PriSTE: From Location Privacy to Spatiotemporal Event Privacy

Yang Cao*¶‖, Yonghui Xiao‡**, Li Xiong†¶, Liquan Bai§¶
¶*Emory University*, Atlanta, U.S.A
‖*Kyoto University*, Kyoto, Japan
**Google Inc.*, Mountain View, U.S.A
Email: *yang@i.kyoto-u.ac.jp †lxiong@emory.edu, ‡yohu@google.com, §liquan.bai@emory.edu

***Abstract*—Location privacy-preserving mechanisms (LPPMs) have been extensively studied for protecting a user's location at each time point or a sequence of locations with different timestamps (i.e., a trajectory). We argue that existing LPPMs are not capable of protecting the sensitive information in user's spatiotemporal activities, such as "visited hospital in the last week" or "regularly commuting between Address 1 and Address 2 every morning and afternoon" (it is easy to infer that Addresses 1 and 2 may be home and office). To address this problem, we define the *spatiotemporal event* as a new privacy goal, which can be formalized as Boolean expressions between location and time predicates. We show that the spatiotemporal event is a generalization of a single location or a trajectory which is protected by existing LPPMs, while some types of spatiotemporal event may not be protected by the existing LPPMs. Hence, we formally define ϵ-*spatiotemporal event privacy* which is an indistinguishability-based privacy metric. It turns out that, interestingly, such privacy metric is orthogonal to the existing indistinguishability-based location privacy metric such as Geo-indistinguishability. We also discuss the potential solution to achieve both ϵ-spatiotemporal event privacy and Geo-indistinguishability.**

I. Introduction

The continued advances and usage of smartphones and GPS-enabled devices have provided tremendous opportunities for Location-Based Service (LBS), such as Google Maps, Yelp and Uber. In location-based services, mobile users have to share their locations or trajectories with the service providers in order to issue snapshot or continuous queries, for example, "where is the nearest gas station" or "continuously report the taxis within one mile of my location". It has raised privacy concerns as users' digital trace can be used to infer sensitive information, such as home and work place, religious places and sexual inclinations [1] [2] [3].

A large number of studies (see surveys [4] [5] [6] [7]) have explored how to protect user's location privacy from different aspects: 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 method for the privacy goal (e.g., Geo-indistinguishability [8] or δ-location set privacy [9]); LPPMs study how to achieve a specified privacy metric.

We argue that existing LPPMs may not fully protect users' sensitive information in their spatiotemporal activities because their privacy goal is focused on protection of either a single location or a trajectory. A user's location data can be represented by a tuple (we consider the single user setting in this paper), i.e., <*location, time*>, which corresponds to information about "where" and "when" in user's real-world activities. Hence, the privacy goals in literature can be categorized into protecting a single *position* or a *trajectory*. Many LPPMs are proposed for these goals based on different privacy metrics. For example, Gruteser et al. [10] designed a spatiotemporal cloaking mechanism satisfying k-anonymity to protect movement trajectories of users; Andrés et al. [8] proposed Planar Laplace mechanism [8] achieving Geo-indistinguishability to protect single locations; Xiao and Xiong [9] designed Planar Isotropic Mechanism for δ-location set privacy to protect each location in a trajectory.

However, the privacy goals in the literature of location privacy only focus on protecting a user's exact location or a trajectory, and cannot cover all cases of complex combination of spatial and temporal information (as shown in Fig.1), which we refer to as *spatiotemporal events* in this paper. Examples of the spatiotemporal event include "visited hospital in the last week" (i.e., the hospital visit may happen once or multiple times at any time in last week) and "regularly commuting between Address 1 and Address 2 every morning and every afternoon" (these periodic spatiotemporal events may happen every day).

We show six cases of the Boolean expression between location and time predicates in Fig.1. It turns out that protecting a single location or a trajectory are only two special cases in protecting a user's spatial and temporal information. Let u^t be a user's position at time t, and $s_i \in \mathbb{S}, i \in [1, m]$ be one of all m locations on the map. The element of a user's secrets in her spatiotemporal activities can be represented by a predicate $u^t = s_i$ (the value can be either *true* or *false*). Then, a *spatiotemporal event* can be defined as a Boolean expression by combining different predicates over spatial and/or temporal dimensions (a predicate alone also can be a spatiotemporal event). As shown in Fig.1, the events representing a sensitive location/area and a trajectory, which are the main focuses

Yang and Yonghui contributed equally to this work.

in previous studies of location privacy, are only two cases (i.e., (b) and (c)) in the six enumerated examples. Even if each location or a trajectory is protected, it is not clear whether or not an adversary can infer the value of a sensitive spatiotemporal event. Protecting the privacy of spatiotemporal events has not been studied in literature.

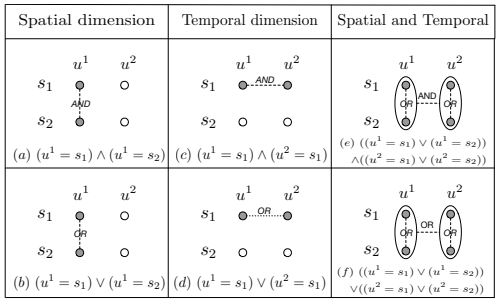

Fig. 1: Examples of spatiotemporal events. s_1 and s_2 are two locations on the map $\mathbb{S}$. u^1 and u^2 are two variables about user's possible locations at time point 1 and time point 2, respectively. Event (a) is always false since a user cannot be at two different locations at the same time. 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.

Although an LPPM protecting a single location or a trajectory (i.e., Fig.1(b) or Fig.1(c)) ensures user's location privacy, it is not clear whether an LPPM also provide a certain level of spatiotemporal event privacy. In this paper, we formalize the new privacy goal of spatiotemporal event privacy by defining ϵ-spatiotemporal event privacy, and propose a framework, i.e., PriSTE (Private SpatioTemporal Event), for protecting spatiotemporal event privacy using existing LPPMs.

II. Spatiotemporal Event Privacy

A. Problem Setting

We study how to protect spatiotemporal event in a single user setting. Consider a user who is sharing her location sequence with a location-based service provider. Since the server may not be trusted or can be compromised by other malicious parties, the user does not want to share her sensitive information with the server (we also assume that the attacker/sever does not have the knowledge of user's predefined spatiotemporal event(s)); instead, she uses a local LPPM that guarantees location privacy at each timestamp. We denote a moving user's true locations as $\{u^1, u^2, \cdots, u^T\}$. The LPPM blurs user's true location u^t to a perturbed one o_t that satisfies a privacy metric such as *geo-indistinguishability* [8] or δ-*location set privacy* [9]. Hence, an LPPM can be considered as an emission matrix that takes user's true location as input and outputs a perturbed one.

B. Spatiotemporal Events

Spatiotemporal events can represent user's secrets in their real-world activities, such as "visited hospital in the last week" or "commuting between Address 1 and Address 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 (we use *state* interchangeably) on the map. A user's trajectory consists of a set of $\{u, t\}$ denoting the user's location at timestamp t in $\{1, 2, \cdots, T\}$. Each pair of location and time can be represented by a predicate. For example, the pair $\{u^1, s_3\}$ can be denoted by a predicate $u^1 = s_3$. 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 II.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 preference for diverse real-world activities. Table I shows some representative examples of EVENT. If a user is in a state s_i at timestamp t, then $u^t = s_i$. If the user is in a region of $\{s_i, s_j, \cdots, s_k\}$ at timestamp t, then $(u^t = s_i) \vee (u^t = s_j) \vee \cdots, \vee (u^t = s_k)$ holds. If the trajectory of the user is $\{s_i, s_j, \cdots, s_k\}$ over timestamps 1 to T, then $(u^1 = s_i) \wedge (u^2 = s_j) \wedge \cdots, \wedge (u^t = s_k)$ holds. Based on the Boolean operations, complicated spatiotemporal events can be defined as follows.

PRESENCE. When the secret is whether or not a user visited a sensitive area (e.g., medical facilities) in a given time period, we can use PRESENCE to represent such secret. A PRESENCE event holds if a user appears in a region during some time. In the simplest case, the region consists of one location, and time period consists of one timestamp, then it becomes one single location shown in Table I. Hence, PRESENCE is a generalization of secrets about single locations. To denote a region, which is a set of locations, we use a vector $\mathbf{s} \in \{0, 1\}^{m \times 1}$ where the ith element is 1 if the region contains s_i. The time period is denoted by $\mathbf{T}$ as a set of timestamps.

Definition II.2 (PRESENCE)**.** *Given a set of regions* $\boldsymbol{S}$ *and a time period* $\boldsymbol{T}$*, if a user appears in* $\boldsymbol{s}$ *at any timestamp* $t \in \boldsymbol{T}$*, then it is a presence event, denoted by* PRESENCE$(\boldsymbol{S}, \boldsymbol{T})$.

Example II.1 (PRESENCE)**.** *Fig.2 shows a map of* $\mathbb{S} = \{s_1, s_2, s_3\}$*. The shaded region shows a* PATTERN *event that the user appears in a region of* s_1 *or* s_2 *during timestamps* 3 *and* 4*. The lines indicate possible trajectories. As long as user's true trajectory passes through the shaded region, the ground truth of the event is true. For this event, the region* $\boldsymbol{s} = [1, 1, 0]^\intercal$ *denoting the states* s_1 *and* s_2*; the time period* $\boldsymbol{T} = \{3, 4\}$ *denoting timestamp* 3 *and* 4*. The* PRESENCE *event is expressed as* $(u^3 = s_1) \vee (u^3 = s_2) \vee (u^4 = s_1) \vee (u^4 = s_2)$.

PATTERN. When the secret is whether or not a user visited multiple sensitive areas successively (e.g., a love hotel and

EVENT	Boolean Expression	Interpretation
single location	$u^t = s_i$	the location at timestamp t is s_i
PRESENCE at s_i during T	$(u^1 = s_i) \vee (u^2 = s_i) \vee \cdots \vee (u^t = s_i)$	appears at location s_i during time $\{1, 2, \cdots, T\}$
PRESENCE at $\mathbf{s}$ and t	$(u^t = s_i) \vee (u^t = s_j) \vee \cdots, \vee (u^t = s_k)$	appears in region $\mathbf{s} : \{s_i, \cdots, s_k\}$ at timestamp t
single trajectory	$(u^1 = s_i) \wedge (u^2 = s_j) \wedge \cdots, \wedge (u^n = s_k)$	a trajectory of locations during a time period
PATTERN of trajectories	$((u^1 = s_i) \vee (u^1 = s_j) \vee \cdots, \vee (u^1 = s_k)) \wedge \cdots$ $\wedge ((u^n = s_l) \vee (u^n = s_m) \vee \cdots, (u^n = s_n))$	a PATTERN of trajectories

TABLE I: EVENTS of Boolean operations on the (location, time) predicates

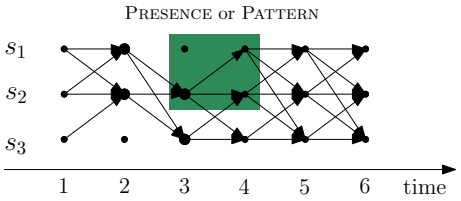

Fig. 2: PRESENCE: the user appears in regions of $\{s_1, s_2\}$ during timestamps 3 and 4; PATTERN: the trajectories all go through $\{s_1, s_2\}$ at timestamp 2 and $\{s_2, s_3\}$ at timestamp 3.

then home), we can use PATTERN to represent such secret. In a simple case, the regions consist of single locations at a set of timestamps, then it becomes single trajectory shown in Table I. Hence, PATTERN is a generalization of secrets about user's trajectories. We define PATTERN as follows where the set of regions $[\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_n]$ is denoted by $\mathbf{S}$.

Definition II.3 (PATTERN)**.** *Given a sequence of regions* $\boldsymbol{S} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n\}$ *and a time period* $\boldsymbol{T}$ *where* $\boldsymbol{s}_i \in \{0, 1\}^{m \times 1}$*, if a user appears in* $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n\}$ *sequentially during* $\boldsymbol{T}$*, then it is a pattern event, denoted by* PATTERN$(\boldsymbol{S}, \boldsymbol{T})$.

Example II.2 (PATTERN)**.** *Fig.2 shows a set of trajectories with a* PATTERN *that all trajectories go through* $\{s_1, s_2\}$ *at timestamp* 2 *and* $\{s_2, s_3\}$ *at timestamp* 3*. For this event, the region at timestamp* 2 *is* $\boldsymbol{s}_2 = [1, 1, 0]^\intercal$ *denoting* s_1 *and* 2_2*; the region at timestamp* 3 *is* $\boldsymbol{s}_3 = [0, 1, 1]^\intercal$ *denoting* s_2 *and* s_3*. The* PATTERN *event is expressed as* $((u^2 = s_1) \vee (u^2 = s_2)) \wedge ((u^3 = s_2) \vee (u^3 = s_3))$.

From the above definitions, we can see that, in terms of privacy goal, spatiotemporal event privacy is a generalization of location privacy. In this paper, we focus on the two representative events defined above, i.e., PRESENCE and PATTERN, which are the two most complicated events in examples of Fig.1. We note that PRESENCE and PATTERN include the cases when the time $\mathbf{T}$ is not consecutive. For simplicity, we assume that the events are defined in consecutive time and use $start$ and end to denote the start point and end point of the defined spatiotemporal event. Users can customize one or multiple spatiotemporal events to be protected. We need a formal privacy metric to preserve user's *plausible deniability* about the truth of her specified spatiotemporal events. We propose such a privacy metric for spatiotemporal event privacy in the next section.

C. ϵ-Spatiotemporal Event Privacy

Inspired by the definition of differential privacy [11], we define ϵ-Spatiotemporal Event Privacy as follows.

Definition II.4 (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.

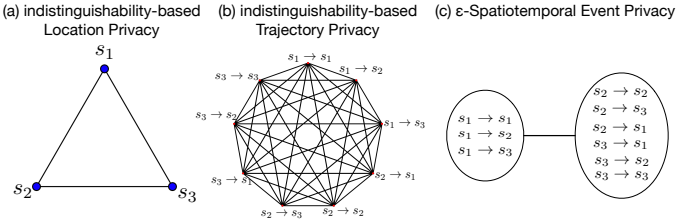

Fig. 3: Illustration of indistinguishability-based privacy metrics for distinct privacy goals when $\mathbb{S} = \{s_1, s_2, s_3\}$ and $T = 2$.

There are two major benefits of adopting such "DP-like" privacy metric. First, it provides a well-defined semantics for spatiotemporal event privacy. Similar to differential privacy that requires the indistinguishability between any two neighboring databases [11], ϵ-Spatiotemporal Event Privacy requires the indistinguishability regarding whether the EVENT is true or false given any observations. Another benefit is that, similar to differential privacy whose privacy guarantee is independent of the prior probability of a given databases, the privacy provided by ϵ-Spatiotemporal Event Privacy is independent of the prior probability of the given spatiotemporal event.

Although the *privacy goal* of spatiotemporal event privacy can be considered as a generalization of location privacy, we note that it may not be true in terms of *privacy metrics*. We illustrate the indistinguishability-based privacy metrics for the three privacy goals in Fig.3, where the lines connecting two secrets indicate the requirements of indistinguishability between the corresponding two possible values of the secrets.

As shown in Fig.3 (a), indistinguishability-based location privacy metrics (such as geo-indistinguishability [8]) require indistinguishability between each pair of locations. Indistinguishability-based trajectory privacy metrics [9] [12] [13] requires indistinguishability between each pair of possible trajectories as shown in Fig.3 (b). Whereas, ϵ-spatiotemporal event privacy only requires indistinguishability between the defined event and its negation. For example, if the spatiotemporal event is defined as PATTERN$(\mathbf{S}, \mathbf{T})$ where $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2\}, \mathbf{s}_1 = \{s_1\}, \mathbf{s}_2 = \{s_1, s_2, s_3\}$ and $\mathbf{T} = \{1, 2\}$ (i.e., a trajectory passes through s_1 and $\{s_1, s_2, s_3\}$) successively, then it only requires the indistinguishability between the set of all possible trajectories that pass through $\{s_1\}$ and $\{s_1, s_2, s_3\}$

and the set of trajectories that do not. Such privacy requirement makes sense; for example, s_1 can be a "love hotel", s_2 is "home", and s_3 is "office". Such spatiotemporal event privacy may not be protected by LPPMs that are designed to ensure indistinguishability between pairs of locations or trajectories as shown in Fig.3(a) and 3(b).

While we can define simple events such as a location or trajectory, the corresponding ϵ-spatiotemporal event privacy does not imply the indistinguishability-based location privacy or trajectory privacy. For example, even if a user specifies all possible trajectories as her requirements for ϵ-spatiotemporal event privacy, it only ensures the indistinguishability between "one trajectory" and "not this trajectory", but no guarantee on the indistinguishability between any two trajectories. Hence, the privacy guarantee of ϵ-spatiotemporal event privacy is orthogonal to geo-indistinguishability.

Location privacy provides general protection against unknown risks when sharing location with the third parties, while spatiotemporal event privacy guarantees flexible and customizable protection which may prevent against profiling attacks such as inferring user's trajectory pattern (location privacy cannot provide such protection). Therefore, it would be preferable that an LPPM achieving location privacy metrics such as geo-indistinguishability also satisfies ϵ-spatiotemporal event privacy, which will be discussed in the next section.

III. PriSTE framework

In this section, we propose a framework PriSTE for achieving both spatiotemporal event privacy and geo-indistinguishability. The idea is to adopt an existing LPPM and adjust its privacy level for achieving spatiotemporal event privacy. For example, we can adjust the privacy parameter of Planar Laplace Mechanism [8], which is designed for geo-indistinguishability, to satisfy the requirement of ϵ-spatiotemporal event privacy w.r.t. the given event(s).

The PriSTE framework is described in Algorithm 1. At each time point t, the algorithm takes the true location as input, and outputs a perturbed location o_t that satisfies ϵ-spatiotemporal event privacy for continuous release and geo-indistinguishability for each single location. In Line 2, a perturbed location o_t is generated based on an LPPM (such as Planar Laplace Mechanism). Since we are not sure whether this location satisfies ϵ-spatiotemporal event privacy, we need a quantification module which is involved in Line 3. The quantification module can properly answer whether such perturbed location may disclose too much information about the predefined spatiotemporal event w.r.t. an informed adversary who has knowledge of the LPPM and the user's mobility pattern, i.e, transition matrix between locations. If such LPPM is not able to provide ϵ-spatiotemporal event privacy w.r.t. the given event, we calibrate the privacy parameter of LPPM (e.g., reduce the privacy parameter of Planar Laplace Mechanism) as shown in Line 4 until it satisfies ϵ-spatiotemporal event privacy. Finally, we release o_t in Line 6.

A significant challenge to implement the framework is the computational complexity of quantifying the spatiotemporal event privacy loss w.r.t. a given LPPM (i.e., Lines 3∼5). For example, given a complex spatiotemporal event, i.e., a Boolean expression, checking its value (true or false) requires enumeration of all possible values of the predicates in the Boolean expression, which is exponential to the number of predicates. Due to the space limitation, we omit the technical details of our solution for addressing this issue and refer reader to the full version of this work [14].

Algorithm 1 PriSTE Framework

Require: true location, ϵ, α, LPPM, **M**, EVENTS
1: **for** t in $\{1, 2, \cdots, T\}$ **do**
2: generate o_t with LPPM w.r.t. the true location;
3: **while** ϵ-Spatiotemporal Event Privacy not hold **do**
4: *calibrate* the privacy level of LPPM and re-generate o_t;
5: **end while**
6: release o_t;
7: **end for**

IV. Conclusion

In this paper, we formalized spatiotemporal event, which is a more general privacy goal than location privacy considered in the literature. We formally defined ϵ-spatiotemporal event privacy, which is orthogonal to the state-of-the-art location privacy metric such as Geo-indistinguishability. We proposed a framework PriSTE that achieves both spatiotemporal event privacy and location privacy.

V. Acknowledgment

This work is supported by NSF under grant No. 1618932, the AFOSR DDDAS program under grant FA9550-121-0240, the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (S) No. 17H06099 and (A) No. 18H04093.

References

[1] P. Golle and K. Partridge, "On the anonymity of Home/Work location pairs," in *Lecture Notes in Computer Science*, 2009, pp. 390–397.
[2] R. Recabarren and B. Carbunar, "What does the crowd say about you? evaluating aggregation-based location privacy," in *WPES*, vol. 2017, 2017, pp. 156–176.
[3] G. Argyros, T. Petsios, S. Sivakorn, A. D. Keromytis, and J. Polakis, "Evaluating the privacy guarantees of location proximity services," *ACM Trans. Priv. Secur.*, vol. 19, no. 4, pp. 12:1–12:31, 2017.
[4] J. Krumm, "A survey of computational location privacy," *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 391–399, 2009.
[5] M. Wernke, P. Skvortsov, F. Dürr, and K. Rothermel, "A classification of location privacy attacks and approaches," *Personal and Ubiquitous Computing*, vol. 18, no. 1, pp. 163–175, 2014.
[6] K. Chatzikokolakis, E. ElSalamouny, C. Palamidessi, and P. Anna, "Methods for location privacy: A comparative overview," *Foundations and Trends in Privacy and Security*, vol. 1, no. 4, pp. 199–257, 2017.
[7] B. Liu, W. Zhou, T. Zhu, L. Gao, and Y. Xiang, "Location privacy and its applications," *IEEE Access*, pp. 17 606–17 624, 2018.
[8] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: differential privacy for location-based systems," in *CCS*, 2013, pp. 901–914.
[9] Y. Xiao and L. Xiong, "Protecting locations with differential privacy under temporal correlations," in *CCS*, 2015, pp. 1298–1309.
[10] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *MobiSys*, 2003, pp. 31–42.
[11] C. Dwork, "Differential privacy: A survey of results," in *TAMC*, 2008, pp. 1–19.
[12] K. Chatzikokolakis, C. Palamidessi, and M. Stronati, "A predictive differentially-private mechanism for mobility traces," in *Lecture Notes in Computer Science*, 2014, no. 8555, pp. 21–41.
[13] G. Theodorakopoulos, R. Shokri, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec, "Prolonging the hide-and-seek game: Optimal trajectory privacy for location-based services," in *WPES*, 2014, pp. 73–82.
[14] Y. Cao, Y. Xiao, L. Xiong, and L. Bai, "PriSTE: from location privacy to spatiotemporal event privacy," *arXiv:1810.09152 [cs]*, 2018.