PriSTE: Protecting Spatiotemporal Event Privacy in Continuous Location-Based Services

Yang Cao
Kyoto University, Japan
yang@i.kyoto-u.ac.jp

Yonghui Xiao
Google Inc., USA
yohu@google.com

Li Xiong
Emory University, USA
lxiong@emory.edu

Liquan Bai
Emory University, USA
bailiquan@gmail.com

Masatoshi Yoshikawa
Kyoto University, Japan
yoshikawa@i.kyoto-u.ac.jp

ABSTRACT

Location privacy-preserving mechanisms (LPPMs) have been extensively studied for protecting a user's location in location-based services. However, when user's perturbed locations are released continuously, existing LPPMs may not protect users' sensitive *spatiotemporal event*, such as "visited hospital in the last week" or "regularly commuting between location 1 and location 2 every morning and afternoon" (it is easy to infer that locations 1 and 2 may be home and office). In this demonstration, we demonstrate PriSTE for protecting spatiotemporal event privacy in continuous location release. First, to raise users' awareness of such a new privacy goal, we design an interactive tool to demonstrate how accurate an adversary could infer a secret spatiotemporal event from a sequence of locations or even LPPM-protected locations. The attendees can find that some spatiotemporal events are quite risky and even these state-of-the-art LPPMs do not always protect spatiotemporal event privacy. Second, we demonstrate how a user can use PriSTE to automatically or manually convert an LPPM for location privacy into one protecting spatiotemporal event privacy in continuous location-based services. Finally, we visualize the trade-off between privacy and utility so that users can choose appropriate privacy parameters in different application scenarios.

PVLDB Reference Format:
Yang Cao, Yonghui Xiao, Li Xiong, Liquan Bai, Masatoshi Yoshikawa. PriSTE: Protecting Spatiotemporal Event Privacy in Continuous Location-Based Services. *PVLDB*, 12(12): 1866 - 1869, 2019.
DOI: https://doi.org/10.14778/3352063.3352086

1. INTRODUCTION

In our modern life, people often use location-based services (LBS) such as Yelp or Uber for snapshot or continuous queries, for example, "where is the nearest restaurant" or "continuously report the taxis within one mile of my location". Mobile users have to share their current location, or a sequence of locations with the service providers, which raises privacy concerns since users' digital trace can be used to infer sensitive information, such as home and workplace, religious places and sexual inclinations [6].

In order to protect location privacy, many studies (see surveys [5]) have explored different aspects of location privacy: privacy goals, adversarial models, location privacy metrics, and location privacy preserving mechanisms (LPPMs). *Privacy goals* indicate what should be protected or what are the secrets (e.g., a single location or a trajectory); *adversarial models* make assumptions about the adversaries; *location privacy metrics* formally define the quantitative measurement of the protection w.r.t. the privacy goal; LPPMs is designed to achieve a specified privacy metric. For instance, Geo-Indistinguishability [1] is a location privacty metrics, which is receiving increasing attention since the protection level does not depend on adversaries' prior knowledge; the privacy goal of Geo-Indistinguishability is to protect a single location; Laplace Planar Mechanism is an LPPM satisfying Geo-Indistinguishability. In this study, we focus on state-of-the-art probabilistic LPPMs, which takes an actual location and a privacy parameter as inputs and outputs a randomly perturbed location. The LPPM privacy parameter controls the location privacy level (take Laplace Planar Mechanism for example, a smaller privacy parameter indicates stronger privacy protection).

We argue that existing LPPMs [1][7][8] may not adequately protect users' sensitive information in their spatiotemporal activities because the *privacy goal* in location privacy is not well-studied. The existing LPPMs focused on the protection of either a single location or a trajectory, which does not completely reflect the secrets that should be protected in users' spatiotemporal activities. To explain this, we need to define "spatiotemporal activities". We define a user's a single location at time t as a predicate $u^t = s_i$ where u^t is the user's position at time t and $s_i \in \mathbb{S}, i \in [1, m]$ is one location on the map $\mathbb{S}$ of m locations. The value of such predicate can be either *true* or *false*, which could be a secret of the user. Then, we can represent users' spatiotemporal activities as Boolean expressions of combining different predicates over spatial and/or temporal dimensions, which is called *spatiotemporal event* in this paper (a predicate alone also can be a spatiotemporal event). As shown in Fig.1,

Proceedings of the VLDB Endowment, Vol. 12, No. 12
ISSN 2150-8097.
DOI: https://doi.org/10.14778/3352063.3352086

we illustrate six representative examples of the Boolean expression between location and time dimensions. It is easy to see that the events representing a sensitive location/area and a trajectory are only two cases (i.e., (b) and (c)) among the six enumerated examples. Therefore, even if an LPPM protects each location or a trajectory, it may not protect a secret spatiotemporal event.

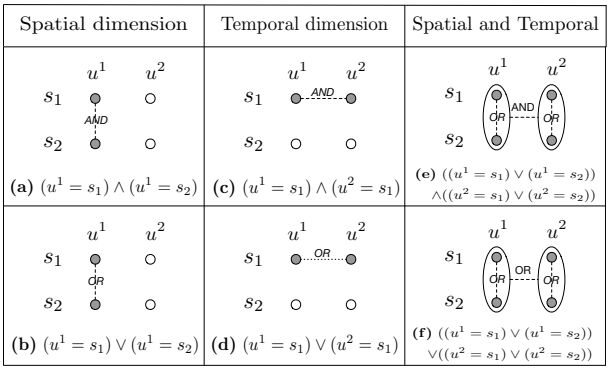

Figure 1: Examples of spatiotemporal events. s_1 and s_2 are two locations. u^1 and u^2 are two variables about a user's possible locations at time 1 and time 2, respectively. Event (a) is always false since a user cannot be at two different locations at the same time. Event (b) means that the secret is a sensitive *area* including locations $\{s_1, s_2\}$. Event (c) represents a sensitive *trajectory* $s_1 \rightarrow s_1$. Event (d) denotes that the secret is the visit to s_1 at time point 1 *or* 2. Event (e) depicts the secret as a type of trajectory **pattern**, i.e., the user may stay at two sensitive areas successively. Event (f) indicates the secret as user's **presence** in sensitive area $\{s_1, s_2\}$ at either time point 1 or 2.

In our recent work [2], we formally defined the such a new privacy goal, i.e., spatiotemporal event, as a Boolean expression between spatial and temporal predicates. We also proposed a new privacy metrics, ϵ-Spatiotemporal Event Privacy, by extending the notion of differential privacy [4]. Although the spatiotemporal event is a generalization of a single location or a trajectory in terms of privacy goal, interestingly, we showed that location privacy metrics and spatiotemporal event privacy metrics could be orthogonal privacy notions. That is to say, an LPPM may not provide spatiotemporal event privacy, while a mechanism for spatiotemporal event privacy has no guarantee of location privacy. We developed a quantification-based method to adapt a given LPPM to protect spatiotemporal event privacy so that a user could enjoy the best of two worlds: the underlying LPPM provides general protection against unknown risks, while spatiotemporal event privacy guarantees flexible and customizable protection which may not be provided by the existing LPPMs.

However, there are three challenges when users use the proposed method in [2] for protecting secret spatiotemporal event in practice. First, it is difficult for users to understand the privacy implications (or the risks) of spatiotemporal event privacy. Second, a user may have different location privacy demands on different locations, so she may want to change the LPPM privacy parameter on the fly depending on the place she is visiting, but such an interface is lacking in [2]. Third, it is hard to know how to set the privacy parameter for striking the right balance between privacy and utility in different scenarios.

To address these challenges, we implement our algorithms in [2] with additional user-friendly modules into PriSTE (Private SpatioTemporal Event).

This demonstration makes the following contributions:

First, to raise users' awareness of such a new privacy goal, we design an interactive tool to quantitatively show how much spatiotemporal event privacy that existing LPPMs can provide. Using this tool, the attendees of the conference could intuitively see how accurate an adversary could infer a user-specific spatiotemporal event from a sequence of locations or even LPPM-protected locations. This tool allows attendees to customize the test under different configurations: the attendees can simulate the input location traces with different mobility patterns, choose one LPPM from the candidates (or without an LPPM) and select different LPPM privacy parameters. The attendees can find that state-of-the-art LPPMs do not always properly protect spatiotemporal event privacy.

Second, we demonstrate how PriSTE can *automatically* or *manually* convert an LPPM for location privacy into one protecting spatiotemporal event privacy in continuous location-based services. Using the algorithms we developed in [2], we can quantify the level spatiotemporal event privacy of an LPPM at each time point, and then adjust the LPPM privacy parameter to satisfy the required level of spatiotemporal event privacy. There are two ways to adjust the LPPM privacy parameter. In the "hands-off" way, a user only needs to initiate an LPPM privacy parameter once at the very beginning (this privacy parameter indicate the location privacy level she wants to enjoy at all time points); then the system will automatically adjust the LPPM privacy parameter for the desired spatiotemporal event privacy. In the "hands-on" way, a user can customize her location privacy protection level at each time, while the system indicates the current level of spatiotemporal event privacy so that user can either reduce or increase her LPPM privacy parameter.

Third, we visualize the trade-off among location privacy, spatiotemporal event privacy, and data utility with respect to different user mobility patterns so that users can explore the interactions among these factors to choose appropriate privacy parameters in different scenarios. Interestingly, a stricter LPPM can satisfy a certain level of spatiotemporal event privacy *without* any adjustment, whereas a more loose LPPM may need to reduce its privacy parameter significantly for protecting the same spatiotemporal event; however, for achieving a specific level of spatiotemporal event privacy, a stricter LPPM is *not* always better in terms of data utility. The attendees can also find that if their simulated trajectories have a significant pattern (e.g., recurrent visits), an LPPM may need a small privacy parameter to achieve the same spatiotemporal event privacy.

2. BACKGROUND

2.1 Spatiotemporal Events

Spatiotemporal events can represent user's secrets about spatiotemporal activities in their real-world, such as "visited a hospital in the last week" or "commuting between Location 1 and Location 2 every morning and afternoon". Let $\mathbb{S} = \{s_1, s_2, \cdots, s_m\}$ be the domain of space, where m is the number of all locations and s_i is one location on the map. A user's trajectory consists of a set of (u, t) denoting the user's location at timestamp t in $\{1, 2, \cdots, T\}$. A predicate

can represent each pair of location and time. For example, a predicate $u^1 = s_3$ indicates (u^1, s_3). If the user is in location s_3 at timestamp 1, then the ground truth of the predicate is true. A spatiotemporal event is defined as a Boolean expression of the (location, time) predicates using the AND, OR, NOT operators, denoted by $\wedge$, $\vee$, $\neg$ respectively.

DEFINITION 2.1 (EVENT). *A spatiotemporal event, denoted by* EVENT*, is a set of (location, time) predicates, i.e.* $u^t = s_i$*, under the Boolean operations.*

Using Boolean logic to define spatiotemporal events enables users to customize their privacy for real-world activities. If a user is at a location s_i at timestamp t, then $u^t = s_i$. If the user is at one location of an area $\{s_i, s_j, \cdots, s_k\}$ at timestamp t, then $(u^t = s_i) \vee (u^t = s_j) \vee \cdots, \vee (u^t = s_k)$ holds. If the user passed through $\{s_i, s_j, \cdots, s_k\}$ over timestamps 1 to T, then $(u^1 = s_i) \wedge (u^2 = s_j) \wedge \cdots, \wedge (u^t = s_k)$ holds.

From the above definitions, we can see that, in terms of privacy goal, spatiotemporal event privacy is a generalization of location privacy; but the privacy notions could be orthogonal as shown in the next section.

2.2 ϵ-Spatiotemporal Event Privacy

Inspired by the definition of differential privacy[4], we define ϵ-Spatiotemporal Event Privacy as follows.

DEFINITION 2.2 (ϵ-SPATIOTEMPORAL EVENT PRIVACY). *A mechanism preserves ϵ-Spatiotemporal Event Privacy for a spatiotemporal* EVENT *if at any timestamp t in $\{1, 2, \cdots, T\}$ given any observations $\{o_1, o_2, \cdots, o_T\}$,*

$$\Pr(o_1, o_2, \cdots, o_t | \text{EVENT}) \leq e^{\epsilon} \Pr(o_1, o_2, \cdots, o_t | \neg \text{EVENT}) \quad (1)$$

where EVENT *is a logic variable about the defined spatiotemporal event and* $\neg$EVENT *denotes the negation of* EVENT. $\Pr(o_1, o_2, \cdots, o_t | \text{EVENT})$ *denotes the probability of the observations $o_1, o_2, \cdots, o_t$ given the value of* EVENT.

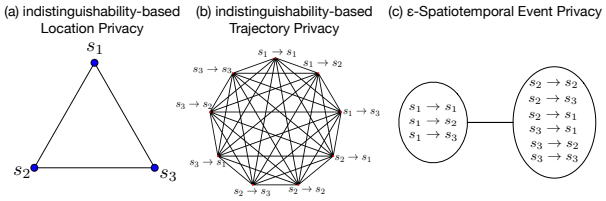

Figure 2: Indistinguishability-based privacy metrics for three types of privacy goals when $\mathbf{S} = \{s_1, s_2, s_3\}$ and $\mathbf{T} = 2$.

To better understand the characteristics of spatiotemporal event privacy, we illustrate the indistinguishability-based privacy metrics for the three privacy goals in Fig.2, where the lines connecting two secrets indicate the requirements of indistinguishability between the corresponding two possible values of the secrets. We can see that these privacy notions are orthogonal due to different structures (see more detailed analysis in an extended version [3] of this work).

3. SYSTEM OVERVIEW

As shown in Fig.3, there are three components in PriSTE framework including LPPM, PrivacyCheck and Visualization. PriSTE calibrates the privacy parameter α of an underlying LPPM (denoted as α-LPPM) at each time point in order to achieve the required ϵ-spatiotemporal event privacy.

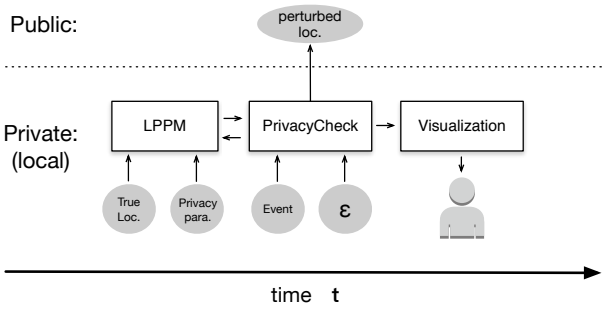

Figure 3: PriSTE framework.

In the LPPM component, we implement two LPPMs for different privacy metrics (i.e., Planar Laplace Mechanism for Geo-Indistinguishability [1] and Planar Isotropic Mechanism for δ-location set privacy [8] [9]) as the candidate LPPMs. Two inputs of LPPM are a true location and the privacy parameter. True locations can be simulated by attendees or selected from a real-world dataset, i.e., Geolife [10]. The privacy parameter determines the location privacy protection level. For both of the two candidate LPPMs, a smaller privacy parameter indicates stronger privacy. In the PrivacyCheck component, the technical challenge is to quantify whether or not the α-LPPM satisfy ϵ-spatiotemporal event privacy w.r.t. user-specified spatiotemporal event. The basic idea is to compute the prior and posterior probabilities of the spatiotemporal event w.r.t. adversaries' prior knowledge π about the initial distribution of the user's possible locations; then we ensure that the spatiotemporal event privacy leakage is bounded w.r.t. any π. In [2], this problem is reduced to a quadratic programming problem, and we designed an algorithm to solve it efficiently.

The interactions between components are described as follows. At each time point, the LPPM component generates a perturbed location from the true location and pass it to PrivacyCheck. The PrivacyCheck component checks whether or not this perturbed location satisfies Equation (1) w.r.t. a user-specified spatiotemporal event. If so, the perturbed location will be released; if not, PrivacyCheck interacts with LPPM to find an appropriate location privacy parameter α either automatically (system decides the next α) or manually (user decides the next α). At each time, PrivacyCheck passes the real-time computation results to the Visualization component for visualized display to users.

4. DEMONSTRATION SCENARIOS

As shown in Fig.3, there are four basic inputs, i.e., true location, LPPM privacy parameter α, the spatiotemporal event to be protected, and its privacy level ϵ. Another input controlled by attendees of this demo is the trajectory pattern. As shown in Fig.4 , in the user dashboard of our demonstration, attendees can simulate input locations by clicking consecutive points on the map, or randomly select a location trace from a real-life database GeoLife [10]. For each click, PriSTE generates a perturbed location and plots the real-time results in the visualization panel. A user can selectively configure the parameters according to different demonstration scenarios below.

4.1 Adversary Posterior

One goal of this demonstration is to raise users' awareness of spatiotemporal event privacy, which may not be protected

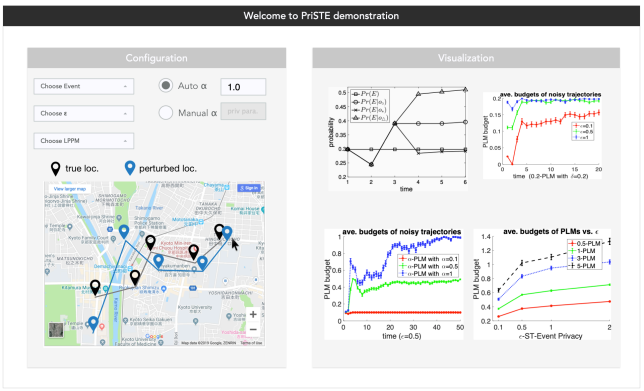

Figure 4: PriSTE Demonstration.

by LPPM. An intuitive and direct way to quantify the spatiotemporal event privacy is to calculate Equation (1) with given prior and observed perturbed (or true) locations. The attendee first configures the value of each input as shown in Fig. 4. Then, she or he can simulate a trajectory by clicking consecutive points on the map. Next, the information about how much adversary prior and posterior w.r.t. the user-specific spatiotemporal event will be plotted in the visualization panel in real time. The difference between prior and posterior indicates the loss of spatiotemporal event privacy. The attendee can observe that for some configurations (for example, recurrent events or a larger privacy parameter α), the spatiotemporal event privacy loss is quite high. It would be interesting to explore how these factors affect spatiotemporal event privacy loss.

4.2 Calibrating LPPM

In the LBS, some "hands-off" users may want to have a uniform location privacy parameter α as an upper bound and do not want to set this value at each time manually; whereas, some "hand-on" users may want more fine-grained controls over the privacy parameter (e.g. allowing a high value or weaker privacy for certain locations such as his/her office). We demonstrate how PriSTE can meet these requirements by *automatically* or *manually* adjusting the privacy parameter of an LPPM. As shown in Fig. 4, the attendee can select "Auto α" or "Manual α" for conservative users and liberal users respectively. In the meantime, a real-time figure of time point (x-axis) versus calibrated LPPM parameter α (y-axis) is plotted. The larger LPPM parameter implies higher data utility (less noise). Even more interestingly, the attendees can select "Manual α" in the configuration panel and then try their best to see if they can outperform (i.e., have larger LPPM parameters than) the system.

4.3 Utility-Privacy Trade-off

One important function of our demonstration is to allow the attendees to explore the data utility under different configurations, which includes different trajectory patterns, LPPMs, location privacy parameters, and spatiotemporal event privacy level. There are two data utility measurements used in our demonstration: one is the LPPM parameters as mentioned in Section 4.2 and the other is Euclidean distance between the perturbed location and the actual location. The attendees can explore the relationship between location privacy and spatiotemporal event privacy and their impact on utility. For example, fixing all parameters in the configuration except LPPM parameter α, we can observe that α does not have a monotonic effect on data utility. An optimal alpha is hard to theoretically capture, but we can find it empirically using this demonstration. This also implies that there is a significant design space for improving the utility of spatiotemporal event privacy mechanisms, and this demonstration can help us discover the insights before we reach elegant theoretical results.

5. CONCLUSION

We demonstrate PriSTE, a framework for protecting spatiotemporal event privacy. PriSTE is featured by quantification-based approach for protecting both location privacy and spatiotemporal privacy so that a user can enjoy the best of two worlds: the underlying LPPM provides general protection against unknown risks, while spatiotemporal event privacy guarantees flexible and customizable protection which may not be provided by the existing LPPMs.

We also present three demonstration scenarios: showing the adversary's posterior about the sensitive spatiotemporal event for raising the users' awareness of such a new privacy goal, providing a user-friendly interface for customizable location privacy protection, and exploring the trade-off between utility and privacy under different configurations.

In summary, the demonstration of PriSTE help the conference attendees to understand the privacy implications and the characteristics of spatiotemporal event privacy. We hope that PriSTE will be a useful tool for protecting users' spatiotemporal event privacy in location-based services.

6. ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 17H06099, 18H04093, 19K20269, National Science Foundation (NSF) under grant No. CNS-1618932 and the AFOSR DDDAS program under grant No. FA9550-121-0240.

7. REFERENCES

[1] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi. Geo-indistinguishability: differential privacy for location-based systems. In *CCS*, pages 901–914, 2013.

[2] Y. Cao, Y. Xiao, L. Xiong, and L. Bai. Priste: From location privacy to spatiotemporal event privacy. In *ICDE*, pages 1606 – 1609.

[3] Y. Cao, Y. Xiao, L. Xiong, L. Bai, and M. Yoshikawa. Protecting spatiotemporal event privacy in continuous location-based services. *arXiv:1907.10814 [cs]*, 2019.

[4] C. Dwork. Differential privacy: A survey of results. In *TAMC*, pages 1–19, 2008.

[5] V. Primault, A. Boutet, S. B. Mokhtar, and L. Brunie. The long road to computational location privacy: A survey. *IEEE Communications Surveys Tutorials*, 2018.

[6] R. Recabarren and B. Carbunar. What does the crowd say about you? evaluating aggregation-based location privacy. In *Proceedings on Privacy Enhancing Technologies*, volume 2017, pages 156–176, 2017.

[7] S. Takagi, Y. Cao, Y. Asano, and M. Yoshikawa. Geo-graph-indistinguishability: Protecting location privacy for LBS over road networks. In *DBSec*, pages 143–163, 2019.

[8] Y. Xiao and L. Xiong. Protecting locations with differential privacy under temporal correlations. In *CCS*, pages 1298–1309, 2015.

[9] Y. Xiao, L. Xiong, S. Zhang, and Y. Cao. LocLok: location cloaking with differential privacy via hidden markov model. *VLDB*, 10(12):1901–1904, 2017.

[10] Y. Zheng, X. Xie, and W.-Y. Ma. GeoLife: a collaborative social networking service among user, location and trajectory. *IEEE Data Eng. Bull.*, 33(2):32–39, 2010.