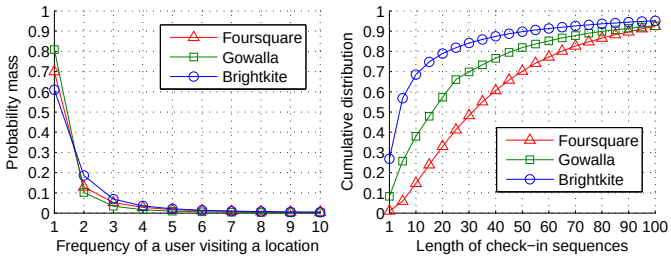
To prevent aggregate transition counts $\mathcal{C}$ of the check-in data set $\mathcal{D}$ of users from adversaries, the probabilistic differential privacy protection module is placed between the check-in data set and the location recommendations module as depicted in Fig. 1. We first describe how to incorporate the constrained maximum contribution of users into the aggregate transition counts $\mathcal{C}$ in Section 5.1, and then present the probabilistic differential privacy-preserving location recommendation framework consisting of a set of compositions that use the modified transition counts $\mathcal{C}$ with noise in Section 5.2. Further, the noise injection mechanism is developed in Section 5.3. Finally, Section 5.4 deals with the issue caused by compositions (i.e., the outputs of compositions may be taken together for joint analysis).

5.1 Constrained Maximum Contribution of Users to Aggregate Statistics

As aforementioned in Section 4.1, most users only check in a little fraction of locations in a geosocial network and contribute a little to aggregate statistics, i.e., the transition counts $\mathcal{C}$ usually have small values. At the same time, there always exist some outlier users who contribute a large number of check-in locations to the transition counts $\mathcal{C}$. As a result, it is required to inject large noise into the relatively small aggregate statistics so as to tackle the outlier or extreme cases at the cost of the utility of noisy outputs. Actually, the large contribution of the outlier users to aggregate statistics offers a little benefit for location recommendation accuracy when without privacy protection since the outlier users are not common; however, this little benefit is far from enough to compensate the loss of recommendation accuracy due to the added large noise when considering privacy protection. **In sum, the large contribution of outlier users is trivial for location recommendations, but it causes a high magnitude of noise for privacy protection and finally degrades recommendation accuracy.** Hence, we propose to constrain the maximum contribution of users to aggregate statistics in order to decrease required noise for the same strength of location privacy protection and then improve recommendation accuracy.

Two constraints for computing aggregate statistics $\mathcal{C}$ from check-in data set $\mathcal{D}$ compared to Section 4.1. When aggregating the transition count matrix $\mathcal{C}$ from the check-in data set $\mathcal{D}$ in Section 4.1, it has the following two constraints. (1) **One-time constraint of transitions.** A user may check in the same location many times, but only the latest transition to the same location is considered. That is, a user cannot contribute two transitions from different sources to the same destination. This constraint is important to determine the lower bound of variety, as shown in Section 5.3.2. (2) **Maximum constraint $n_{\max}$ of transitions.** For each user's check-in sequence, only the $n_{\max}$ differently latest transitions are taken into account. For example, in the works [8], [27], the maximum constraint is set to $n_{\max} = 20$.

Facts regarding the two constraints. To validate the reality of the two constraints, we analyze the check-in behaviors of users on three publicly available large-scale real data sets collected from Foursquare for 7 months [15], Gowalla for 21 months and Brightkite for 19 months [1]. Fig. 3 depicts the distributions of check-in frequency and length on the three real-world data sets.



(a) Check-in frequency (b) Check-in length

Fig. 3. Verification of the two constraints in real data

(1) **Check-in frequency.** According to Fig. 3(a), users check in the same location only once with the probability of around 0.7, 0.8 and 0.6 in Foursquare, Gowalla and Brightkite, respectively. (2) **Check-in length.** In terms of Fig. 3(b), the length of check-in sequences is less than 100 with the probability of around 0.9 on the three data sets. In sum, the two constraints are in accordance with objective reality and most check-ins meet the two constraints when the maximum constraint of transitions is $n_{\max} = 100$. Accordingly, the two constrains will not significantly affect the resulting in aggregate statistics.

5.2 PLORE Framework

True aggregate statistics. To recommend the top-k new locations with the highest probability $\Pr(l|u)$ for user u, the proposed location recommendation model given by Equation (5) in Section 4.2 needs to compute a true aggregate statistic from check-in data set $\mathcal{D}$ for each new location $l \in L$, i.e.,

$$\mathcal{B}_l(\mathcal{D}) = \sum_{i=1}^{n} 2^{-\alpha i} \cdot \mathcal{C}(l_i \to l), \forall l \in L, \tag{6}$$

where $\mathcal{C}$ is a transition count matrix for each pair of locations in L and is aggregated from $\mathcal{D}$ as discussed in Section 4.1, with the two constraints defined in Section 5.1. Thus, it is essential to keep the true aggregate statistics $\mathcal{B}_l(\mathcal{D})$ (not to mention the check-in location sequence S_u) from adversaries residing in untrusted recommender systems so as to protect the users participating in the check-in data set $\mathcal{D}$.

Privacy protection for aggregate statistics. To protect the true aggregate statistics $\mathcal{B}_l(\mathcal{D})$ from adversaries, we contrive a random algorithm $\mathcal{A}$ based on (ε, δ)-probabilistic differential privacy that provides mathematically rigorous privacy guarantees against adversaries with background knowledge and does not assume what kinds of background knowledge an adversary has. In general, our differentially private random algorithm $\mathcal{A}$ can be achieved through a noise adding mechanism to $\mathcal{B}_l(\mathcal{D})$,

$$\mathcal{A}_l(\mathcal{D}) = \mathcal{B}_l(\mathcal{D}) + X, \forall l \in L, \tag{7}$$

where (a) the distribution of the noise random variable X is carefully chosen according to the distribution of the variety of $\mathcal{B}_l(\mathcal{D})$ caused by the addition or removal of a single user's check-in record from $\mathcal{D}$, as presented in Section 5.3, and (b) the random algorithm $\mathcal{A}$ is decomposed into a set of compositions $\mathcal{A}_l, \forall l \in L$, as discussed in Section 5.4.

Accordingly, rather than publishing the true aggregate statistics $\mathcal{B}_l(\mathcal{D})$, the noisy outputs $\mathcal{A}_l(\mathcal{D})$ of the random algorithm $\mathcal{A}$ over data set $\mathcal{D}$ are released to untrusted recommender systems for making location recommendations. In practice, we cannot directly inject the noise X into $\mathcal{B}_l(\mathcal{D})$, because $\mathcal{B}_l(\mathcal{D})$ is dynamic and computed for each individual user u. Instead, we add the noise X

into each entry of transition count matrix $\mathcal{C}$. Subsequently, we can rewrite Equation (5) into

$$\begin{aligned}\Pr(l|u) &\propto \mathcal{A}_l(\mathcal{D}) = \mathcal{B}_l(\mathcal{D}) + X \\ &\propto \sum_{i=1}^{n} 2^{-\alpha i} \cdot [\mathcal{C}(l_i \to l) + X], \forall l \in L. \end{aligned} \quad (8)$$

It is worth emphasizing that: Only the output of $\mathcal{A}$ is available to untrusted recommender systems. Differential privacy formally ensures that the output of $\mathcal{A}$ is insensitive to the variety of any single user's record, i.e., the output distribution is nearly the same whether or not that record is present in the data set. Consequently, with the output of $\mathcal{A}$, an adversary cannot significantly learn more information than her background knowledge. That is, for a record owner any privacy breach will not be a result of participating in the data set $\mathcal{D}$.

5.3 Noise Injection

5.3.1 Mechanism for Probabilistic Differential Privacy

Why not using ε-differential privacy. In general, to accomplish ε-differential privacy, the noise X in Equation (7) or (8) can be drawn from a Laplace distribution with the probability density function $f_X(x|b) = \frac{1}{2b}\exp\left(-\frac{|x|}{b}\right)$, where the scale parameter $b = \frac{\Delta_{\max}}{\varepsilon}$, ε is the privacy budget and $\Delta_{\max}$ is the *sensitivity* of $\mathcal{B}_l(\mathcal{D})$, i.e., the *maximum variety* of $\mathcal{B}_l(\mathcal{D})$ resulting from the addition or removal of a single user's check-in record from check-in data set $\mathcal{D}$. Unfortunately, it is unfeasible to straightforwardly apply this Laplace mechanism in location recommendations because of two reasons. (1) **A large amount of small values in aggregate statistics.** Most users actually check in a little fraction of locations in a geosocial network and then contribute a little to $\mathcal{B}_l(\mathcal{D})$. In other words, most $\mathcal{B}_l(\mathcal{D})$ are aggregated from highly sparse check-in data and hence have many small values (e.g., 0 or 1 as shown in Fig. 2(b)). (2) **Relatively high sensitivity.** The sensitivity of $\mathcal{B}_l(\mathcal{D})$ is still high compared to the aggregate statistics with small values, although we have constrained the maximum contribution of users to aggregate statistics to some extent in Section 5.1. As a result, the noise required by the high sensitivity will dominate the true aggregate statistics $\mathcal{B}_l(\mathcal{D})$, which makes the published noisy statistics $\mathcal{A}_l(\mathcal{D})$ worthless and severely deteriorates the recommendation accuracy.

Using (ε, δ)-probabilistic differential privacy. To this end, we strive for a good trade-off between *privacy* and *accuracy* by relaxing ε-differential privacy into (ε, δ)-probabilistic differential privacy that achieves ε-differential privacy with probability at least $1-\delta$, i.e., having a breach of ε-differential privacy with probability at most δ. At first, we devise a mechanism to implement (ε, δ)-probabilistic differential privacy as proved in Theorem 1.

Theorem 1. Mechanism for (ε, δ)-probabilistic differential privacy. Let Y be the random variable on the variety of $\mathcal{B}_l(\mathcal{D})$ in Equation (6) over two random neighboring data sets $\mathcal{D}_1$ and $\mathcal{D}_2$, i.e.,

$$Y = |\mathcal{B}_l(\mathcal{D}_1) - \mathcal{B}_l(\mathcal{D}_2)|, \text{ where } |\mathcal{D}_1 - \mathcal{D}_2| = 1. \quad (9)$$

The proposed random algorithm $\mathcal{A}_l(\mathcal{D})$ in Equation (7) satisfies (ε, δ)-probabilistic differential privacy, if the noise random variable X in Equation (7) has the following Laplace probability density function

$$f_X(x|\Delta(\delta)) = \frac{\varepsilon}{2\Delta(\delta)}\exp\left(-\frac{\varepsilon|x|}{\Delta(\delta)}\right), \quad (10)$$

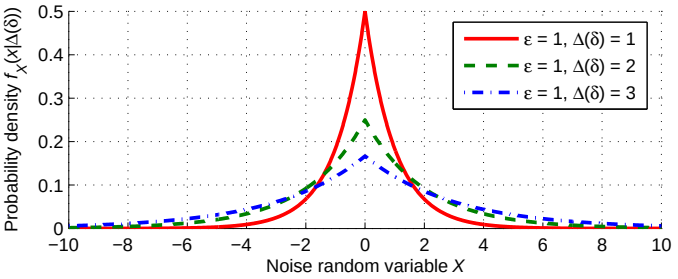


Fig. 4. Distribution of noise with respect to $\Delta(\delta)$

where $\Delta(\delta)$ denotes a function of δ and

$$\Pr[Y \le \Delta(\delta)] \ge 1 - \delta. \quad (11)$$

Proof: In terms of Equations (7) and (10), we have: $\forall o \in Range(\mathcal{A}_l(\mathcal{D}))$,

$$\begin{aligned}\Pr[\mathcal{A}_l(\mathcal{D}_1) = o] &= \Pr[\mathcal{B}_l(\mathcal{D}_1) + X = o] \\ &= \Pr[X = o - \mathcal{B}_l(\mathcal{D}_1)] \\ &\propto \frac{\varepsilon}{2\Delta(\delta)}\exp\left(-\frac{\varepsilon|o - \mathcal{B}_l(\mathcal{D}_1)|}{\Delta(\delta)}\right).\end{aligned}$$

For the same reason,

$$\Pr[\mathcal{A}_l(\mathcal{D}_2) = o] \propto \frac{\varepsilon}{2\Delta(\delta)}\exp\left(-\frac{\varepsilon|o - \mathcal{B}_l(\mathcal{D}_2)|}{\Delta(\delta)}\right).$$

Hence,

$$\begin{aligned}&\left|\ln\left(\frac{\Pr[\mathcal{A}_l(\mathcal{D}_1) = o]}{\Pr[\mathcal{A}_l(\mathcal{D}_2) = o]}\right)\right| \\ &= \left|\frac{\varepsilon}{\Delta(\delta)}|o - \mathcal{B}_l(\mathcal{D}_1)| - \frac{\varepsilon}{\Delta(\delta)}|o - \mathcal{B}_l(\mathcal{D}_2)|\right| \\ &\le \frac{\varepsilon}{\Delta(\delta)}|\mathcal{B}_l(\mathcal{D}_1) - \mathcal{B}_l(\mathcal{D}_2)| \quad \text{(Triangle Inequality)} \\ &\le \frac{\varepsilon}{\Delta(\delta)}Y. \quad \text{(From Equation (9))}\end{aligned}$$

Thus, when $Y \le \Delta(\delta)$, $\mathcal{A}_l(\mathcal{D})$ satisfies ε-differential privacy according to Definition 5. Further, in terms of Equation (11), $Y \le \Delta(\delta)$ holds with at least $1-\delta$ probability, i.e., $\mathcal{A}_l(\mathcal{D})$ satisfies ε-differential privacy with at least $1-\delta$ probability. Therefore, $\mathcal{A}_l(\mathcal{D})$ satisfies (ε, δ)-probabilistic differential privacy according to Definition 6 and Theorem 1 holds. □

The lower bound of variety $\Delta(\delta)$ in Theorem 1. In terms of Theorem 1, the noise X in Equation (7) or (8) can be drawn from the distribution defined by Equation (10), as depicted in Fig. 4. As $\Delta(\delta)$ decreases, the drawn noise is close to 0 with a larger probability. Thus, here **the key or goal is to determine the lower bound $\Delta(\delta)$ of the variety random variable Y that meets Equation (11), referred to lower bound variety $\Delta(\delta)$ hereafter**.

5.3.2 Determining Lower Bound Variety $\Delta(\delta)$

To discover the boundary relation between δ and Δ that satisfies Equation (11), we need to estimate the distribution of the variety random variable Y in Equation (9). Y denotes the variety of $\mathcal{B}_l(\mathcal{D})$ resulting from the addition or removal of a single user's check-in record from the check-in data set $\mathcal{D}$.

Let Y_i be the random variable on the variety of single $\mathcal{C}(l_i \to l)$ $(i = 1, 2, \ldots, |L|)$ caused by adding or removing a user's check-in record, i.e.,

$$Y_i = |\mathcal{C}(l_i \to l) \text{ on } \mathcal{D}_1 - \mathcal{C}(l_i \to l) \text{ on } \mathcal{D}_2|, \quad (12)$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ are two random neighboring data sets differing in only one record. In terms of Equations (6) and (9), Y can be represented as

$$Y = \sum_{i=1}^{n} 2^{-\alpha i} \cdot Y_i, \tag{13}$$

where n denotes the number of locations visited by a user. To facilitate the estimation of Y, we also define an auxiliary random variable Z as

$$Z = \sum_{i=1}^{|L|} 2^{-\alpha i} \cdot Y_i, \tag{14}$$

where $|L|$ denotes the total number of locations in a geosocial network and $|L| \gg n$, and the order of the first n locations in Z is exactly the same with that in Y, i.e., $Z = Y + \sum_{i=n+1}^{|L|} 2^{-\alpha i} \cdot Y_i$. Thus, $Y \leq Z$ which is independent of the order of the remaining locations in Z. The relation between Y and Z is given in Theorem 2.

Theorem 2. Relation between Y and Z. Given two random variables Y and Z in Equations (13) and (14), if $\Pr[Z \leq \Delta(\delta)] \geq 1-\delta$, then $\Pr[Y \leq \Delta(\delta)] \geq 1-\delta$.

Proof: Due to $Y \leq Z$,

$$\Pr[Y \leq \Delta(\delta)] \geq \Pr[Y \leq Z \leq \Delta(\delta)] = \Pr[Z \leq \Delta(\delta)] \geq 1-\delta.$$

Then, $\Pr[Y \leq \Delta(\delta)] \geq 1-\delta$ is sound. □

Distribution of Z. Based on Theorem 2, we can determine the lower bound variety $\Delta(\delta)$ of Y through estimating the distribution of the larger random variable Z than Y. Due to the one-time transition constraint that only the latest transition of a user to the same location is taken into account when building the transition counts $\mathcal{C}$ in Section 5.1, a user cannot contribute more than one transition to the same location $l \in L$. That is, at most only one random variable in the set of Y_i $(i = 1, 2, \ldots, |L|)$ can take value of 1 and the others are 0. Here we further enlarge Z a little to model its distribution through assuming that there always exists one random variable in the set of Y_i $(i = 1, 2, \ldots, |L|)$ equaling to 1. Moreover, because the differential privacy does not assume any background knowledge on users, each random variable Y_i equals 1 with the probability of $|L|^{-1}$, formalized by

$$\sum_{i=1}^{|L|} Y_i = 1, \text{ and} \tag{15}$$

$$\Pr(Y_i = 1) = |L|^{-1}. \tag{16}$$

If an adversary has background knowledge on users, the differential privacy ensures that the adversary cannot significantly learn more information than her background knowledge using the released noisy data. That is, any privacy breach will not be a result of participating in the data set. Then, the distribution of Z is shown in Theorem 3.

Theorem 3. Distribution of Z. The random variable Z defined in Equation (14) has the following probability mass function

$$\Pr(Z = 2^{-\alpha i}) = |L|^{-1}, (i = 1, 2, \ldots, |L|). \tag{17}$$

Proof: According to Equations (14) and (15), $Z = 2^{-\alpha i}$ if and only if $Y_i = 1$ (note that Y_i only taking value of 1 or 0). Moreover, because of $\Pr(Y_i = 1) = |L|^{-1}$ in Equation (16), Equation (17) is sound. □

Further, the lower bound of variety $\Delta(\delta)$ is shown in Theorem 4.

Theorem 4. The lower bound of variety $\Delta(\delta)$. Given random variable Y defined in Equation (9) or (13), the lower bound of variety $\Delta(\delta)$ that satisfies $\Pr[Y \leq \Delta(\delta)] \geq 1-\delta$ (Equation (11)) is given by

$$\Delta(\delta) = 2^{-\alpha \lfloor |L|\delta + 1 \rfloor}, (0 \leq \delta < 1). \tag{18}$$

Note that it is not meaningful for $\delta = 1$, since it means no privacy protection and there is no need to add noise (i.e., $\Delta(\delta) = 0$); $\lfloor (\cdot) \rfloor$ denotes the largest integer that is less than or equal to $(\cdot)$.

Proof: At first, based on Theorem 2, we can turn to finding the lower bound of variety $\Delta(\delta)$ that satisfies $\Pr[Z \leq \Delta(\delta)] \geq 1-\delta$. Then, from Theorem 3, we have

$$\Pr(Z \leq z) = \begin{cases} 0, & z < 2^{-\alpha|L|}; \\ |L|^{-1}, & 2^{-\alpha|L|} \leq z < 2^{-\alpha(|L|-1)}; \\ \vdots & \vdots \\ (|L|-1)|L|^{-1}, & 2^{-2\alpha} \leq z < 2^{-\alpha}; \\ 1, & z \geq 2^{-\alpha}. \end{cases}$$

$$= \begin{cases} 0, & z < 2^{-\alpha|L|}; \\ \left(|L| + \lfloor \alpha^{-1} \log_2 z \rfloor + 1\right)|L|^{-1}, & 2^{-\alpha|L|} \leq z < 2^{-\alpha}; \\ 1, & z \geq 2^{-\alpha}. \end{cases}$$

Hence:
(1) If $\delta = 0$, to meet $\Pr[Z \leq \Delta(\delta)] \geq 1-\delta$, i.e., $\Pr[Z \leq \Delta(0)] \geq 1$, $\Delta(0) \geq 2^{-\alpha}$ must hold. Thus, we can get the lower bound of variety $\Delta(0) = 2^{-\alpha}$ that identifies with Equation (18) for the case $\delta = 0$.
(2) If $0 < \delta < 1$, to meet $\Pr[Z \leq \Delta(\delta)] \geq 1-\delta$, it is required:

$$\left(|L| + \lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor + 1\right)|L|^{-1} \geq 1-\delta,$$

$$\text{or } \lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor \geq -\left(|L|\delta + 1\right).$$

Since $\lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor$ is an integer (i.e., the largest integer less than or equal to $\alpha^{-1} \log_2 \Delta(\delta)$), it is further required:

$$\lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor \geq \lceil -\left(|L|\delta + 1\right) \rceil = -\lfloor \left(|L|\delta + 1\right) \rfloor.$$

Note that $\alpha^{-1} \log_2 \Delta(\delta) \geq \lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor$ and $\lfloor \alpha^{-1} \log_2 \Delta(\delta) \rfloor$ is monotonically increasing with respect to $\Delta(\delta)$, so the lower bound of variety is given by

$$\alpha^{-1} \log_2 \Delta(\delta) = -\lfloor \left(|L|\delta + 1\right) \rfloor$$

$$\text{or } \Delta(\delta) = 2^{-\alpha \lfloor |L|\delta + 1 \rfloor}.$$

In sum, Theorem 4 is sound. □

According to Theorem 4, the lower bound of variety $\Delta(\delta)$ is monotonically decreasing with respect to the increase of δ as depicted in Fig. 5 for three decay rates α (Note that $|L|$ is the total number of locations in a geosocial network and is usually very large). This relation between δ and Δ reflects the fact that

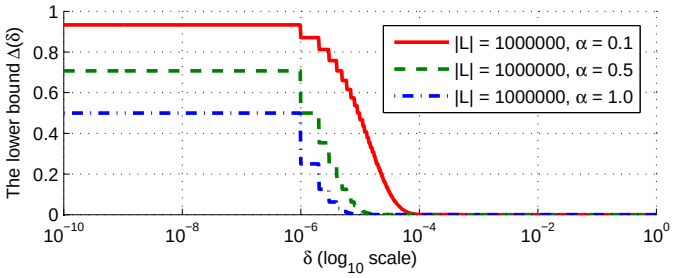


Fig. 5. The lower bound of variety $\Delta(\delta)$ with respect to δ

stricter privacy (i.e., lower δ) needs larger noise (i.e., higher Δ). In sum, the noise X in Equation (7) or (8) can be drawn from the distribution defined by Equations (10) and (18).

5.4 Composition Analysis

As given by Equation (7) in Section 5.2, the proposed differentially private random algorithm $\mathcal{A}$ for protecting $\mathcal{B}_l(\mathcal{D})$ has been decomposed into a set of compositions $\mathcal{A}_l, \forall l \in L$, and the noise X for each $\mathcal{A}_l$ is independently drawn from the distribution defined by Equations (10) and (18) in Section 5.3. Thus, the random algorithm $\mathcal{A}$ must deal with the issue caused by compositions that several outputs of $\mathcal{A}$ may be taken together for joint analysis, and should still provide privacy guarantees even when it is subjected to joint analysis.

Compositions on the privacy budget ε. On the compositions of the privacy budget ε in differential privacy, the study [42] gives one important parallel composition property (Lemma 1).

Lemma 1. Parallel composition property [42]. Let $\mathcal{A}_i$ be a differentially private composition of the random algorithm $\mathcal{A}$. If each $\mathcal{A}_i$ provides ε_i-differential privacy operating on the disjoint subset $\mathcal{D}_i$ of the input data set $\mathcal{D}$, then $\mathcal{A}$ achieves $\max_i \varepsilon_i$-differential privacy.

In terms of Lemma 1, we can prove that our differentially private random algorithm $\mathcal{A}$ given by Equation (7) in Section 5.2 satisfies the parallel composition property, as shown in Theorem 5.

Theorem 5. Parallel compositions of $\mathcal{A}$ on the privacy budget ε. If each differentially private composition $\mathcal{A}_l, \forall l \in L$ provides ε-differential privacy, then $\mathcal{A}$ achieves ε-differential privacy.

Proof: Based on Lemma 1, we only need to prove each composition $\mathcal{A}_l$ operating on the disjoint subset of the check-in data set $\mathcal{D}$. As shown in Fig. 6, $\mathcal{A}_l$ operates on disjoint transition counts $\mathcal{C}(l_i \rightarrow l)(i = 1, 2, \ldots, |L|)$ aggregated from disjoint transition subsets $l_i \rightarrow l$ to the same location l of the check-in data set $\mathcal{D}$. Thus, Theorem 5 holds. □

Therefore, the proposed differentially private random algorithm $\mathcal{A}$ for protecting each $\mathcal{B}_l(\mathcal{D})$ in Equation (7) of Section 5.2 accomplishes the privacy budget ε with the same level as its compositions by operating on the disjoint subset of the input data set $\mathcal{D}$, which is important to overcome the joint analysis. Otherwise, the privacy guarantee of $\mathcal{A}$ would degrade if the compositions of $\mathcal{A}$ operated on the joint subsets of the input data set $\mathcal{D}$, i.e., the total privacy budget ε is the sum of ε_i in each composition, $\varepsilon = \sum_i \varepsilon_i$ [42].

Compositions on the breach probability δ. On the compositions of the breach probability δ in probabilistic differential privacy, we have the following result (Theorem 6).

Theorem 6. Independent compositions of $\mathcal{A}$ on the breach probability δ. If each composition $\mathcal{A}_l, \forall l \in L$ provides (ε, δ')-probabilistic differential privacy, then $\mathcal{A}$ achieves (ε, δ)-probabilistic differential privacy,

$$\delta = 1 - (1 - \delta')^{n_{\max}}, \tag{19}$$

where $n_{\max}$ is the maximum constraint of transitions when aggregating transition counts from the sequences in Section 5.1.

Proof: The compositions on ε is given by Theorem 5, so we focus on the compositions on δ. Since only the $n_{\max}$ latest transitions are considered when building the transition counts from a user's check-in sequence, any sequence can lead to the breach of ε-differential privacy over at most $n_{\max}$ compositions $\mathcal{A}_l$. Further, because each of the $n_{\max}$ compositions independently draws the noise X from the distribution of Equation (10) and has the breach at most δ' probability, $\mathcal{A}$ achieves ε-differential privacy with at least $(1 - \delta')^{n_{\max}}$ probability, i.e., the breach with at most $\delta = 1 - (1 - \delta')^{n_{\max}}$ probability. □

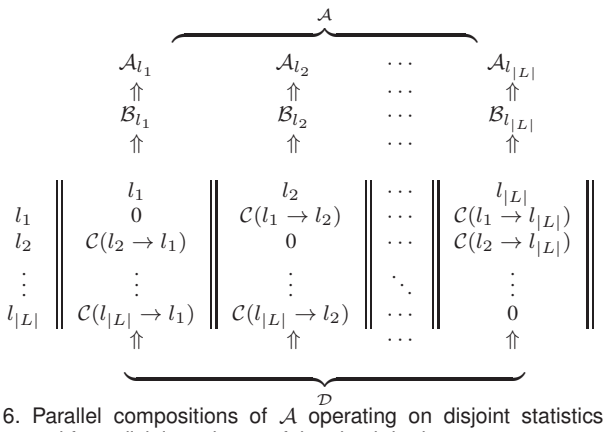


Fig. 6. Parallel compositions of $\mathcal{A}$ operating on disjoint statistics $\mathcal{C}$ aggregated from disjoint subsets of the check-in data set $\mathcal{D}$

Based on Theorem 6, the total breach probability δ becomes higher with the increase of $n_{\max}$. Thus, it is very important to constrain $n_{\max}$ in Section 5.1.

6 Experimental Settings

In this section, we describe our experiment settings for evaluating the performance of PLORE.

6.1 Three Real Data Sets

We use three publicly available large-scale real check-in data sets that were crawled from (i) Foursquare over the period from Jan. 2011 to Jul. 2011 [15], (ii) Gowalla over the period from Feb. 2009 to Oct. 2010, and (iii) Brightkite over the period from Apr. 2008 to Oct. 2010 [1], in which the locations are distributed all over the world. The statistics of the data sets are shown in TABLE 2. In the pre-processing, we split each data set into the training set and the testing set in terms of the check-in time rather than using a random partition method, because in practice we can only utilize the past check-in data to predict the future check-in events. A half of check-in data with earlier timestamps are used as the training set and the other half of check-in data are used as the testing set. In the experiments, the training set is used to learn the recommendation models of the evaluated techniques described in Section 6.2 to predict the testing data.

6.2 Evaluated Techniques

The evaluated state-of-the-art techniques are classified into two categories, as listed below.

(1) Location recommendation category without privacy protection:

- Non-Personalized Method (NPM): It recommends the same locations with the highest popularity to all users [10].
- Coarse-Grained Method (CGM): It applies the classical higher-order Markov chain by utilizing the counts of n-grams [8].
- First-order Markov Chain (FMC): It derives the transition probability of a user to a new location using only her latest

TABLE 2
Statistics of the three real data sets

	Foursquare	Gowalla	Brightkite
Number of users	11,326	196,591	58,228
Number of locations ($\|L\|$)	182,968	1,280,969	772,965
Number of check-ins	1,385,223	6,442,890	4,491,143
User-location matrix density	2.3×10^{-4}	2.4×10^{-5}	1.9×10^{-5}

TABLE 3
Parameter settings

Parameter	Range	Default
No. of visited locations by a user (n)	relied on data	all
No. of recommended locations for a user (k)	2 to 50	all
Decay rate (α)	0.0625 to 2	0.5
Maximum sequence length ($n_{\max}$)	100	100
Privacy budget (ε)	10^{-5} to 10^{0}	0.1
Breach probability (δ)	10^{-5} to 10^{-1}	0.01

visited location based on Equation (3) in Section 4 and is widely used in existing works [2], [3], [4], [5], [6].

- Additive Markov Chain (AMC): Our AMC [7] deduces the transition probability of a user to a new location using all her visited locations based on Equation (5) in Section 4.

(2) Differential privacy category for highly sensitive data:

- PLORE: Our PLORE provides (ε, δ)-probabilistic differential privacy for AMC based on Equations (8), (10) and (18) in Section 5.
- GS: It achieves ε-differential privacy for AMC through preprocessing counts by Grouping and Smoothing them via averaging [12].
- DPSS: In the work [13], it is also called DPSense-S that completes ε-differential privacy for AMC through sensitivity control by the normalization of counts.

6.3 Performance Metrics

Recommendation accuracy. In general, recommendation techniques compute a score for each candidate item (i.e., a location in this paper) regarding a target user and return locations with the top-k highest scores as a recommendation result to the user. To evaluate the quality of location recommendations, it is important to find out how many locations actually visited by the user in the testing data set are discovered by the recommendation techniques. For this purpose, we employ two standard metrics: Precision and Recall [7]. Another two metrics are Normalized Discounted Cumulative Gain (NDCG) and Mean Average Precision (MAP) that are used to measure the ranking quality of the top-k recommendations. The relevance values of recommendations are their the check-in counts made by the target user.

Recommendation efficiency. We show the running time of PLORE with respect to various numbers of check-in locations of users. All algorithms were implemented in Matlab and run on a machine with 3.4GHz Intel Core i7 Processor and 16GB RAM.

6.4 Parameter Settings

TABLE 3 lists the parameter settings in our experiments.

(1) **The number of visited locations by a user in the training set (n).** n is not tunable because it is specified by users. Unless otherwise specified, the performance of evaluated recommendation techniques is averaged over all users with various numbers of check-in locations relied on the training set.

(2) **The number of recommended locations for a user (k).** The top-k is set to the large range from 2 to 50. By default, the average performance of all users is further averaged over all these k values.

(3) **The decay rate (α).** We set α to the range with small values from 0.0625 to 2 having the default value of 0.5, since the decay rate is exponential and fast.

(4) **The maximum constraint of transitions for building transition count matrix ($n_{\max}$).** According to Fig. 3(b) in Section 5.1, $n_{\max}$ is set to 100.

(5) **The privacy budget (ε).** We consider the range of ε from 10^{-5} to 10^{0} with the default value of 0.1 that is referred to a typical small value in the literature on differential privacy. A lower ε value signifies a stricter privacy protection.

(6) **The breach probability (δ).** Given the total breach probability δ, we can set each composition $\mathcal{A}_l, \forall l \in L$ with $\delta' = 1 - \sqrt[n_{\max}]{1-\delta}$ based on Equation (19). We examine the range of δ from 10^{-5} to 10^{-1} with the default value of 0.01 that is referred to a high confidence level in statistics.

TABLE 4
NDCG in top-10 recommendations

	NPM	CGM	FMC	AMC	PLORE	GS	DPSS
Foursquare	0.0216	0.0306	0.0832	0.1943	0.1925	0.0942	0.0751
Gowalla	0.0123	0.0187	0.0468	0.1409	0.1402	0.0689	0.0509
Brightkite	0.0068	0.0114	0.0130	0.0954	0.0938	0.0487	0.0401

TABLE 5
MAP in top-10 recommendations

	NPM	CGM	FMC	AMC	PLORE	GS	DPSS
Foursquare	0.0189	0.0295	0.0769	0.1488	0.1464	0.0833	0.0702
Gowalla	0.0118	0.0146	0.0399	0.1098	0.1068	0.0612	0.0476
Brightkite	0.0064	0.0101	0.0119	0.0806	0.0793	0.0455	0.0309

7 Experimental Results

This section analyzes the experimental results: recommendation accuracy in Section 7.1, the trade-off between recommendation accuracy and location privacy in Section 7.2, and recommendation efficiency in Section 7.3.

7.1 Recommendation Accuracy

We compare the recommendation accuracy of all evaluated techniques on the three data sets in TABLES 4 and 5, and Fig. 7 and 8. We have the following general and important findings.

The category without privacy protection. (1) The non-personalized method NPM [10] suffers from very low recommendation accuracy, since it returns the same locations to all users and thus cannot satisfy the personal preference of different users. The coarse-grained method CGM [8] also generates inaccurate recommendations for users, because it severely depends on the counts of n-grams ($n > 2$) and almost all these counts are zero due to the highly sparse check-in data, as depicted in TABLE 2. These results show that the two methods are not applicable to the personalized and fine-grained location recommendations, even without considering the location privacy protection. (2) FMC [2], [3], [4], [5], [6] computes the transition probability of a user to a new location by utilizing only her latest visited location

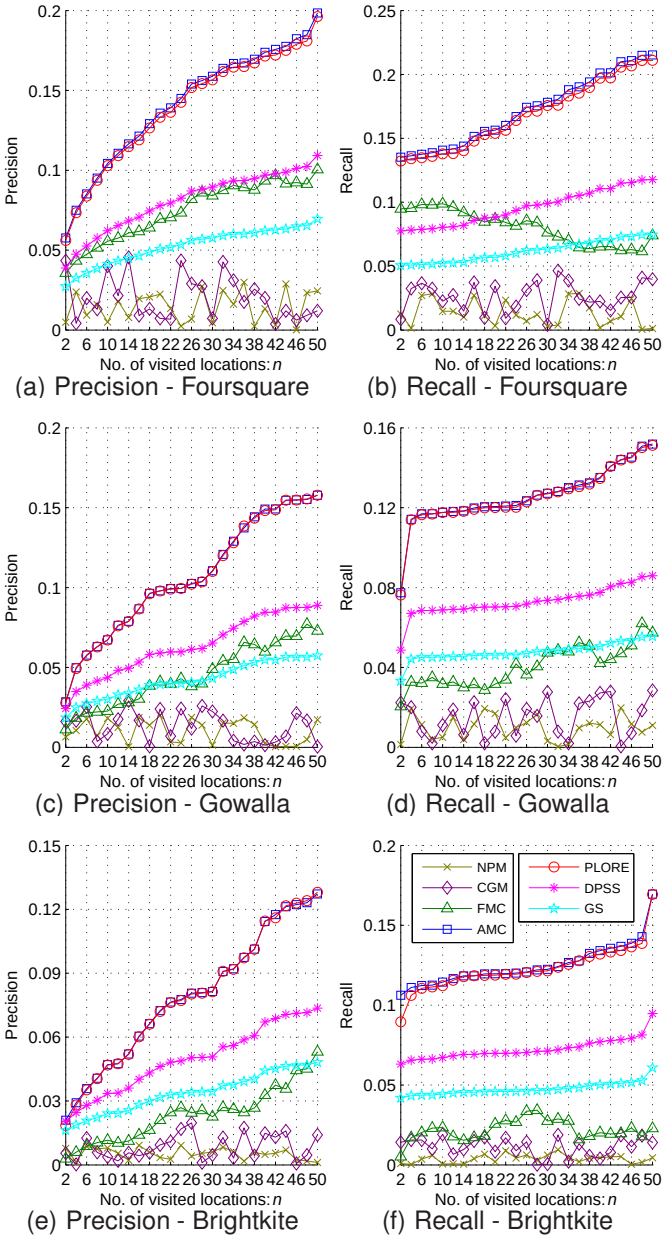


Fig. 7. Recommendation accuracy on visited location number n ($\alpha = 0.5$, $\varepsilon = 0.1$, and $\delta = 0.01$)

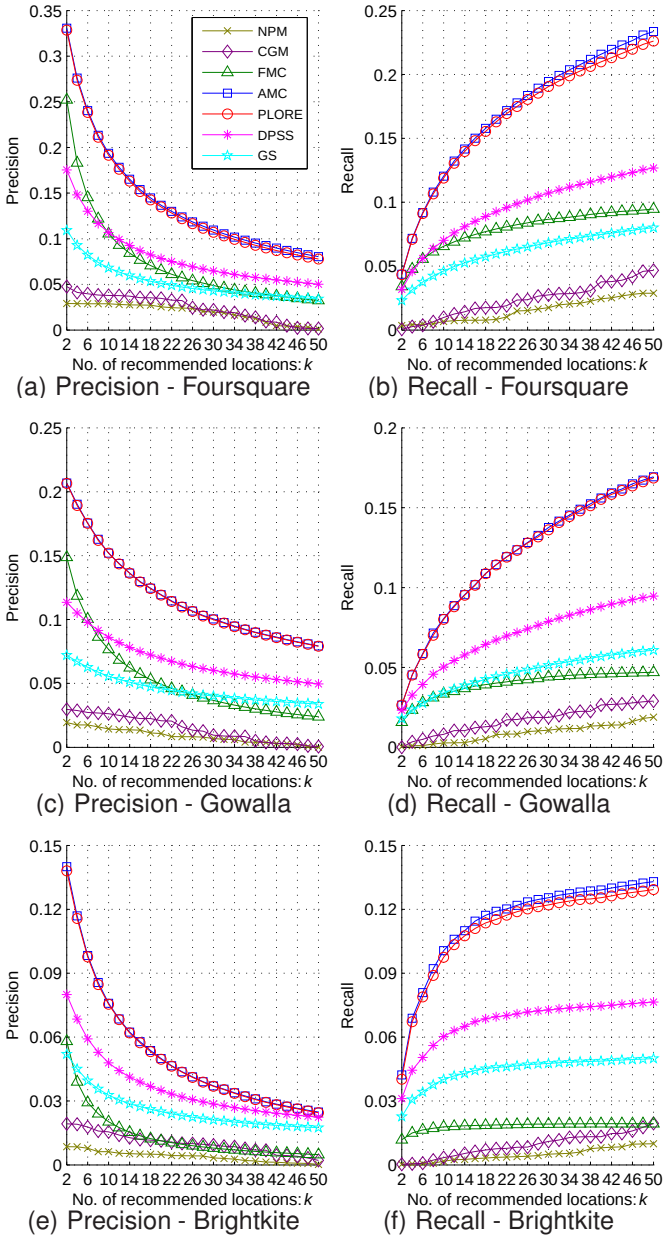


Fig. 8. Recommendation accuracy on recommended location number k ($\alpha = 0.5$, $\varepsilon = 0.1$, and $\delta = 0.01$)

in the check-in sequence. As a result, it often cannot take full advantage of sequential patterns in location recommendations, since it ignores the impact of the earlier visited locations in the sequence on the new likely visiting locations. Thus, FMC returns inaccurate locations in terms of NDCG, MAP, precision and recall. (3) To overcome the limitation of FMC, AMC derives the transition probability of a user to new locations based on all her visited locations and leans the weight towards recently visited locations. Accordingly, AMC significantly increases the NDCG, MAP, precision and recall of FMC on the three data sets. These results verify the superiority of exploiting the impact of the whole location sequence for location recommendations proposed in our recent study [7] over only considering the latest visited location in the current works [2], [3], [4], [5], [6].

The category with privacy protection. In this category, all evaluated privacy-preserving techniques are integrated with AMC to observe how each of them deteriorates the recommendation accuracy of AMC. (1) By greatly reducing the noise added into

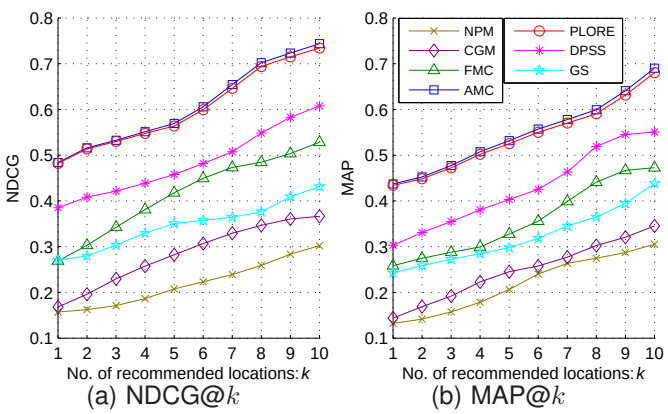


Fig. 9. Accuracy on another Foursquare data set with higher data density ($\alpha = 0.5$, $\varepsilon = 0.1$, and $\delta = 0.01$)

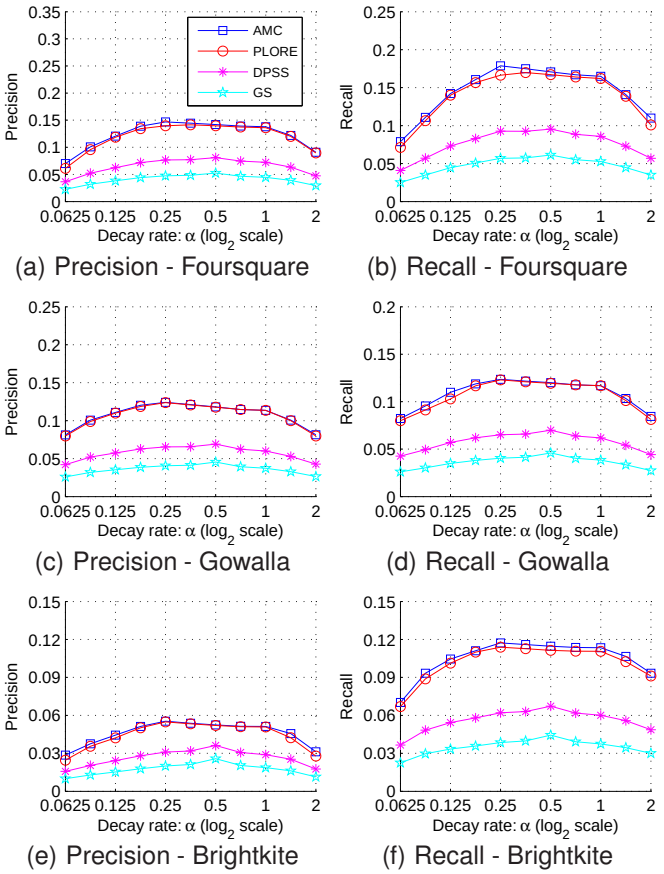


Fig. 10. Accuracy on decay rate α ($\varepsilon = 0.1$ and $\delta = 0.01$)

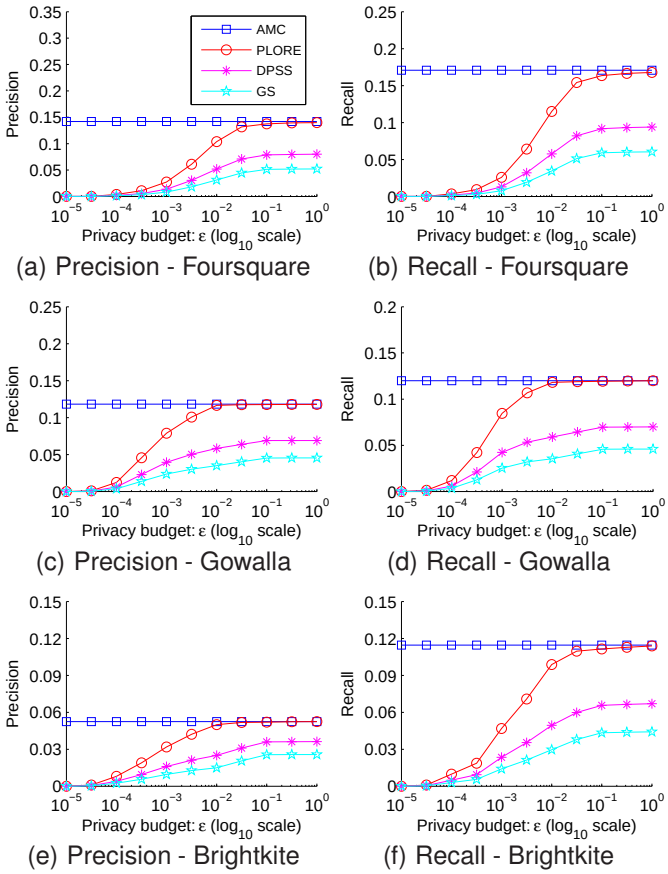


Fig. 11. Accuracy on privacy budget ε ($\alpha = 0.5$ and $\delta = 0.01$)

transition counts, our PLORE not only achieves the same level of recommendation accuracy as AMC, but also protects the location privacy of users rigorously, i.e., ensuring strict differential privacy ($\varepsilon = 0.1$, a typical low value) with high probability, at least $1 - \delta = 0.99$. (2) Although DPSS and GS guarantee stricter differential privacy by injecting much more noise, their accuracy is severely degraded into a low level, as depicted in TABLES 4 and 5, and Fig. 7 and 8. As a result, their utility is very limited in location recommendations. We will discuss the trade-off between accuracy and privacy in details in Section 7.2.

Effect of visited and recommended location numbers on accuracy. (1) Using more check-in data of users (i.e., n with a larger value), most methods can learn users' preferences more accurately and then the accuracy inclines, as depicted in Fig. 7. However, the accuracy of FMC is a little fluctuant since it only uses the latest visited location to predict the probability of a user visiting new locations, i.e., it is independent of the value of n. (2) By returning more locations to users (k with a larger value), they always can discover more preferred locations but the extra recommended locations are less possible to be liked by them, so the recall becomes higher and the precision gets lower, as shown in Fig. 8.

Effect of data density on accuracy. It is important to note that the accuracy of location recommendation techniques for geosocial networks is usually not high, because the density of a user-location check-in matrix is pretty low. For example, the three data sets used in our experiments have very low densities, i.e., 2.3×10^{-4}, 2.4×10^{-5} and 1.9×10^{-5} in the Foursquare, Gowalla and Brightkite data sets, respectively (TABLE 2). Further, we have conducted experiments on another Foursquare data set with the higher density of 7.2×10^{-3}; this data set contains 824 users, 38,336 locations and 227,428 check-ins [43]. Fig. 9 shows the accuracy regarding different top-k values, in which the NDCG@10 and MAP@10 ($k = 10$) of all evaluated recommendation techniques are much higher than that in TABLES 4 and 5, since the new check-in data set is at least one order-of-magnitude denser. More importantly, our PLORE not only significantly improves the recommendation accuracy compared to the state-of-the-art competitors, but also preserves the location privacy of users rigorously.

Effect of decay rate on accuracy. Fig. 10 depicts the effect of the decay rate α on the precision and recall. When α changes in a large range from 0.25 to 1, AMC and PLORE perform stably, which is an important feature for us to choose a default value of α instead of finding the optimal value that usually costs much more effort and suffers from over-fitting. In contrast, when α is out of the range between 0.25 and 1, the precision and recall of AMC and PLORE become lower. The reason is that a large value of α tends to overestimate the effect of the recently visited locations on the newly possible visiting locations (i.e., underestimating the effect of the anciently visited locations), whereas a small value of α is prone to weighing all visited locations equally. Importantly, the accuracy of PLORE always approaches that of AMC, although PLORE provides the strict location privacy protection for users. Again we have observed that DPSS and GS achieve stricter differential privacy at the expense of much lower accuracy.

7.2 Trade-off between Accuracy and Privacy

This section discusses the relation between *recommendation accuracy* and *location privacy*. Due to similar results and space limitations, we omit the result on NDCG and MAP. The results

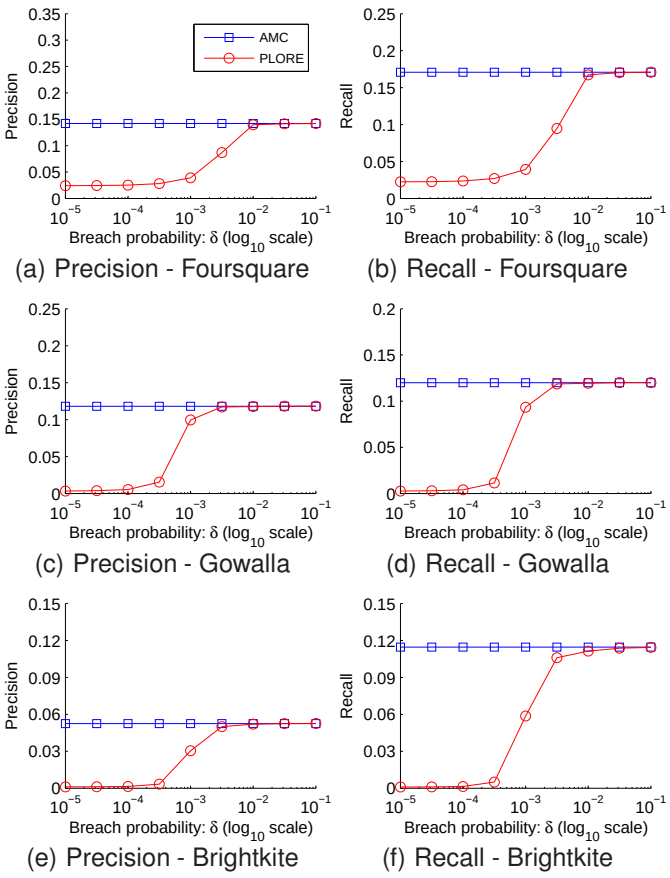


Fig. 12. Accuracy on breach probability δ ($\alpha = 0.5$ and $\varepsilon = 0.1$)

of AMC are also plotted to observe the accuracy loss of PLORE, DPSS and GS due to privacy protection.

Privacy budget vs. accuracy. Fig. 11 depicts the recommendation accuracy with respect to different privacy budgets (ε). In general, a lower privacy budget results in less accuracy since greater noise is required to inject into the aggregate statistics used by recommendation models, as observed in Fig. 11 on the three data sets. For example, when the privacy budget ε gets smaller, the accuracy of DPSS and GS quickly deteriorates. Moreover, DPSS and GS report much lower accuracy than PLORE because they inject a relatively large amount of noise for the small transition statistics from sparse check-in data, which is consistent with the results in TABLES 4 and 5, and Fig. 7, 8, and 9. Promisingly, the decrement on the accuracy of PLORE is insensitive when lowering privacy budgets from 1 to a very small value 0.01. Our explanation is that the lower bound of variety in Equation (18) dominates the effect of the privacy budget ε on the distribution of noise in Equation (10). Subsequently, without compromising the accuracy, PLORE can achieve considerably strict differential privacy (i.e., the very low ε) with some probability (i.e., $1 - \delta$). Under this case, the privacy strictness is mainly measured by the breach probability δ.

Breach probability vs. accuracy. Fig. 12 depicts the recommendation accuracy of PLORE with respect to different breach probabilities (δ). As expected, a lower breach probability of ε-differential privacy generally leads to lower accuracy, because it is required to add larger noise into the aggregate statistics for recommendation models based on Equations (10) and (18). Importantly, PLORE is able to accomplish a low breach probability (i.e., $\delta = 0.01$) of ε-differential privacy with negligible accuracy loss on the three data sets. These promising results benefit from the proposed probabilistic differential privacy mechanism based on the lower bound of varieties caused by removing or adding a single user's check-in sequence for any required breach level δ; the lower bound variety minimizes the required noise for a given breach level δ.

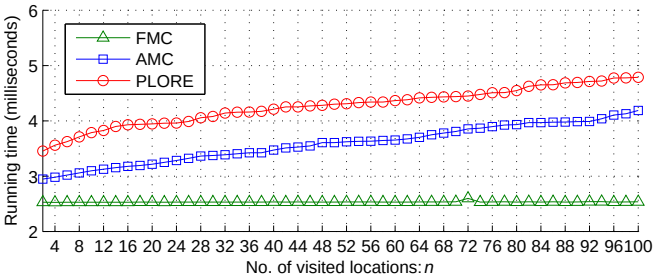


Fig. 13. Recommendation efficiency

7.3 Recommendation Efficiency

Due to similar results and space limitations, Fig. 13 depicts the running time of the evaluated recommendation techniques respecting the number of visited locations (n) by a user on the Foursquare data set only. (1) FMC maintains constant running time because it only utilizes the latest visited location of a user to derive transition probabilities. That is, its running time is independent of the length of check-in sequence of the user. Unfortunately, this high efficiency is at the cost of low accuracy as shown in Section 7.1. (2) AMC takes linearly and slowly increasing running time as the given check-in sequence of users gets longer, which is much faster than the classical nth-order Markov chain that has the exponential complexity. (3) PLORE requires some extra time to inject noise for privacy protection in comparison to AMC. Fortunately, the increment of time also remains constant independent of the length of check-in sequences of users. Hence, PLORE is competitive to AMC and both have the same computational complexity $O(n)$. Note that both DPSS and GS cost more time than PLORE to achieve ε-differential privacy, so it is not fair to compare them.

8 Conclusion and Future Work

Location recommenders access the raw check-in data of users to mine their preferences and raise serious location privacy breaches. Most existing studies apply differential privacy to provide formal location privacy guarantees against adversaries with background knowledge at the cost of recommendation accuracy, i.e., recommending non-personalized or coarse-grained locations for users. To address the two key challenges on *recommendation accuracy* and *location privacy* caused by the high sensitivity and small count problems in the context of personalized and fine-grained location recommendations, this paper proposes a new probabilistic differential privacy-preserving location recommendation framework called PLORE for geosocial networks. PLORE improves recommendation accuracy using nth-order additive Markov chain and strives for a good trade-off between recommendation accuracy and location privacy through the developed probabilistic differential privacy mechanism. The proposed privacy mechanism exploits the *lower bound of variety* instead of the sensitivity (i.e., the *maximum variety*) to reduce the noise injected into aggregate statistics. Finally, extensive experimental results on three real-world data sets show that PLORE achieves *high recommendation accuracy* and *strict location privacy*.

In the future, we plan to study two directions: (1) how to recommend a trip of locations with differential privacy protection, and (2) how to devise differential privacy mechanisms for other recommendation methods, e.g., geographical techniques.

ACKNOWLEDGMENTS

This research was partially supported by the National Natural Science Foundation of China under Grant 61772445 and the City University of Hong Kong under Grant 7004684.

REFERENCES

[1] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *ACM KDD*, 2011.
[2] Z. Chen, H. T. Shen, and X. Zhou, "Discovering popular routes from trajectories," in *IEEE ICDE*, 2011.
[3] A.-J. Cheng, Y.-Y. Chen, Y.-T. Huang, W. H. Hsu, and H.-Y. M. Liao, "Personalized travel recommendation by mining people attributes from community-contributed photos," in *ACM MM*, 2011.
[4] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *IJCAI*, 2013.
[5] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geotags in photo sharing sites," in *ACM CIKM*, 2010.
[6] Y.-T. Zheng, Z.-J. Zha, and T.-S. Chua, "Mining travel patterns from geotagged photos," *ACM TIST*, vol. 3, no. 3, pp. 56:1–56:18, 2012.
[7] J.-D. Zhang, C.-Y. Chow, and Y. Li, "LORE: Exploiting sequential influence for location recommendations," in *ACM SIGSPATIAL*, 2014.
[8] J.-D. Zhang, G. Ghinita, and C.-Y. Chow, "Differentially private location recommendations in geosocial networks," in *IEEE MDM*, 2014.
[9] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *ACM CCS*, 2013.
[10] D. Riboni and C. Bettini, "Differentially-private release of check-in data for venue recommendation," in *IEEE PerCom*, 2014.
[11] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *TCC*, 2006.
[12] G. Kellaris and S. Papadopoulos, "Practical differential privacy via grouping and smoothing," *PVLDB*, vol. 6, no. 5, pp. 301–312, 2013.
[13] W.-Y. Day and N. Li, "Differentially private publishing of high-dimensional data using sensitivity control," in *ASIA CCS*, 2015.
[14] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: A survey," *GeoInformatica*, vol. 19, no. 3, pp. 525–565, 2015.
[15] H. Gao, J. Tang, and H. Liu, "gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks," in *ACM CIKM*, 2012.
[16] J.-D. Zhang and C.-Y. Chow, "CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations," *Information Sciences*, vol. 293, pp. 163–181, 2015.
[17] ——, "GeoSoCa: Exploiting geographical, social and categorical correlations for point-of-interest recommendations," in *ACM SIGIR*, 2015.
[18] ——, "TICRec: A probabilistic framework to utilize temporal influence correlations for time-aware location recommendations," *IEEE TSC*, vol. 9, no. 4, pp. 633–646, 2016.
[19] ——, "CRATS: An LDA-based model for jointly mining latent communities, regions, activities, topics, and sentiments from geosocial network data," *IEEE TKDE*, vol. 28, no. 11, pp. 2895–2909, 2016.
[20] J.-D. Zhang, C.-Y. Chow, and Y. Li, "iGeoRec: A personalized and efficient geographical location recommendation framework," *IEEE TSC*, vol. 8, no. 5, pp. 701–714, 2015.
[21] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," *IEEE TKDE*, vol. 27, no. 5, pp. 1167–1179, 2015.
[22] J.-D. Zhang, C.-Y. Chow, and Y. Zheng, "ORec: An opinion-based point-of-interest recommendation framework," in *ACM CIKM*, 2015.
[23] J.-D. Zhang and C.-Y. Chow, "Spatiotemporal sequential influence modeling for location recommendations: A gravity-based approach," *ACM TIST*, vol. 7, no. 1, pp. 11:1–11:25, 2015.
[24] S. Mascetti, D. Freni, C. Bettini, X. S. Wang, and S. Jajodia, "Privacy in geo-social networks: Proximity notification with untrusted service providers and curious buddies," *VLDBJ*, vol. 20, no. 4, pp. 541–566, 2011.
[25] J.-D. Zhang and C.-Y. Chow, "REAL: A reciprocal protocol for location privacy in wireless sensor networks," *IEEE TDSC*, vol. 12, no. 4, pp. 458–471, 2015.
[26] M. Ghasemzadeh, B. C. Fung, R. Chen, and A. Awasthi, "Anonymizing trajectory data for passenger flow analysis," *Transportation Research Part C: Emerging Technologies*, vol. 39, pp. 63–79, 2014.
[27] R. Chen, G. Acs, and C. Castelluccia, "Differentially private sequential data publication via variable-length n-grams," in *ACM CCS*, 2012.
[28] G. Cormode, C. Procopiuc, D. Srivastava, E. Shen, and T. Yu, "Differentially private spatial decompositions," in *IEEE ICDE*, 2012.
[29] R. Chen, B. C. Fung, B. C. Desai, and N. M. Sossou, "Differentially private transit data publication: A case study on the montreal transportation system," in *ACM KDD*, 2012.
[30] Y. Hong, J. Vaidya, H. Lu, P. Karras, and S. Goel, "Collaborative search log sanitization: Toward differential privacy and boosted utility," *IEEE TDSC*, vol. 12, no. 5, pp. 504–518, 2015.
[31] A. Inan, M. Kantarcioglu, G. Ghinita, and E. Bertino, "A hybrid approach to private record matching," *IEEE TDSC*, vol. 9, no. 5, pp. 684–698, 2012.
[32] N. Li, W. Qardaji, D. Su, and J. Cao, "PrivBasis: Frequent itemset mining with differential privacy," *PVLDB*, vol. 5, no. 11, pp. 1340–1351, 2012.
[33] W. Qardaji, W. Yang, and N. Li, "Understanding hierarchical methods for differentially private histograms," *PVLDB*, vol. 6, no. 14, pp. 1954–1965, 2013.
[34] M. A. Pathak and B. Raj, "Large margin Gaussian mixture models with differential privacy," *IEEE TDSC*, vol. 9, no. 4, pp. 463–469, 2012.
[35] Z. Jorgensen and T. Yu, "A privacy-preserving framework for personalized, social recommendations," in *EDBT*, 2014.
[36] C. Li, B. Palanisamy, and J. Joshi, "Differentially private trajectory analysis for points-of-interest recommendation," in *IEEE BigData*, 2017.
[37] V. Bindschaedler and R. Shokri, "Synthesizing plausible privacy-preserving location traces," in *IEEE S&P*, 2016.
[38] E. ElSalamouny and S. Gambs, "Differential privacy models for location-based services," *Transactions on Data Privacy*, vol. 9, no. 1, pp. 15–48, 2016.
[39] S.-S. Ho, "Preserving privacy for moving objects data mining," in *IEEE ISI*, 2012.
[40] S.-S. Ho and S. Ruan, "Preserving privacy for interesting location pattern mining from trajectory data," *Transactions on Data Privacy*, vol. 6, no. 1, pp. 87–106, 2013.
[41] M. Götz, A. Machanavajjhala, G. Wang, X. Xiao, and J. Gehrke, "Publishing search logs — a comparative study of privacy guarantees," *IEEE TKDE*, vol. 24, no. 3, pp. 520–532, 2012.
[42] F. D. McSherry, "Privacy integrated queries: An extensible platform for privacy-preserving data analysis," in *ACM SIGMOD*, 2009.
[43] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE TSMC*, vol. 45, no. 1, pp. 129–142, 2015.

Jia-Dong Zhang received the M.Sc. degree from Yunnan University, China, in 2009, and the Ph.D. degree from City University of Hong Kong in 2015. He is currently a research fellow in Department of Computer Science, City University of Hong Kong. His research work has been published in premier conferences (e.g., *ACM SIGIR*, *CIKM* and *SIGSPATIAL*), transactions (e.g., *ACM TIST*, *IEEE TKDE*, *TDSC TSC* and *TITS*), and journals (e.g., *Pattern Recognition* and *Information Sciences*). His research interests include data mining, location-based services and location privacy.

Chi-Yin Chow received the M.S. and Ph.D. degrees from the University of Minnesota-Twin Cities, USA in 2008 and 2010, respectively. He is currently an assistant professor in Department of Computer Science, City University of Hong Kong. His research interests include big data analytics, data management, GIS, mobile computing, location-based services, and data privacy. He is the co-founder and co-chair of the ACM SIGSPATIAL MobiGIS 2012 to 2016, and the editor of the ACM SIGSPATIAL Newsletter. He received the VLDB "10-year award" in 2016, and the best paper awards in ICA3PP 2015 and IEEE MDM 2009.