A Unified Framework for Location Privacy

Reza Shokri, Julien Freudiger and Jean-Pierre Hubaux

LCA, EPFL, Switzerland
firstname.lastname@epfl.ch
EPFL-REPORT-148708 June 2010

Abstract. We introduce a novel framework that provides a logical structure for identifying, classifying and organizing fundamental components, assumptions, and concepts of location privacy. Our framework models mobile networks and applications, threats, location-privacy preserving mechanisms, and metrics. The flow of information between these components links them together and explains their interdependencies. We demonstrate the relevance of our framework by showing how the existing achievements in the field of location privacy are embodied appropriately in the framework. Our framework provides "the big picture" of research on location privacy and hence aims at paving the way for future research.

1 Introduction

Location-based services are becoming ubiquitous, fueled by the proliferation of mobile devices, notably smart phones. There exist numerous applications that take advantage of the latest capabilities of mobile devices, in order to share information between users in a wireless peer-to-peer manner [1–3], or to exchange location-based information with the service providers [4, 5, 17, 20, 46, 62].

Despite the popularity of these services, privacy issues such as the undesired leakage of users' location information to location-based service operators, or to external eavesdroppers is a major concern. In this paper, we are concerned with *location privacy* that, according to Duckham and Kulik [27], is defined as a *special type of information privacy which concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others.*

The problem of protecting the location privacy of mobile users has attracted researchers from various backgrounds such as database security and anonymous communication. Several works approach this problem from different perspectives and, hence, various protection mechanisms have been proposed. However, existing proposals usually aim at solving location-privacy problem, based on incomplete (and sometimes not fully explicit) set of assumptions, that are not fully in accordance with network and adversary models and users' privacy requirements. Without a systematic identification of the possible threats, specification of available countermeasures, and designing of appropriate evaluation metrics, there is no guarantee that the location privacy of users can be actually protected in

different scenarios. In particular, the adversary's knowledge, her available technologies, access rights, potential actions, and her goals should be formalized in order to enable the design of effective countermeasures. The metrics used for measuring location privacy should be carefully designed in order to reflect the users' actual gain in employing a location-privacy preserving mechanism, by considering the users' privacy requirements and the adversary's knowledge.

Due to these considerations, we are motivated to construct a unified framework for location privacy in which the different components, that affect location privacy, are defined and their interrelations are identified. A consistent structure and terminology, proposed in this framework, allows us to better understand different works and thus position them appropriately in the field of location-privacy research. We provide a thorough study of previous works and place each of them appropriately in our framework, based on the role it plays in protecting location privacy. The framework, further, paves the way for future research in this field, by introducing a common notion for location privacy. It will help identifying the shortcomings of the existing approaches and discovering neglected aspects and hence open problems of location privacy.

In our framework, we introduce the various components of location privacy according to the flow of information from the users to the adversary. A location-privacy preserving mechanism acts as a noisy channel that modifies the information that is communicated from the users (as the source of information) to the adversary (as the observer/receiver). Users' location privacy is maximized if the adversary cannot correctly link their location and identity over time. In other words, using the information she obtains by observing users' activities from behind the curtain of privacy preserving mechanisms, more distorted is users' location in the adversary's eyes, the higher their location privacy is.

The structure of the paper is as follows. First, in Section 2, we model *mobile networks* and formalize the state of the real world (i.e., the users' mobility). Further, in Section 3, we model the spatiotemporal image of users' activities after being distorted by *location-privacy preserving mechanisms*. In Sections 4 and 5, we provide an elaborate model for the *adversary*, different categories of location privacy, the methodology for measuring location privacy, and an exhaustive study of existing *metrics*. In Section 6, using our framework, we briefly model and study location privacy in an example scenario: location-based services. Finally, in Section 7, we survey the existing models in the literature, before concluding the paper in Section 8.

To the best of our knowledge, this is the first paper that takes the many aspects of location privacy into account, clarifies their interrelation, and proposes a unified framework in which the existing achievements are embodied.

2 Mobile Networks

A mobile wireless network consists of a set of mobile users equipped with wireless devices that are capable of establishing ad hoc communication among themselves and/or connecting to infrastructures (e.g., cellular networks, and WiFi access

points) in order to use a *common service* using the appropriate *application*. There exist many applications that can help users to access the provided services in mobile networks. Location-based services, mobile social networking, mobile recommender systems, friend finder services, people-centric sensing systems, and ad hoc networking are some examples of the services provided in mobile networks.

Users can employ a wide range of applications in mobile networks. These applications can be categorized into subsets based on their different features. Because our focus is on how users' location can be used by an application, we divide the applications into two dimensions: *automatic* vs. *manual*, based on the way information is shared by the application; *continuous* vs. *discrete*, based on the time distribution of information sharing. As an example, people-centric sensing applications, where users upload information about their environment to a central server, are *automatic* and *continuous*. Electronic ticketing, which is used in public transportation systems, is *automatic* and *discrete*. Some of the applications such as location-based services encompass a variety of applications, each of which has different features. However, most of the popular location-based services are *manual* and *discrete*.

2.1 Users

We denote by $\mathcal{U}$ the set of users who are members of the mobile network. Depending on the provided service, users might also belong to online social networks. We assume that each and every user is associated with two distinct types of identities: A *real identity* and a set of *pseudonyms*.

The real identity of a user is any subset of his attributes (e.g., name, national identity number, and private key) that uniquely identifies the user within the set of users $\mathcal{U}$ [61]. The real identity of each user is unique and invariant over time. Let $\mathcal{I}$ denote the set of real identities associated with all users. There is a bijective relation between a user and his real identity. This is captured by function $\mathbf{name} : \mathcal{U} \rightarrow \mathcal{I}$ that maps each user with his real identity.

Every user is also associated with a set of *pseudonyms*. Pseudonyms are mostly temporary identifiers that facilitate identification and authentication of a user in a communication without revealing the *user*'s real identity. MAC addresses of wireless devices, IP addresses, public keys, group signatures, physical layer fingerprints of wireless devices, signature of the applications, and the combination of these identities are different examples of pseudonyms. The user to whom a particular pseudonym refers is called the *holder* of the pseudonym. If it is shared by multiple users, a pseudonym is called a *group pseudonym*. Based on the information derived from a pseudonym *per se*, no one (except the pseudonym's holder) can identify the holder of the pseudonym [61]. Let $\widehat{\mathcal{I}}$ denote the set of pseudonyms used by all the users. The set of real identities and pseudonyms are disjoint, i.e., $\widehat{\mathcal{I}} \cap \mathcal{I} = \emptyset$. We define function $\mathbf{nyms} : \mathcal{U} \rightarrow \mathcal{P}(\widehat{\mathcal{I}})$ that gives the set of pseudonyms associated with each user, where $\mathcal{P}(\widehat{\mathcal{I}})$ is the power set of $\widehat{\mathcal{I}}$. A user can use various pseudonyms over time. Pseudonyms can expire or be revoked from the network. Lastly, we define $\widehat{\mathcal{I}}(t) \subseteq \widehat{\mathcal{I}}$ to denote the set of pseudonyms that are still valid at time t.

2.2 Time and Space

The notion of time that we employ in this model is discrete, and each unit of time, which is a natural number, is called a *time instance*. We define a *time period* (t_s, t_e) to be the set of all time instances between t_s and t_e including t_s and t_e. Let $\mathcal{T}$ be the time period in which we model the system.

We employ a three-layer model in order to represent the space in which users can move, and its associated contextual information. The first layer models the *geographical space*. The second layer models the places or *location sites*, and the third layer models the *type of sites*.

We model the first layer, i.e., the geographical space in which users can move, in a discrete way. Let $\mathcal{L} \subset \mathbb{N}^2$ be the grid that represents the two dimensional space. We name any member of the set $\mathcal{L}$ a *location instance*. We also call a non-empty set of location instances a *location area*. Each location instance belongs to a *location site* that has a unique name and address (e.g., a hospital, an avenue, or a house with its unique ZIP code). In other words, any location instance has a tag name, and clearly, multiple location instances can have the same name because they belong to the same site. The location sites, which are constructed on top of the location instances, build the second layer of the model. The third layer captures the type of location sites and their similarity. A location type represents the usage of the location, e.g., shopping, cultural, sporting, residential.

Note that the granularity of time and location depends on the application and we do not make any assumption about that.

2.3 The Spatiotemporal State of Users

As users are mobile, their location is a time-dependent value. To model the users mobility, let $\mathbf{whereis} : \mathcal{U} \times \mathcal{T} \rightarrow \mathcal{L}$ be a function that gives the *actual location* of users at any time instance in $\mathcal{T}$. Note that this function gives the exact location of each user, regardless of the knowledge of any entity about it (i.e., maybe even the user himself does not know his accurate position, because of, for example, the inaccuracy of his GPS device).

We model a mobile network in terms of the location-based events that occur in the network. Events in our framework reflect the spatiotemporal state of the users in the real world and also in the eyes of an observer.

We define an *event* as a 3-tuple $\langle i, t, l \rangle$, where $i \in \mathcal{I} \cup \widehat{\mathcal{I}}$ is the real identity or a pseudonym of a user, t is the time instance at which the event occurs (referred to as the *time-stamp* of the event), and $l \in \mathcal{P}(\mathcal{L})$ is the location area associated with the event (referred to as the *location-stamp* of the event). Let $\mathcal{E}$ denote the set of all possible events. We define three functions $\mathbf{id} : \mathcal{E} \rightarrow \mathcal{I} \cup \widehat{\mathcal{I}}$, $\mathbf{tm} : \mathcal{E} \rightarrow \mathcal{P}(\mathcal{T})$, and $\mathbf{loc} : \mathcal{E} \rightarrow \mathcal{P}(\mathcal{L})$ that give the identity, time-stamp, and location-stamp of an event, respectively. Thus, for any event $e \equiv \langle i, t, l \rangle$ we have $\mathbf{id}(e) = i$, $\mathbf{tm}(e) = t$, and $\mathbf{loc}(e) = l$. We further define a *trace* to be a non-empty set of events.

An event e is called an *actual event* associated with a user u if the following conditions hold: $\mathbf{id}(e) = \mathbf{name}(u)$, $\mathbf{loc}(e) = \mathbf{whereis}(u, \mathbf{tm}(e))$, and $\mathbf{tm}(e) \in \mathcal{T}$. Thus, an actual event represents the spatiotemporal status of a user in the real

world. Following the definition of actual events, the actual trajectory of a user is the trace of all his actual events. Let $\mathcal{R}_u$ denote the trajectory of a user during $\mathcal{T}$. Hence, $\mathcal{R}_u = \{\langle \mathbf{name}(u), t, \{\mathbf{whereis}(u,t)\}\rangle, \forall t \in \mathcal{T}\}$.

We define relation $\sim$ between two actual events e_i and e_j, if and only if they are associated with the same user, i.e., $e_i \sim e_j \Leftrightarrow \exists u \in \mathcal{U}$ s.t. $e_i, e_j \in \mathcal{R}_u$. Let $\mathcal{R}$ denote the set of all actual events of all users. Thus, $\mathcal{R} = \bigcup_u \mathcal{R}_u$, and $\forall u, v \in \mathcal{U}$ we have $\mathcal{R}_u \cap \mathcal{R}_v = \emptyset$. Hence, the relation $\sim$ partitions the trace $\mathcal{R}$ into subsets, each representing the actual trajectory of one user. We refer to the set partition associated with $\sim$ as the *actual set partition* of $\mathcal{R}$.

3 Location-Privacy Preserving Mechanisms

Actual events represent the state of the real world, i.e., the actual locations of users over time. A user has absolutely no location privacy if an adversary has access to the set of his actual events. To protect a user's location privacy, some privacy preserving mechanisms should alter the information observable by the adversary. Ideally, the amount of information leakage should be minimal, while enabling users a proper use of the service. There are three entities that play a role in preserving location privacy: *users*, *applications*, and *privacy tools*. Each entity controls the amount of shared information and thus affects user privacy. Users and applications might *intentionally* (e.g., by being cautious about sharing unnecessary information) or *unintentionally* (e.g., by sharing incorrect information) reduce the amount of information revealed. Privacy policies influence the way applications can share information with different entities, and they are applied to the application based on the users' decisions. Various privacy tools, also, use sophisticated algorithms to guarantee users' privacy.

In order to capture the effect of the three above-mentioned entities in preserving location privacy of users, in our framework, we abstract away the entities and model a location-privacy preserving mechanism as a single unit that separates actual events of the users (i.e., the ground truth) and the adversary. Formally we define a *location-privacy preserving mechanism* as a transformation function that modifies the users' actual events before they can become observable by any observer.

Privacy tools (as the main entity in location-privacy preserving mechanisms) work in three architectures: (i) *Distributed* (user-side): They can work in a distributed way by being implemented on individual mobile devices, where each device itself transforms its events and modifies what an observer can see about the user's spatiotemporal state. This can be done either with the help of information that a device gets from other devices or exclusively with the information that the user has himself. (ii) *Centralized* (server-side): They can work in a centralized manner by using a trusted central server that acts as a privacy preserving proxy and modifies users' messages (correspond to events in our model) before being observable by an untrusted entity. (iii) *Hybrid*: They can be a hybrid of both distributed and centralized architectures.

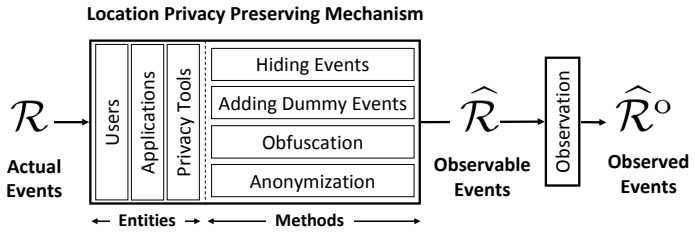

Fig. 1. Location Privacy Preserving Mechanisms

We define function $\mathbf{trns} : \mathcal{E} \rightarrow \mathcal{E} \cup \{\text{HIDDEN}\}$ to denote the location privacy preserving mechanisms, where HIDDEN stands for a *hidden event* (i.e., an unobservable event). The output events of the transformation function on the set of actual events is called the set of *observable events* and is denoted by $\widehat{\mathcal{R}}$, i.e., $\widehat{\mathcal{R}} = \overrightarrow{\mathbf{trns}}(\mathcal{R}) \setminus \{\text{HIDDEN}\}$, where $\overrightarrow{\mathbf{trns}}()$ is the image of function $\mathbf{trns}()$ on a trace. Mechanisms perform the transformation function by means of four different *methods* (which can also be called the primitives of location-privacy preserving mechanisms): *hiding events*, *adding dummy events*, *obfuscation*, and *anonymization*. These four methods together can model any transformation on the events: the first two (*hiding events*, and *adding dummy events*) modify the set of the events, *obfuscation* modifies time- and location-stamps, and *anonymization* modifies identity of events. Figure 1 illustrates the role of different entities that can employ a combination of these methods in order to alter the accessible information to an observer. The methods are described below.

- **Hiding Events** The most basic method for protecting users' location privacy consists in hiding information about the trajectories of users. A subset of events are removed in the transformation process. This is modeled by replacing the candidate events by HIDDEN in the **trns** function, which will not further appear in $\widehat{\mathcal{R}}$. This method is implemented mostly in distributed architectures where mobile devices refrain from transmitting information by being silent during certain time periods. Privacy sensitive users, or privacy tools (e.g., [10, 44, 45, 47, 53]) make use of mainly this method (along with other method). It can also be implemented in the centralized architecture. A service provider who follows some privacy policies, in practice, is applying this method. The mechanisms proposed in [40, 42] are examples of using event-hiding privacy tools, especially, in centralized architecture.
- **Adding Dummy Events** The other method to achieve privacy is to mislead an observer by adding some *dummy* events through the event injection method of the transformation function. This method can be effectively implemented in the centralized architecture. Mechanisms proposed in [19, 49, 51, 55, 71] employ mainly this method. Generating a trace of events that looks like a normal user's trajectory is one of the main challenges of this set of papers.

- **Obfuscation** Using this method, the location-stamp and/or the time-stamp of the actual events in $\mathcal{R}$ can be altered. Obfuscation methods result in *inaccuracy* or *imprecision* of the location/time of the events [26]. This is done by adding noise to the location- and/or time-stamp of the events or by coarse graining them. The method can be implemented in both distributed and centralized (and hence the hybrid) architectures using various algorithms. In the existing privacy preserving mechanisms, obfuscation is achieved mostly through *perturbation* [38, 56] or *generalization* [7, 8, 18, 35, 38] algorithms.
- **Anonymization** Using the anonymization method, the identity of an event in $\mathcal{R}$ is altered in order to break the link between a user and his events or to make a user's events unlinkable to each other. To this end, in the transformation function, the real identity of a user on each event can be replaced by one of his valid pseudonyms. In the centralized architecture, this method is implemented mainly by replacing all the events' identities with a single group pseudonym (full anonymization by having no identity, i.e., null pseudonym [66]), e.g., Cornielius *et al.* [21]. In distributed architectures, users themselves change their pseudonyms from time to time. This pseudonym change is done usually in some predetermined places called *mix zones* (Beresford and Stajano [10]) where users remain silent when they are inside the mix zone and change their pseudonyms when they leave the zone. Thus, these mechanisms employ the hiding method as well, e.g., Jiang *et al.* [47]. The pseudonym change can also be done in a self-organized way by using group-signatures (e.g., Calandriello *et al.* [16]) or ring-signatures (e.g., Freudiger *et al.* [32]) as group pseudonyms. The mechanism proposed by Li *et al.* [53] also makes use of group pseudonyms, and users exchange their group pseudonyms with each other when they leave mix zones. Buttyan *et al.* [15] and Freudiger *et al.* [31, 30] proposed formal models to evaluate the effectiveness of static and dynamic mix zones, respectively, (users decide on-the-fly whether to change their pseudonyms or not) in mobile ad hoc networks.

We denote the observable trace of a user $u \in \mathcal{U}$, which is the output of the location privacy preserving mechanism applied on the user's actual trace, by $\widehat{\mathcal{R}}_u = \overrightarrow{\mathbf{trns}}(\mathcal{R}_u)$. Thus, $\widehat{\mathcal{R}} = \bigcup_u \widehat{\mathcal{R}}_u$. If two actual events e_i and e_j are associated with the same user (i.e., $e_i \sim e_j$) and $\hat{e}_i = \mathbf{trns}(e_i)$ and $\hat{e}_j = \mathbf{trns}(e_j)$, then we define a relation $\sim_o$ between $\hat{e}_i$ and $\hat{e}_j$. In other words, we define $\hat{e}_i \sim_o \hat{e}_j$ if $\hat{e}_i, \hat{e}_j \in \widehat{\mathcal{R}}_u$ for some $u \in \mathcal{U}$, which represents the linkability of observable events.

4 Threat Model

Depending on the characteristics of the mobile network and the services provided for users, location privacy of users can be threatened in different ways. The adversary can be an entity who eavesdrops on wireless communications between users, or she can be the operator who provides a location-based service for her subscribers, or even she can be one of the users in the network who participates with other users in running a protocol. In our framework, the adversary

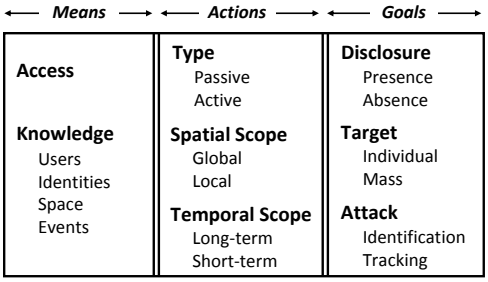

Fig. 2. Structure of the threat model and adversary's attributes

is actually the entity who observes the output of privacy preserving mechanisms and, hence, has access to a subset of the observable events $\widehat{\mathcal{R}}$. The subset of $\widehat{\mathcal{R}}$ that is accessible to a given adversary is called the set of *observed events* by that adversary and is denoted by $\widehat{\mathcal{R}}^{\mathrm{o}} \subseteq \widehat{\mathcal{R}}$. The properties of this subset and the implication of this observation on the users' location privacy depend on the characteristics of the adversary. Note that the adversary might have multiple observation points from each of which she can observe a different set of events. At each observation point, the adversary, observes different transformations of the same actual events. However, the structure of transformations (not their settings) is the same. Thus, in this section, we focus on the set of adversary's observed events at a single observation point, which is shown by $\widehat{\mathcal{R}}$.

In this framework, we model an adversary based on the following three factors: her *Means*, *Actions*, and *Goals*. Each of these factors are explained in the following sections and illustrated in Figure 2.

4.1 Means

The means of the adversary are the *technologies* available to her for capturing events, her *access* credentials in the system, and her a priori *knowledge* about the system.

Access — The adversary might eavesdrop on the wireless communication of users. Based on the level of sophistication of her eavesdropping devices, the accuracy of observed events changes. It can also be the entity in the organization that provides the service for users. For example, the adversary might be the insurance company who periodically collects positions of vehicles in a Pay-As-You-Drive application [68], or can be the operator of an automatic toll collection system who can get more sparse information about location of vehicles. Fingerprinting wireless devices [14, 22, 29] is also one technique that can be used in order to extract pseudonyms correlated to the hardware used by a user rather than based on the content of his messages (e.g., IP address). As another example, the adversary might have access to the high-level transactions of a location-based

service (LBS) or an environment monitoring network [39].

Knowledge — The a priori knowledge of the adversary is composed of multiple pieces. Here, we categorize the adversary's knowledge into multiple classes. The precision and confidence of the adversary's knowledge about each class determines her a priori knowledge. Her knowledge in each class can be deterministic or probabilistic and this must be clarified in each threat model.

- **Users** The adversary might know the (exact or estimated) number of users at any time, or more precisely the set of users $\mathcal{U}$, that implies knowing the real identity of active users. This knowledge can evolve over time, or she may remain oblivious about the dynamics of the set of users and their joining/leaving. This class of adversary's knowledge also includes the adversary's knowledge about the relation between users, i.e., social network graph.
- **Identities** This class specifies to what extent the adversary knows about the users' identities and the pseudonyms used by them. The adversary might know the relation between pseudonyms of a user, and also the constraints on the set of pseudonyms (e.g., how many pseudonyms a user can have). The extent to which each pseudonym is linkable to its holder's real-identity is also part of the knowledge of the adversary in this class.
- **Space** The knowledge of the adversary on the three-level model for the space in which users move falls into this class. The connection between users and places must also be specified here. For example, does the adversary know the address of the users' homes or their workplaces, which are modeled in the second layer of our space model?
- **Events** The adversary might have access to some actual events that are performed before the observation time. Moreover, in many cases the adversary has some statistics about the typical behavior of users. For example, she knows the (im)possibility or the probability that one specific actual event can be performed by a user, or that two specific events belong to the same user. Knowledge of the adversary about mobility profile of users (which represents how probable/possible it is for a specific user or a mass of users to move from one location to another location in a specific time period) falls into this class.

We assume that the adversary knows the application, employed privacy tools, and also the location-privacy metric that users (or system designers) use.

4.2 Actions

The adversary might be *passive* and only observe the network, or, in addition to that, become *active* and interact with the users in order to obtain more information about them. In most of the cases in the literature, the adversary is only a passive observer. However, one can imagine some sophisticated active attacks, such as the following ones. By relaying [60] the traffic between two different mix zones [10], the adversary can create a fake high-density area and encourage neighboring users to change their pseudonyms. Thus, users have an erroneous feeling

of high privacy while the adversary can easily distinguish between pseudonyms in each mix zone and link the users' pseudonyms. The adversary can also create fake toll readers in vehicular networks by using a relay attack and force users to reveal their presence in critical locations. Here, the adversary actively participates in asking users to generate some events.

The action scope of an adversary is determined by the size of the location areas and the duration of time periods in which the adversary observes the system. Considering these factors, adversaries can consequently be divided into different categories. An adversary is *global* if she observes the observable events occurred at any location in the space. Whereas, she is called *local* if during the observation period she cannot observe the transformation of some events that are generated in specific location areas. Similarly, based on the observation time, an adversary is referred to as a *short-term* attacker if the transformation of events performed at some time periods are not observable by the adversary. In the case there is no such time restrictions, she is called a *long-term* attacker. In the case an attack is global and long-term, we have $\widehat{\mathcal{R}}^{\mathrm{o}} = \widehat{\mathcal{R}}$.

4.3 Goals

Presence vs. Absence Disclosure — An adversary's goals of observing users' activities in a mobile network can be divided into two main categories: *presence disclosure* or *absence disclosure*. In the former category, the adversary's goal is to find out if a given user or a set of users are present at some place(s). Whereas, in the latter category, the adversary wants to know whether a specific set of users are *not* present at some place(s). Virtually all of the attacks presented in the literature fall into the first category. However, there are some reports about the consequences of absence disclosure attacks on people. As an obvious example, by misusing her access to an LBS database, the adversary can find out the best time to break into a person's house, or blackmail them. An implementation of this attack can be found here [6]. Obfuscating the victim's location as big as the whole North America cannot protect him from absence disclosure attacks if he lives in Europe, despite the fact that the adversary cannot locate his exact location. However, using that obfuscation, if the user lives in the US, his location privacy is protected against both presence and absence disclosure attacks.

Individual vs. Mass Target — The attacks can disclose the private information of a specific user in an *individual target* attack, or it might be targeting a set of users, collectively, in a *mass target* attack where the adversary does not distinguish users in the set, for example when they belong to a community.

Tracking vs. Identification — The two main known attacks on users' location privacy, which are used usually to disclose users' presence, are *tracking* and *identification*. These two attacks are tightly related to each other, although they have different ways of obtaining users' private location information.

In *tracking* attacks, the adversary's goal is to reconstruct the users' actual trajectories (which might have been distorted by privacy preserving mechanisms) and subsequently identify the locations that the users have visited. This information can be used to predict the future locations of users. The tracking can

be done in various manners depending on the adversary's goal. The adversary might want to know the trace of location instances (i.e., coordinates) visited by the users in a given time period, or the location sites (e.g., specific hospital) where they have been to, or only the type of places that the users are used to periodically visit.

In *identification* attacks, the adversary wants to discover the real identity of her targets. This can be done on small scale where the adversary is interested in de-anonymizing a specific observed event, or on a large scale where the adversary is interested in finding the identity of users from whom the adversary has observed some anonymous traces of events. The identification is done using some inference attacks based on the adversary's knowledge on the linkability of the users to sensitive areas such as their homes or work places [37, 41, 50]. Identification can also leverage on the mobility pattern of users, because users tend to visit certain places regularly [23]. It can also be done indirectly through de-anonymizing [58] the social network that is linked (e.g., friend-finder applications) to the users' observed events.

It is clear that the success of each of the two above-mentioned attacks also paves the way for the other. In the case the adversary manages to discover the actual trajectory of a user, the identification of the user is not a difficult task. Especially if the adversary has access to the information about location sites such as homes or work places of the users, which contain a lot of information about their identities. In the case that the adversary has already de-anonymized some events of a user, the recovery of the user's actual trajectory (i.e., tracking him) can be done more easily, if the adversary has access to the mobility profile of the users.

Tracking and *identification* attacks have been individually studied in the literature. However, there is not much research done on *modeling* the inter-relation between them, which the adversary is likely to make use of (a combined attack). Considering that many privacy preserving mechanisms focus on protecting users from one of these two attacks, it is of upmost importance to analyze to what extent the adversary can break the privacy of users by running unexpected attacks, e.g., de-anonymizing the traces using the social network graph available to the adversary. Lastly, some inference attacks can be developed in order to recognize the activities of the targets and track the types of places they visit (which may eventually leads to their identification) [54].

5 Location-Privacy Measurement

Every location-privacy preserving mechanism is designed based on an assumed location-privacy metric. In the literature, various metrics have been used to capture location privacy in different scenarios. Needless to say, choosing the right metric for each specific setting is of upmost importance to increase the actual users' location-privacy against possible attacks. Considering different users' privacy needs, we define location privacy on macro and micro levels. We define the notion of *macroscopic* location privacy to be a user's privacy level throughout

his trajectory, whereas the *microscopic* location privacy is defined to represent a user's privacy on a small scale, for example at the time an event related to a given user is observed. These two metrics are tightly related to each other, and reflect location privacy in two different scales. The selection of the right scale highly depends on the threat model and on the specification of location-privacy requirements. Using these notions, we classify the metrics that have been used so far, and thus we can discuss their effectiveness in representing the true privacy level of users.

Microscopic Location-Privacy We expect the location privacy of a user at the micro level, depending on the user's privacy requirements, to be inversely proportional to the success of the adversary in identifying his real identity when an event is observed of a user, or to locate the user at a given time instance and find out his presence/absence at a given location. The more accurately the adversary can locate a user, the poorer the location privacy of the user will be. Depending on the threat model, the adversary might be interested to find out the coordinate of the user's location, his location site, or the type of his location.

The most popular metric for this micro location privacy is based on the *uncertainty* of the adversary. This metric was originally proposed by Diaz *et al.* [25] and Serjantov and Danezis [65] for anonymous networks (known as entropy-based or information theoretic-based metrics), and by Samarati and Sweeney [63, 67] for database privacy (known as k-anonymity metric, where, assuming maximum uncertainty for the adversary, k is equal to the effective anonymity set size [65]). These metrics were adapted to measure microscopic location privacy by Gruteser and Grunwald [38], and later used in many papers such as [7, 12, 34, 35, 48, 57, 64, 70, 72]. Virtually all of the various versions of uncertainty-based metrics for micro location-privacy, measure *only* the adversary's success in the presence-disclosure identification attacks. The metric reflects a given user's privacy, from whom an event is observed by the adversary, as the size of the effective anonymity set in which the user is hidden.

Macroscopic Location-Privacy How accurately the adversary can track a user throughout his trajectory (i.e., tracking attack), or how closely she can find out the identity of a user after observing a set of events (i.e., identification attack), is reflected by macroscopic location privacy metrics. The set of macro metrics can be divided into two main categories, based on the set of the criteria that is used in each metric: *uncertainty*-based and *error*-based metrics.

— The first set of macro metrics reflects the *uncertainty* of the adversary in tracking users. Similarly to the micro metrics, in this category, entropy-based metrics and k-anonymity are the two most popular measures. The set of observed events are linked to each other, based on the a priori knowledge of the adversary, in a way that each link shows the possibility and also probability of observing the two linked events from the same user. The entropy-based metrics exploit this data structure and compute the adversary's uncertainty at the outgoing links of each observed event. These values reflect the user's privacy throughout

his trajectory, however, only at the time instances that an event is observed of the user. The overall privacy level is measured mainly as the fraction of times when the uncertainty is below a threshold [9, 10, 30, 40, 42, 44, 45, 47, 53]. The k-anonymity metric is extended also to the macro level by measuring how many trajectories at a time period are indistinguishable in an anonymity set [12, 36, 59, 69].

— The second set of macro metrics are based on the adversary's *error* in tracking/identifying users. This category is divided into multiple subclasses: (i) *clustering-error*, (ii) *probability of error*, and (iii) *distortion-based* metrics.

(i) *Clustering-error* metrics: The adversary's goal is defined to be the clustering of the observed events into partitions, based on the relation $\sim_o$, each partition for one user. Formally, the adversary is looking for $\widehat{\mathcal{R}}_u$ for all $u \in \mathcal{R}$. Two slightly different versions of this metric are used in [28, 39]. Note that both versions aim at measuring the success of adversary's tracking attack.

(ii) *Probability of error* metrics: The adversary's probability of error in finding the real identity of a user, or linking his observed events, is considered as the metric. For identification attacks, in [50] various algorithms using machine learning techniques are proposed to identify the *homes* of mobile users in the users' observed events and subsequently find their identities based on the adversary's knowledge. Similarly, in [41] an algorithm is proposed to identify users based on their home addresses. The higher the average adversary's probability of error is, the higher the users' location privacy is in their model. In [33] the probability of error is used as the metric to evaluate users' location privacy against tracking attacks in mix zones.

(iii) *Distortion-based* metrics: Having prior knowledge about the system, and after observing a set of events, the goal of the adversary of perpetrating the tracking attack is to reconstruct the actual trajectory of the users. The distortion-based metric [66] reflects how distorted the reconstructed trajectory of each user will be for the adversary. To measure the distortion, it is enough to condition the possible actual trajectories of the targets to the observed events, and compute the expected distance of the predicted location of a user with his actual location at any time instance. It is shown that this metric is superior to the other macro metrics that focus on tracking attacks, in terms of the accuracy of the metric. A set of criteria, derived from the definition of location privacy, is also proposed to compare the effectiveness of the metrics.

Here, we discuss the next steps towards the definition of more accurate and realistic location-privacy metrics. The departure point is the threat model. Concerning presence disclosure attacks, we need metrics that represent users' location privacy in the presence of identification/tracking attacks. In identification attacks, *uncertainty*-based metrics (when we are measuring the system level privacy) and metrics based on the adversary's *probability of error* (when we measure user level privacy) are more representative of users' location privacy. Regarding tracking attacks, *distortion*-based metrics are shown [66] to be more accurate than other metrics, as the adversary's goal is to predict as closely as possible the

users' actual locations over time, and the more distorted her prediction is the higher users' location privacy is.

6 Application Scenario: Location-based Services (LBSs)

Using the proposed framework and terminology, we briefly model location-privacy in a typical LBS, as a small example to show the effectiveness of the framework.

Consider a LBS in a region, by means of which users can obtain the list of their nearby public places, by sending their GPS coordinates to the server. Users might either subscribe to the server in order to rate the places they visit, or simply use the service without authenticating to the server. As users decide when to use the system and this is done infrequently, the application is *manual* and *discrete*. In this kind of LBS, users do not need to form any social network on the server to use the service. Thus, their relation is not disclosed to the service provider through this service.

The three entities (users, applications, and privacy tools) employ location-privacy preserving methods in the following way. Users connect to the server with *pseudonyms* and their real identities are hidden. A pseudonym is the concatenation of IP address, cookie id, and username in the application. Hence, the first method for *users* in **trns** function is *anonymization*. The more basic method used by users is *hiding events*, as they do not connect to the server at every time instance and are hidden most of the times. The application on the mobile devices that connects to the LBS, also unintentionally uses *obfuscation* method by perturbing the actual events' location-stamps (due to the error of GPS devices). The third entity that is active in **trns** is the set of *privacy tools* that can be implemented in both distributed or centralized architectures, by using all four methods in **trns**. However, compared to the distributed architecture, the centralized form is more powerful but less practical. Especially, the method of *adding dummy events* can highly increase users' location-privacy in centralized architecture, where the privacy tool can fully anonymize users (by using the same group pseudonym for all users).

We assume the LBS operator aims at identifying users, and hence she is the adversary in our threat model. This explains the *technology* and *access* credentials of the adversary. Her *knowledge* can be modeled as following. She knows a subset of $\mathcal{U}$ and their real identities in $\mathcal{I}$. This is because only a fraction of the users in a region are known to the adversary. Moreover, she does not certainly know $|\mathcal{U}|$. However, she can estimate how many different users actively use the system at a specific time. She also has access to the home and work addresses of some users in $\mathcal{U}$, for example those have made this information available online, or by accessing a governmental database that stores this information. She is well aware of the space in which users move (GIS information modeled as the 3-layer structure in our framework), because it is part of the provided service. She has some statistics (with some precision) about users' typical mobility in the region. Any part of her knowledge is subject to error, as the adversary has some level of

uncertainty in them. This must be modeled and quantified in the computational threat model. These collectively model the adversary's *means*.

In terms of the adversary's *actions*, in our model, she can be classified as a *passive*, *global*, and *long-term* observer. Let us assume her *goal* is to perpetrate *presence disclosure* attack on *individual* targets. Then, she is able to execute both *identification* and *tracking* attacks. Here, we explain that these two attacks reinforce each other and have a tight dependency to each other in this scenario. She first *tracks* users (while they are pseudonymous) and clusters their observed events, by using her knowledge about users' pseudonyms and mobility patterns. Then, she tries to *identify* each pseudonymous user, as she has access to the information about their sensitive places (i.e., home and work). After a user is identified, his observed events are de-anonymized and thus he can be tracked more accurately. This information might even help the adversary to find out the users' locations from which they did not connect to the server. Thus, there is a strong dependency between *identification* and *tracking* attacks in this application.

The notion of location privacy that users are more likely to be concerned about is *macroscopic* location-privacy, as both of the above-mentioned attacks work at a large scale. The location-privacy metrics must capture the adversary's success in both attacks, considering their dependency. From our model of location privacy in LBSs, as discussed above, there are some open problems yet to be addressed in location privacy of LBSs: Modeling the adversary's a priori knowledge and incorporating her uncertainty in her knowledge; Modeling the interrelation of the two attacks; Capturing ultimate success of the adversary by a metric; Evaluating the users' location-privacy without employing privacy tools and depending only on the mechanisms that are (un)intentionally used by *users* and the *application* on his mobile device; and Measuring the impact of the method *adding dummy events*, especially in centralized architecture.

Other applications can similarly be modeled within our framework and interdependency between various components of location privacy can be identified. As it is shown, this results in finding the drawbacks of existing approaches and suggestions for improving them.

7 Related Work

In this section, we briefly survey the papers that formalize location privacy or give an overview of location-privacy problem.

Hong and Landay [43] introduce a basic toolkit, called Confab, for developing privacy-sensitive ubiquitous computing applications. The requirements of end-users (e.g., decentralized control, special exceptions for emergencies, and plausible deniability) and also application developers needs (e.g., support for optimistic and pessimistic applications, access control mechanisms, and logging) are considered.

As an important technique to protect users' location privacy in LBSs, Duckham and Kulik [26] propose a formal model for obfuscation mechanisms. The

authors provide an algorithm to balance each user's desired quality of service against their need for location-privacy.

Bettini *et al.* [11] model the microscopic location-privacy of users in location-based services. The authors take a few different kinds of knowledge the adversary could acquire, and evaluate users' privacy using uncertainty-based metrics.

Decker [24] gives an overview of location-privacy problems in LBSs and divides them into two main classes: direct and indirect attacks. Furthermore, the technical approaches to prevent misuse of location data are classified in the following categories: policy approaches, anonymization, and deliberate impairment of locating. The role of legal regulations in protecting users' location privacy is also discussed.

Blumberg and Eckersley [13] present a list of emerging threats and opportunities of location-aware services that create digital repositories of people's movement and activities. As a way to protect people's location privacy in the short run, the authors refer to "using cryptographic tools" for building systems that blindly provide location-based services and cannot infer information about people's location. The authors believe that, in the long run, "the decision about when we retain our location privacy (and the limited circumstances under which we will surrender it) should be set by democratic action and lawmaking."

Krumm [52] provides a literature review of computational location privacy. The authors discuss the need for sharing location information and also the value that people put on preserving location privacy. Going through a list of threats and countermeasures, the author, state that the progress in computational location privacy is dependent to the accuracy of location privacy metrics.

Shokri *et al.* [66] propose a framework for modeling and evaluating macroscopic location-privacy metrics. Within this framework, they formalize various metrics and, based on a set of criteria derived from the definition of location privacy, the authors study the effectiveness of existing metrics in reflecting the actual users' location privacy. Finally, they propose a distortion-based metric and show that it is superior to other existing macro metrics.

As discussed, all these works focus on formalizing a specific problem of location privacy, e.g., particular protection mechanisms, and therefore do not provide a generic framework that encompasses all location-privacy components.

8 Conclusion

In this paper, we propose a framework for location privacy that unifies its relevant components, considering users' actual location-privacy requirements. We identify different categories of threats, and establish a methodology for measuring location privacy in different scenarios in order to identify appropriate location-privacy metrics. The proposed framework enables us to design and build appropriate location-privacy protection mechanisms, identify the drawbacks of existing works, express different works with the same terminology, and discover new directions for research in location privacy.

Acknowledgment

The authors would like to thank Marco Gruteser, Claudia Diaz, and Carmela Troncoso for their insights and suggestions on earlier versions of this work.

References

1. http://en.wikipedia.org/wiki/Bluedating.
2. http://www.aka-aki.com.
3. http://csg.ethz.ch/research/projects/Blue_star.
4. http://www.madeinlocal.com.
5. http://www.loopt.com.
6. http://www.pleaserobme.com.
7. C. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati. Location privacy protection through obfuscation-based techniques. In *Data and Applications Security XXI*, 2008.
8. B. Bamba, L. Liu, P. Pesti, and T. Wang. Supporting anonymous location queries in mobile environments with privacygrid. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 237–246, New York, NY, USA, 2008. ACM.
9. A. R. Beresford. *Location privacy in ubiquitous computing.* PhD thesis, University of Cambridge Computer Laboratory, 2005.
10. A. R. Beresford and F. Stajano. Mix zones: User privacy in location-aware services. In *PERCOMW '04: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*, page 127, Washington, DC, USA, 2004. IEEE Computer Society.
11. C. Bettini, S. Mascetti, X. S. Wang, and S. Jajodia. Anonymity in location-based services: Towards a general framework. In *MDM '07: Proceedings of the 2007 International Conference on Mobile Data Management*, pages 69–76, Washington, DC, USA, 2007. IEEE Computer Society.
12. C. Bettini, X. S. Wang, and S. Jajodia. Protecting privacy against location-based personal identification. In *In 2nd VLDB Workshop SDM*, pages 185–199, 2005.
13. A. J. Blumberg and P. Eckersley. On locational privacy, and how to avoid losing it forever. Technical report, Electronic Frontier Foundation (EFF), 2009.
14. V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *MobiCom '08: Proceedings of the 14th ACM international conference on Mobile computing and networking*, pages 116–127, New York, NY, USA, 2008. ACM.
15. L. Buttyán, T. Holczer, and I. Vajda. On the effectiveness of changing pseudonyms to provide location privacy in vanets. In *ESAS*, pages 129–141, 2007.
16. G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy. Efficient and robust pseudonymous authentication in vanet. In *VANET '07: Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks*, pages 19–28, New York, NY, USA, 2007. ACM.
17. A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric urban sensing. In *ACM WICON*, 2006.
18. C.-Y. Chow, M. F. Mokbel, and X. Liu. A peer-to-peer spatial cloaking algorithm for anonymous location-based service. In *GIS '06: Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems*, pages 171–178, New York, NY, USA, 2006. ACM.

19. R. Chow and P. Golle. Faking contextual data for fun, profit, and privacy. In *WPES '09: Proceedings of the 8th ACM workshop on Privacy in the electronic society*, pages 105–108, New York, NY, USA, 2009. ACM.
20. C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. Anonysense: privacy-aware people-centric sensing. In *ACM MobiSys*, 2008.
21. C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. Anonysense: privacy-aware people-centric sensing. In *MobiSys '08: Proceeding of the 6th international conference on Mobile systems, applications, and services*, pages 211–224, New York, NY, USA, 2008. ACM.
22. B. Danev and S. Capkun. Transient-based identification of wireless sensor nodes. In *IPSN '09: Proceedings of the 2009 International Conference on Information Processing in Sensor Networks*, pages 25–36, Washington, DC, USA, 2009. IEEE Computer Society.
23. Y. De Mulder, G. Danezis, L. Batina, and B. Preneel. Identification via location-profiling in gsm networks. In *WPES '08: Proceedings of the 7th ACM workshop on Privacy in the electronic society*, pages 23–32, New York, NY, USA, 2008. ACM.
24. M. Decker. Location privacy-an overview. In *ICMB '08: Proceedings of the 2008 7th International Conference on Mobile Business*, pages 221–230, Washington, DC, USA, 2008. IEEE Computer Society.
25. C. Diaz, S. Seys, J. Claessens, and B. Preneel. Towards measuring anonymity. In R. Dingledine and P. Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.
26. M. Duckham and L. Kulik. A formal model of obfuscation and negotiation for location privacy. In *Proceedings of Pervasive Computing, Third International Conference, PERVASIVE*, Munich, Germany, 2005.
27. M. Duckham and L. Kulik. Location privacy and location-aware computing. In *Dynamic and Mobile GIS: Investigating Change in Space and Time*, 2006.
28. L. Fischer, S. Katzenbeisser, and C. Eckert. Measuring unlinkability revisited. In *WPES '08: Proceedings of the 7th ACM workshop on Privacy in the electronic society*, pages 105–110, New York, NY, USA, 2008. ACM.
29. J. Franklin, D. McCoy, P. Tabriz, V. Neagoe, J. Van Randwyk, and D. Sicker. Passive data link layer 802.11 wireless device driver fingerprinting. In *USENIX-SS'06: Proceedings of the 15th conference on USENIX Security Symposium*, Berkeley, CA, USA, 2006. USENIX Association.
30. J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. C. Parkes. On non-cooperative location privacy: a game-theoretic analysis. In *CCS '09: Proceedings of the 16th ACM conference on Computer and communications security*, pages 324–337, New York, NY, USA, 2009. ACM.
31. J. Freudiger*, M. H. Manshaei*, J.-Y. Le Boudec, and J.-P. Hubaux. On the Age of Pseudonyms in Mobile Ad Hoc Networks. In *IEEE Infocom*, 2010.
32. J. Freudiger, M. Raya, and J.-P. Hubaux. Self-Organized Anonymous Authentication in Mobile Ad Hoc Networks. In *Conference on Security and Privacy in Communication Networks (Securecomm)*, pages 350–372, 2009.
33. J. Freudiger, R. Shokri, and J.-P. Hubaux. On the optimal placement of mix zones. In *PETS '09: Proceedings of the 9th International Symposium on Privacy Enhancing Technologies*, pages 216–234, Berlin, Heidelberg, 2009. Springer-Verlag.
34. B. Gedik and L. Liu. Location privacy in mobile systems: A personalized anonymization model. In *ICDCS '05: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems*, pages 620–629, Washington, DC, USA, 2005. IEEE Computer Society.

35. B. Gedik and L. Liu. Protecting location privacy with personalized k-anonymity: Architecture and algorithms. *IEEE Transactions on Mobile Computing*, 7(1):1–18, 2008.
36. A. Gkoulalas-Divanis and V. S. Verykios. A free terrain model for trajectory k—anonymity. In *DEXA '08: Proceedings of the 19th international conference on Database and Expert Systems Applications*, pages 49–56, Berlin, Heidelberg, 2008. Springer-Verlag.
37. P. Golle and K. Partridge. On the anonymity of home/work location pairs. In *Pervasive '09: Proceedings of the 7th International Conference on Pervasive Computing*, pages 390–397, Berlin, Heidelberg, 2009. Springer-Verlag.
38. M. Gruteser and D. Grunwald. Anonymous usage of location-based services through spatial and temporal cloaking. In *MobiSys '03: Proceedings of the 1st international conference on Mobile systems, applications and services*, pages 31–42, New York, NY, USA, 2003. ACM.
39. B. Hoh and M. Gruteser. Protecting location privacy through path confusion. In *SECURECOMM '05: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks*, pages 194–205, Washington, DC, USA, 2005. IEEE Computer Society.
40. B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J.-C. Herrera, A. M. Bayen, M. Annavaram, and Q. Jacobson. Virtual trip lines for distributed privacy-preserving traffic monitoring. In *MobiSys '08: Proceeding of the 6th international conference on Mobile systems, applications, and services*, pages 15–28, New York, NY, USA, 2008. ACM.
41. B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Enhancing security and privacy in traffic-monitoring systems. *IEEE Pervasive Computing*, 5(4):38–46, 2006.
42. B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Preserving privacy in gps traces via uncertainty-aware path cloaking. In *CCS '07: Proceedings of the 14th ACM conference on Computer and communications security*, pages 161–171, New York, NY, USA, 2007. ACM.
43. J. I. Hong and J. A. Landay. An architecture for privacy-sensitive ubiquitous computing. In *MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 177–189, New York, NY, USA, 2004. ACM.
44. L. Huang, H. Yamane, K. Matsuura, and K. Sezaki. Towards modeling wireless location privacy. In *Proceedings of Privacy Enhancing Technologies Workshop (PET 2005)*, pages 59–77, 2005.
45. L. Huang, H. Yamane, K. Matsuura, and K. Sezaki. Silent cascade: Enhancing location privacy without communication qos degradation. In *Security of Pervasive Computing (SPC)*, pages 165–180, 2006.
46. B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden. Cartel: a distributed mobile sensor computing system. In *ACM SenSys*, 2006.
47. T. Jiang, H. J. Wang, and Y.-C. Hu. Preserving location privacy in wireless lans. In *MobiSys '07: Proceedings of the 5th international conference on Mobile systems, applications and services*, pages 246–257, New York, NY, USA, 2007. ACM.
48. P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias. Preventing location-based identity inference in anonymous spatial queries. *Knowledge and Data Engineering, IEEE Transactions on*, 19(12):1719–1733, Dec. 2007.
49. H. Kido, Y. Yanagisawa, and T. Satoh. An anonymous communication technique using dummies for location-based services. In *Pervasive Services, 2005. ICPS '05. Proceedings. International Conference on*, pages 88–97, July 2005.

50. J. Krumm. Inference attacks on location tracks. In *In Proceedings of the Fifth International Conference on Pervasive Computing (Pervasive), volume 4480 of LNCS*, pages 127–143. Springer-Verlag, 2007.
51. J. Krumm. Realistic driving trips for location privacy. In *Pervasive '09: Proceedings of the 7th International Conference on Pervasive Computing*, pages 25–41, Berlin, Heidelberg, 2009. Springer-Verlag.
52. J. Krumm. A survey of computational location privacy. *Personal Ubiquitous Comput.*, 13(6):391–399, 2009.
53. M. Li, K. Sampigethaya, L. Huang, and R. Poovendran. Swing & swap: user-centric approaches towards maximizing location privacy. In *WPES '06: Proceedings of the 5th ACM workshop on Privacy in electronic society*, pages 19–28, New York, NY, USA, 2006. ACM.
54. L. Liao, D. Fox, and H. Kautz. Location-based activity recognition using relational markov networks. In *IJCAI'05: Proceedings of the 19th international joint conference on Artificial intelligence*, pages 773–778, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.
55. H. Lu, C. S. Jensen, and M. L. Yiu. Pad: privacy-area aware, dummy-based location privacy in mobile services. In *MobiDE '08: Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, pages 16–23, New York, NY, USA, 2008. ACM.
56. J. Meyerowitz and R. Roy Choudhury. Hiding stars with fireworks: location privacy through camouflage. In *MobiCom '09: Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 345–356, New York, NY, USA, 2009. ACM.
57. M. F. Mokbel, C.-Y. Chow, and W. G. Aref. The new casper: query processing for location services without compromising privacy. In *VLDB '06: Proceedings of the 32nd international conference on Very large data bases*, pages 763–774. VLDB Endowment, 2006.
58. A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *SP '09: Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, pages 173–187, Washington, DC, USA, 2009. IEEE Computer Society.
59. M. E. Nergiz, M. Atzori, and Y. Saygin. Towards trajectory anonymization: a generalization-based approach. In *SPRINGL '08: Proceedings of the SIGSPATIAL ACM GIS 2008 International Workshop on Security and Privacy in GIS and LBS*, pages 52–61, New York, NY, USA, 2008. ACM.
60. P. Papadimitratos, M. Poturalski, P. Schaller, P. L. and D. Basin, S. Čapkun, and J.-P. Hubaux. Secure Neighborhood Discovery: A Fundamental Element for Mobile Ad Hoc Networking. *IEEE Communications Magazine*, 46(2):132–139, February 2008.
61. A. Pfitzmann and M. Köhntopp. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management, v0.32. Accessible through http://dud.inf.tu-dresden.de/Anon_Terminology.shtml, 2009.
62. O. Riva and C. Borcea. The urbanet revolution: Sensor power to the people! *IEEE Pervasive Computing*, 6(2), 2007.
63. P. Samarati and L. Sweeney. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. In *IEEE Symposium Research in Security and Privacy*, 1998.
64. K. Sampigethaya, L. Huang, M. Li, R. Poovendran, K. Matsuura, , and K. Sezaki. Caravan: Providing location privacy for vanet. In *The 3rd workshop on Embedded Security in Cars (ESCAR)*, 2005.

65. A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In R. Dingledine and P. Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.
66. R. Shokri, J. Freudiger, M. Jadliwala, and J.-P. Hubaux. A distortion-based metric for location privacy. In *WPES '09: Proceedings of the 8th ACM workshop on Privacy in the electronic society*, pages 21–30, New York, NY, USA, 2009. ACM.
67. L. Sweeney. k-anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5), 2002.
68. C. Troncoso, G. Danezis, E. Kosta, and B. Preneel. Pripayd: privacy friendly pay-as-you-drive insurance. In *WPES '07: Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pages 99–107, New York, NY, USA, 2007. ACM.
69. T. Xu and Y. Cai. Exploring historical location data for anonymity preservation in location-based services. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 547–555, April 2008.
70. T. Xu and Y. Cai. Feeling-based location privacy protection for location-based services. In *CCS '09: Proceedings of the 16th ACM conference on Computer and communications security*, pages 348–357, New York, NY, USA, 2009. ACM.
71. T.-H. You, W.-C. Peng, and W.-C. Lee. Protecting moving trajectories with dummies. In *Mobile Data Management, 2007 International Conference on*, pages 278–282, May 2007.
72. G. Zhong and U. Hengartner. A distributed k-anonymity protocol for location privacy. *IEEE International Conference on Pervasive Computing and Communications*, 0:1–10, 2009.