Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2011, Article ID 486975, 12 pages
doi:10.1155/2011/486975

Research Article

Time in Privacy Preserving LBSs: An Overlooked Dimension

Luciana Marconi,[1] Roberto Di Pietro,[2] Bruno Crispo,[3] and Mauro Conti[4]

[1] *Dipartimento di Informatica, Sapienza Università di Roma 00185, Roma, Italy*
[2] *Dipartimento di Matematica, Università di Roma Tre, Roma, Italy*
[3] *Dipartimento di Ingegneria e Scienza dell'Informazione, Università di Trento 38122, Trento, Italy*
[4] *Department of Computer Science, Vrije Universiteit Amsterdam, 1081 Amsterdam HV, The Netherlands*

Correspondence should be addressed to Luciana Marconi, marconi@di.uniroma1.it

Received 16 October 2010; Revised 23 January 2011; Accepted 27 January 2011

Academic Editor: Cristina Pinotti

A new privacy model for Location-Based Services (LBSs) has been recently proposed based on users' footprints-these being a repre-sentation of the amount of time a user spends in a given area. Unfortunately, while the model is claimed to be independent from the specific knowledge of the adversary about users' footprints, we argue that an adversary, that has a more structured knowledge over time, can pose a threat to the privacy guarantees of the model. The major contribution of this paper is to show that time is a relevant dimension that needs to be taken into consideration when investigating LBSs privacy issues. In particular, we show that applying our considerations, user privacy can be violated. We support our claim with analysis and a concrete example. Furthermore, by analyzing a real data set of vehicular traces, we show that the threat is actually present in a real scenario and that its effect on jeopardizing user privacy is relevant.

1. Introduction

Location-Based Services (LBSs) can be defined as services that add value to a user integrating his mobile device's location with additional information. Hence, the localization feature can be considered the main characteristic of a Location-based service. LBSs can be regarded as a subset of context-aware applications [1], the most basic context being the user's location. Context information is used to deliver a service and to add value to the service by adapting it to the user's personal context.

LBSs are widely spreading, particularly leveraging the use of mobile devices. As an example, we can consider the vehicular services that many national transportation infrastructures are developing: traffic monitoring, hazard warning, congestion-based, and "pay-as-you-go" road pricing [2, 3]. However, this type of services are subject to a privacy threat: the possibility to identify the user that requests a given service and her location at the time of the request. Even when privacy mechanisms are taken into consideration to anonymize the users, a user might be reidentified correlating the access information with other kind of information (e.g., the mobility of the user or some specific location-bound feature). In particular, there are three main issues related to the privacy of users in LBSs (i) how to anonymize a user; (ii) how to specify the level of anonymity; and, (iii) how to guarantee to a given user the same level of desired anonymity for all of her requests. Common anonymization techniques leverage the concept of k-anonymity (i) consisting in cloaking the user within a set of k potential users. The *feeling*-based model, recently introduced [4, 5], also leverages the concept of k-anonymity. However, this model is motivated by the fact that specifying a practical value of k could be a difficult choice for the user. Hence, the feeling-based model allows a user to define her desired level of anonymity (ii) by specifying a given area (e.g., a shopping mall). The entropy of the selected area is used to describe its popularity. In turns, the popularity is expressed in terms of footprints of the visitors in the selected area. The popularity of the user-specified area is considered later on, in the subsequent user's LBSs requests, as the anonymization level that the LBS has to guarantee to the user (iii).

While the feeling-based approach seems to be promising from the point of view of user's awareness of privacy, we

argue that the specific proposed solution is missing an important variable: time. In fact, the threat model considered in the proposals [4, 5] assumes an adversary having the same amount of information on the users as the one leveraged by the anonymizer. While this might seem a strong adversary model, it actually does not take into consideration practical aspects related to the distribution of such a knowledge over time. In particular, we consider both of the following situations to be practical. First, the adversary might have the information of the users footprints structured over time (e.g. how many footprints in the *morning* and how many in the *afternoon*). Second, the adversary might just be able to observe a subset of the footprints (e.g., the adversary is only able to get footprints information during the *morning*). While in the first case the adversary is stronger than the one consider in [4, 5]—having more structure data—the second scenario describes a weaker (but more realistic) adversary—basing its decisions on depleted information. We further underline that both of these adversary models fall into the assumption given in [4, 5] about the adversary.

In this work, we highlight the importance of the time when providing privacy in Location-based Services. We first show how user privacy can be violated leveraging time, with respect to the solutions in [4, 5]. In particular, we investigate on the provided privacy considering a different, more realistic adversary model. We argue that the newly introduced adversary is realistic and that it can also be weaker in terms of the amount of users information available, but still effective. We introduce our claim through a practical example; we then support and verify the claim with simulations and analysis on a real data set of vehicles' traces. The rest of the paper is organized as follows. Section 2 describes the related work in the area. Section 3 defines the notion of time and presents the threat model and the feeling-based privacy model. Section 4 shows how user privacy can be violated applying our considerations; we support our claim with both analysis and a practical example. Section 5 discusses and compares results from the analysis of a real data set. Section 6 argues about a viable approach under assumptions sligthly different from the ones in [4, 5]. Finally, Section 7 reports some concluding remarks.

2. Related Work

One of the main issues that slow down the large-scale adoption of LBSs is privacy [6]. In particular, given the peculiarity of these services (e.g., particularly relevant to mobile user devices), the privacy solutions already designed for other environments—like the ones based on k-anonymity [7–9]—result not portable into this context.

The main aspect related to the anonymization of LBSs regards the users mobility. In fact, mobile users ask for LBSs from different locations that correspond either to their current position or other positions of their interest. The first approach [10] for location anonymity aimed at applying the k-anonymity concept. The proposal was to reduce the accuracy of the definition of the user location (defined by both space and time) when asking for an LBS. The aim of reducing this accuracy was to cloak the requesting user within $k - 1$ other users, present in a broader area and consider a broader time frame. However, increasing the area would lead to a coarser service, while increasing the time frame would lead to a delay of the user's request.

Several works leveraged on the basic concept introduced in [10]. For example, the CliqueCloak algorithm [11] aims at minimizing the size of the cloaking area, while allowing the user to specify the value of k. However, this solution is practicable only for small values of k and requires a high computation overhead. The work in [12] generates a cloaking area in polynomial time and also considers attacks that correlate periodic location updates. The possibility of choosing k is also considered in [13], without considering the minimization of the cloaking area. Further work [14] provides a solution for mobile peer-to-peer environment, where the cloaking area is determined in a distributed way. The spatial cloaking algorithm proposed in [15] distinguishes between location privacy (i.e., a user willing to hide her location) and query privacy (i.e., a user can have her location revealed, but not her query). The aim is to prevent the adversary to link the user location to the submitted query. The motivation stems from the fact that in many applications the locations of mobile users is publicly known.

All these works do not explicitly consider the fact that nodes move, and their location-related request might be correlated. This issue has been first addressed by some works [16, 17] intended to cope with nodes tracing. However, these solutions were not developed having LBSs privacy in mind. In fact, they all report the actual user location. In particular, the work in [16] introduced the concept of mix zone—a zone where nodes avoid reporting their locations and exchange their identification instead. The aim of a mix zone is to make it hard for an adversary to correlate the pseudonym that a node used before entering the mix zone, and its pseudonym once it is out of the mix zone. Selfish behaviour of the nodes in mix zones has also been considered recently [18], as well as how pseudonyms aging affects privacy [19]. An idea similar to the one of mix zone is *path confusion* [17, 20]—pseudonyms are exchanged between nodes that have paths close to each other. The mix zone concept is also applied in [21] to protect the location privacy of drivers in vehicular networks (Vanet). The idea is to combine mix zones with mix networks that leverage on the mobility of vehicles and the dynamics of road intersections to mix identifiers.

The solution proposed in [22] requires that each LBSs request comes together with at most $k - 1$ dummy requests that simulate the movement of nodes. However, the dummy traces do not take into consideration the actual geography of the area where the corresponding dummy user is expected to be—such type of anomalies could let the adversary identify the dummy requests. Trajectory anonymization is also considered in [23], increasing the cloaking area to include exactly $k - 1$ other users. Unfortunately, continuously increasing the cloaking area degrades the precision of the LBSs.

The special case of providing location privacy in Vanet has been addressed in [24]. In this work, authors observed how Vanet poses specific constraints to mobility of nodes (vehicles)—the movement being spatially restricted (to lanes and freeways) and dependent (among vehicles). To tackle these unique characteristics and they proposed a scheme

that leverages pseudonyms with some enhancing features (i) increasing silent period between subsequent broadcast messages to obtain pseudonyms' unlinkability; (ii) grouping vehicles in geographical proximity to avoid overhearing of pseudonyms. As for vehicle-to-infrastructure communication, a privacy preserving mechanism can be found in [25].

A slightly different problem, that is avoiding reporting information about sensitive areas (e.g., a night club), has also been addressed [26]. Here, anonymization is achieved using areas instead of users. In fact, the reported location should include k sensitive areas instead of k users. Similarly, the framework proposed in [27] provides obfuscation of sensitive semantic locations-based on the privacy preference specified by each user. The solution uses a probabilistic model of space—the semantic locations being expressed in terms of spatial features—and does not take time into account. The solution proposed in [28] aims to avoid reporting the user location. The technique applies a Private Information Retrieval protocol to let the user of the service to download directly the LBSs information without requiring a trusted anonymizer. However, as the amount of data to be downloaded by the user depends on the total amount of data stored by the service provider, it may be impractical for a mobile device.

A problem strictly related to the protection of the user location privacy is the quantification of the "privacy level" guaranteed by several solutions. The solution in [20] quantifies location privacy as the duration over which an attacker could track a subject. The expected error in distance between a person's current location and an attacker's uncertain estimate of that location is used in [17]. The number of users k represents the level of privacy in [10] where k-anonymity is introduced for location privacy. Other works derive metrics from information theory [29]. For instance, entropy is the privacy quantifier used in [5, 16]. Whatever location privacy metric is adopted, it is maximized if no one knows a subject's location. Hence, the majority of the proposed solutions can be considered a trade-off between location privacy and quality of service. Some interesting solutions to location privacy in WSNs (Wireless Sensor Networks), sharing some common points with LBSs, have already been proposed. In particular, solutions in [30, 31] achieve privacy when querying a WSN, but sensors are required to partake logical hierarchy. Open problems highlights and related solution guidelines for a general privacy model in WSNs are in [32].

The problem of anonymity of trajectories has also been considered in other contexts. For instance, the work in [33] proposes a privacy-aware data publishing perspective. Differently from the LBS context, where the anonymity is centered on the Location-based service, authors of [33] consider an off-line and data-centric anonymity on a database of moving objects. The anonymization is enforced before the database is made public—the aim being to preserve privacy of people releasing the data to the public.

In our work, we show that leveraging time-frame provides an adversary with a powerful tool to compromise privacy in LBSs. A preliminary investigation, without real data analysis and consequent discussion, appeared in [34]. In particular, we show an application of this concept by compromising the privacy claimed in [4, 5], where the feeling-based model is introduced. Being a reference also for this paper, we recall this model in Section 3.3. Finally, our findings are consistent with the recent proposal in [35] where time is considered one of the aspects to take into account to protect user location.

3. Preliminaries and Notation

In this section, we propose models and definitions used in the paper. Section 3.1 introduces the system model. Section 3.2 formalizes the notion of time applied to time-related concepts analyzed throughout this work. Section 3.3 gives an overview of the solutions proposed in [4, 5], while the threat model description can be found in Section 3.4.

3.1. System Model. We consider the same system architecture used in [4, 5]. We assume mobile nodes (users) communicating with location-based services (LBSs) providers through a central anonymity server, the location depersonalization server (LDS), which is considered trusted. The LDS is managed by some mobile service provider allowing the (mobile) users to access to wireless communications. The provider offers the depersonalization service as an added value service and supplies the LDS with an initial footprints database derived from users phone calls.

3.2. Formalizing Time. Consistently, with the literature [35], we consider a discrete timeline, starting from time t_0—this time corresponding to the deployment of the system. Hence, we formalize the notion of time with the following definitions.

Definition 1 (time unit). The smallest measurable time unit we consider in our discrete time-line.

Definition 2 (time period). A time period is a predetermined number (ℓ) of contiguous time units, $\ell \in \mathbb{N}^+$. We denote periods with p_i, $0 \leq i \leq \rho$, ρ being the number of periods from the system start-up.

Definition 3 (time slice). A time slice of a period p is defined to be a time interval of a predetermined length $s < \ell$. We denote time slice j of time period p with T_j^p.

Thus, a time period is composed of $q = \ell/s$ time slices. We assume, without loss of generality, that $q \in \mathbb{N}$.

Definition 4 (time frame). A time frame is defined to be the set obtained as the union of the ith time slice of each period. We denote a time frame with $\widehat{T}_i$. Hence, $\widehat{T}_j = \{T_j^{p_0}, T_j^{p_1}, \ldots, T_j^{p_\rho}\}$.

For a practical discussion, time parameters to be fixed are thus the length ℓ of the period and the length of the time slice s. As an example, if we fix ℓ to be one week, and s to be one day, the period p is set to be the pth week, T_1^p = Sunday, T_2^p = Monday, $\ldots$, T_7^p = Saturday represent the days of the pth week.

3.3. Feeling-Based Privacy Model. The aim of the work in [5] is to provide location privacy protection for users requesting location-based services enhancing the k-anonymity model. The privacy model proposed introduces the concept of feeling-based privacy, based on the intuition of privacy being mainly a matter of feeling. The user is allowed to express a privacy requirement by specifying a spatial region in which she would feel comfortably cloaked (public region). Their solution then transforms the intuitive notion of user privacy feeling, in a quantitative evaluation of the level of protection provided, using the user-specified region. They define the entropy of a spatial region to measure the popularity of that region. This popularity is then used as the quantity describing the user privacy requirement: the popularity of the location disclosed by the anonymizer on behalf of the user, is required to be at least that of the specified public region. Formally, they provide the following definitions.

Definition 5 (entropy). Let R be a spatial region and $S(R) = \{u_1, u_2, \ldots, u_m\}$ be the set of users having footprints in R. Let $n_i (1 \leq i \leq m)$ be the number of footprints that user u_i has in R, and $N = \sum_{i=1}^{m} n_i$. The entropy of R is defined as $E(R) = -\sum_{i=1}^{m} (n_i/N) \cdot \log(n_i/N)$.

Definition 6 (popularity). The popularity of R is defined as $P(R) = 2^{E(R)}$.

The entropy is used to address the problem of the possible dominant presence of some users in a certain region. This phenomenon makes the number of visitors of a region not sufficient to quantify its popularity. The property that $P(R)$ is higher if m is larger is preserved even using entropy: a region is more popular if it has more visitors. Also, a skewed distribution of footprints significantly reduces the $P(R)$ with respect to a symmetric distribution. The entropy is also intended by the authors as the amount of additional information needed for the adversary to identify the service user from $S(R)$ when R is reported as her location in requesting an LBSs.

3.4. Threat Model. In this section, we present the threat model we consider. In particular, we define two types of adversary: ADV and ADV^T, both satisfying the assumptions provided in [5]. In particular, ADV mimics the adversary considered in [4, 5]. ADV is able to identify users in a cloaking region correlating with restricted spaces. However, it will not be able to reidentify the user who requests the service. We assume the adversary being present from time t_0, that is from the system deployment. Hence, we observe that the adversary may coincide with the LBSs provider. In fact, it could be highly interested in exploiting the location knowledge (historical) of the LDS anonymizer—potentially motivated by commercial or marketing purposes. Thus, ADV and LBSs will be used interchangeably throughout the paper.

Some existing techniques use current location of k neighbours of the service requester to protect from the adversary and to calculate the cloaking area. These techniques protect the anonymity of the service users but not their location privacy. An adversary identifying the users in the cloaking area knows their locations as it is aware of their presence in the cloaking area at the time of the service request.

The idea to use footprints, that is historical data, makes the adversary weaker as it is not able to know neither who requested the service nor who was really there at the time of the service request. From this core idea, introduced in [4] and applied by the same authors to mobile user's trajectory can be extracted in [5], another implicit assumption: the indistinguishability for the ADV between current and historical visitors of the cloaking area. This is equivalent to assume that ADV can not have instantaneous access to current users location data. If this will be the case, the usage of historical locations would not be suitable to compute the cloaking box for depersonalization. As an example, let us suppose the LDS reporting a cloaking area for a user, based on a five footprints (historical) calculation. If the user is the only one actually in that area and the LBSs knows the user location at each time instant, the latter would immediately identify the service requester. Thus, we also assume the users location knowledge held by the adversary to be the footprints information provided by the LDS anonymizer. We denote such a knowledge with LK.

In this work, we also consider a time-aware adversary, ADV^T, that has just additional information on time frames. Hence, we assume ADV^T has the same knowledge of ADV (the footprints information database), with the difference that such a knowledge takes also into account the different time frames $\hat{T}_j$. We denote ADV^T knowledge with LK^T. We can observe that the knowledge of ADV^T might be lower than the knowledge of ADV as it could know footprints information regarding just a portion of the time slices. Figure 1 illustrates the comparison of the knowledge of the two adversaries. For example, Table *daily* stands for the footprints data information of ADV. Table *morning* and Table *afternoon* stand for the footprints data information in Table *daily*, split on two time frames. We assume that ADV^T may know both Table *morning* and Table *afternoon* or, in a weaker version, just one of the two.

Hence, two scenarios may apply to ADV^T: it has the same user footprints information of ADV split on time frames, or ADV^T has less user footprints information than ADV, having footprints information only for some time frames.

Table 1 summarizes the notation used in this work.

4. Time Warp: Facing The Time-Aware Adversary

In this section, we aim to investigate on the privacy guaranteed by the solution in [5] when facing ADV^T. Section 4.1 introduces the adversary model used and an example showing how user privacy can be violated. Section 4.2 provides an evaluation of the adversary effectiveness against the privacy guarantees of the protocol in [5].

4.1. The Time-Aware Adversary Model. Our adversary model is motivated by the fact that the privacy of user's location may be highly influenced by the time frames considered. For instance, we might refer to several real scenarios: a theatre is

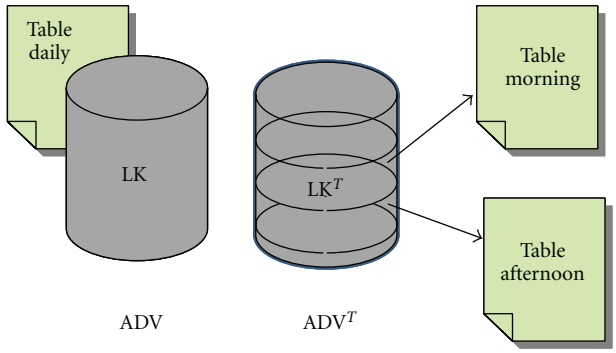


Figure 1: ADV and ADV^T footprints location knowledge.

Table 1: Notation table.

R	A spatial region
$S(R)$	Set of users having footprints in R
$E(R)$	Entropy of region R
$P(R)$	Popularity of region R
p_i	ith time period, $0 \le i \le \rho$
ρ	Number of periods from system start-up
T_i^p	ith time slice of a period p
q	l/s, number of slices composing a period
$\hat{T}_j$	Time frame $\hat{T}_j = \{T_j^{p_0}, T_j^{p_1}, \ldots, T_j^{p_\rho}\}$
$E(R, \hat{T}_j)$	Entropy of region R, during time slice x of time period p
$P(R, \hat{T}_j)$	Popularity of region R, for time frame $\hat{T}_j$
u_i	Generic ith user of a set of users, $1 \le i \le m$, $m \in \mathbb{N}$
$u_{i,\hat{T}_j}$	Generic ith user who have footprints in R in time frame $\hat{T}_j$
n_i	Number of footprints of user u_i in R
$n_{i,\hat{T}_j}$	Number of footprints of user u_i in R in time frame $\hat{T}_j$
N	Total number of footprints in a region R

a physical place where users concentrate only on particular days and in specific time frames, restaurants are most likely to be crowded at lunch and dinner time, and, office buildings are supposed to be almost empty during night. We show that with the knowledge held by ADV^T, the LDS is no more able to guarantee to users the claimed level of privacy. Further, we will also show scenarios where the entropy of the user public region is actually lower than the entropy calculated by the LDS. Therefore, the adversary will need less effort—with respect to what assumed by the LDS—to identify the user. Further, We will show that ADV^T may be more effective than ADV even if provided with less knowledge. This, as we will formally show at the end of this section, is due to the fact that time severely affects the entropy and the popularity of a cloaking region. This may result in a reduced amount of additional information needed for the adversary to identify the service user (see Section 3.3).

Definition 7 (entropy in $\hat{T}_j$). Let R be a spatial region and $S(R, \hat{T}_j)$ the set of users who have footprints in R, if observed during time frame $\hat{T}_j$, that is, $S(R, \hat{T}_j) = \{u_{1,\hat{T}_j}, u_{2,\hat{T}_j}, \ldots, u_{m,\hat{T}_j}\}$, where $n_{i,\hat{T}_j} (1 \le i \le m)$ is the number of footprints that user u_i has in R during the time frame $\hat{T}_j$ and $N_{\hat{T}_j} = \sum_{i=1}^{m} n_{i,\hat{T}_j}$. We define the entropy of R at time $\hat{T}_j$ as $E(R, \hat{T}_j) = -\sum_{i=1}^{m} (n_{i,\hat{T}_j}/N_{\hat{T}_j}) \cdot \log(n_{i,\hat{T}_j}/N_{\hat{T}_j})$.

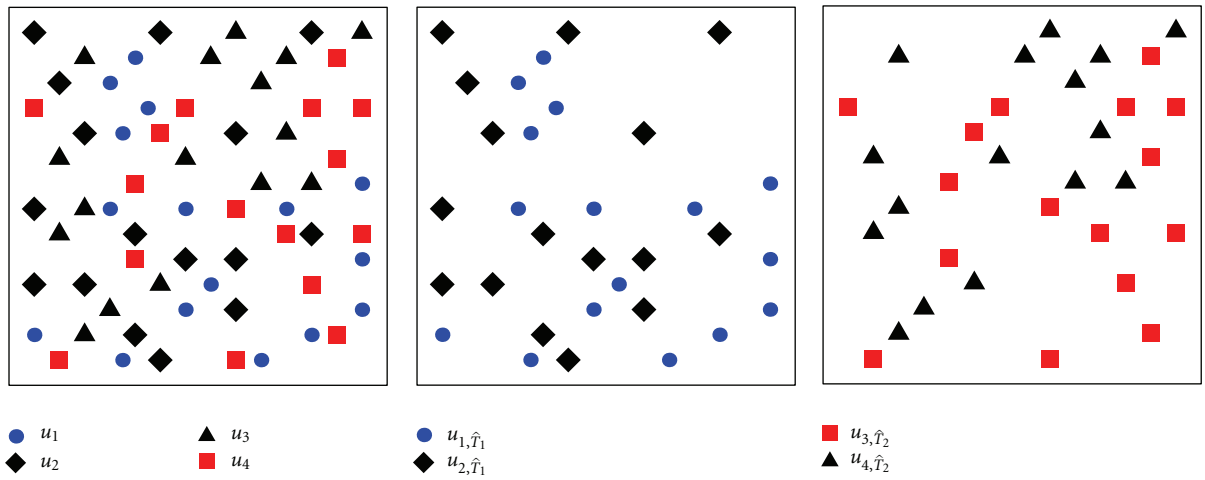
Definition 8 (popularity in $\hat{T}_j$). We define the popularity of R at time frame $\hat{T}_j$ as $P(R, \hat{T}_j) = 2^{E(R,\hat{T}_j)}$.

We observe that we can rewrite the quantities in Definition 5, using our Definition 7. More formally, we consider: $N = \sum_{x=1}^{q} N_{\hat{T}_j}$ and $n_i = \sum_{x=1}^{q} n_{i,\hat{T}_j}$.

We use the following example to support our discussions and to compare with the privacy model in [4, 5].

Example 1. Let us consider a user, Alice, requesting a LBS from her office building. She feels her privacy is preserved when specifying her office as the public region. In Alice's office, employees are organized on work shifts. Part of the employees are on a *morning* shift and the remaining ones on an *afternoon* shift. Let us consider $m = 4$ users (u_1, u_2, u_3, u_4) for the region corresponding to Alice's office (later on also referred as region R_1), each of them having 16 footprints in the LDS footprints database. This scenario is depicted in Figure 2. The corresponding footprints data for u_1, u_2, u_3, u_4 are provided and highlighted in the first column of Tables 2(a), 2(b), and 2(c), respectively.

Data in Table 2(a) represent the footprints information used by the LDS to calculate the entropy and the popularity of Alice's office. The results of the calculation determine a corresponding spatial region R_j (column labels in Table 2) used to cloak the user location. Hence, Table 2(a) also represents the knowledge of ADV. Tables 2(b) and 2(c) instead represent the structured knowledge of ADV^T, that is the same information of ADV when taking into account two time frames: $\hat{T}_1$ = *morning* and $\hat{T}_2$ = *afternoon*. Each table is provided with additional column data to show that both the entropy and the popularity depend on footprints distribution among visitors. In fact, it is possible to check that in each reported scenario the total number of footprints per user remains unchanged. Let us take the values of entropy and popularity in Table 2(a) as reference point to evaluate entropy and popularity calculations reported for each data column in Tables 2(b) and 2(c). As it is shown in Table 2(a) column 1, the maximum is obtained from footprints uniform distribution (column 1). We can observe that a more structured knowledge, like that of ADV^T in Tables 2(b) and 2(c) may result in the following possible scenarios. (i) ADV^T entropy and popularity values are strictly less than that of ADV. This is the case for the first and the second data columns in Table 2(c) and for the first column in Table 2(b), compared to the corresponding columns in Table 2(a). (ii) ADV^T entropy and popularity values are equal to that of ADV (see Tables 2(b) and 2(c)



(a) ADV: daily (b) ADVT: *morning* (c) ADVT: *afternoon*

FIGURE 2: ADV and ADV^T knowledge.

TABLE 2: ADV and ADV^T table data.

(a) ADV: *daily*

User	R_1	R_2
u_1	**16**	9
u_2	**16**	16
u_3	**16**	18
u_4	**16**	21
$E(R)$	2	1.94
$P(R)$	4	3.84

(b) ADV^T: *morning*

User	R_1	R_2	R_3
$u_{1,\hat{T}_1}$	**16**	4	8
$u_{2,\hat{T}_1}$	**16**	8	8
$u_{3,\hat{T}_1}$	**0**	9	8
$u_{4,\hat{T}_1}$	**0**	11	8
$E(R,\hat{T}_1)$	1	1.92	2
$P(R,\hat{T}_1)$	2	3.78	4

(c) ADV^T: *afternoon*

User	R_1	R_2	R_3
$u_{1,\hat{T}_2}$	**0**	5	8
$u_{2,\hat{T}_2}$	**0**	8	8
$u_{3,\hat{T}_2}$	**16**	9	8
$u_{4,\hat{T}_2}$	**16**	10	8
$E(R,\hat{T}_2)$	1	1.96	2
$P(R,\hat{T}_2)$	2	3.88	4

column 3). (iii) ADV^T entropy and popularity values are greater than that of ADV. This is the case for the second column in Table 2(b) with entropy 1.51—greater than the corresponding 1.49 in Table 2(a).

In the following, we formally prove that an anonymizer using the aggregated data can guarantee the level of privacy requested by the user only if it is facing the adversary ADV. In fact, we prove that when the anonymizer is facing ADV^T, the following two cases can also happen: (i) the anonymizer is not able to guarantee to the user the requested level of privacy. (ii) the anonymizer is degrading the accuracy of the location information for the LBSs, exceeding the level of privacy requested by the user.

Theorem 1. *Given a spatial region R and footprints data $\hat{T}_i$ related to the ith time slice, footprints distributions exist such that $E(R,\hat{T}_i) \neq E(R)$.*

Proof. The proof is a direct consequence of the two following cases.

Case 1. If $n_{i,\hat{T}_j}$ satisfies $n_{i,\hat{T}_j} \leq n_i \cdot N_{\hat{T}_j}/N$, then $E(R,\hat{T}_j) \leq E(R)$. In fact, the condition can be rewritten as: $(n_{i,\hat{T}_j}/N_{\hat{T}_j}) \leq (n_i/N)$. Since the log function is monotonically increasing, $\log n_{i,\hat{T}_j}/N_{\hat{T}_j} \leq \log n_i/N$. As a consequence, $E(R,\hat{T}_j) \leq E(R)$.

Case 2. If $n_{i,\hat{T}_j}$ satisfies $n_{i,\hat{T}_j} > n_i \cdot (N_{\hat{T}_j}/N)$, then $E(R,\hat{T}_j) > E(R)$. The proof is similar to the proof of Case 1. □

Case 1 shows that with a time-aware adversary, ADV^T, and the LDS is not always able to guarantee the level of privacy requested by the user. This happens when $E(R,\hat{T}_i) < E(R)$. In fact, if this is the case, the region R does not achieve an entropy at least equivalent to the public region specified by the user in order to meet her privacy requirement. Case 2 shows that with a time-aware adversary, ADV^T, the LDS is not always able to guarantee the maximum level of accuracy for the LBSs service requested by the user. This happens when $E(R,\hat{T}_i) > E(R)$. If this is the case, the LDS introduces a loss

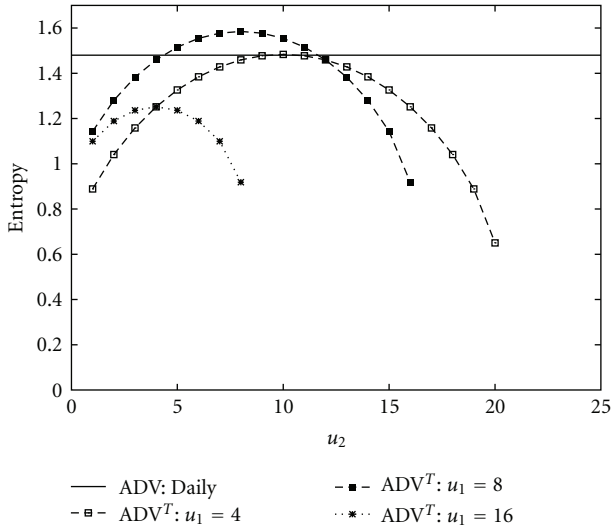


FIGURE 3: Comparing entropy between ADV and ADV^T: $\hat{T}_2$ (*afternoon*) footprints distribution, $u_{1,\hat{T}_2} = 4,\ 8,\ 16$.

in service accuracy—since a region larger than necessary is used to guarantee the user requested level of privacy.

4.2. Evaluating the Adversary Effectiveness. In this section, we highlight the importance of the time when providing LBSs privacy.

To show the influence of the time frames, we evaluated the adversary effectiveness against the privacy guarantees of the protocol in [5]. To do so, we plot the analytical results of some example data. The aim of the graph is to show how footprints distribution impacts the entropy values used to measure the required adversary effort. We remind that the entropy is a measure for the adversary effort needed to compromise the user privacy. Let us assume the user selected a desired level of privacy (entropy). On the one hand, if the anonymizer behaves in such a way that the effort required to ADV^T to compromise privacy is less than the expected one, the anonymizer is failing in guaranteeing the claimed level of privacy. On the other hand, each time the actual level of entropy for ADV^T is greater than the one sufficient for guaranteeing the user's chosen level of privacy, the anonymizer is decreasing the quality of the LBSs.

In our example, we assume the user sets the entropy value (that is the privacy level) to 1.48, represented by the straight line parallel to the x-axis in Figure 3. We also assume—as for the example in Section 4—three users being visiting the region for a total of 48 footprints, while the ADV^T knowledge is split in two time frames: $\hat{T}_1$ = *morning* and $\hat{T}_2$ = *afternoon*. We use the fixed entropy value (as the one that would be considered by the solution in [5]) to compare with different ADV^T footprints distributions, sampled as possible ADV^T knowledge at time frame $\hat{T}_2$ = *afternoon*. The different scenarios for footprints in $\hat{T}_2$ are obtained as follows: (i) we fix the subset of total ADV footprints for the time frame $\hat{T}_2$, 24 out of 48 in our example, (ii) we fix the number of footprints for user $u_{1,\hat{T}_2}$, and (iii) we let $u_{2,\hat{T}_2}$ vary (x-axis), $u_{3,\hat{T}_2}$ being determined once u_1 and u_2 are known. We report the entropy values computed for $u_{1,\hat{T}_2}$, $u_{2,\hat{T}_2}$, and $u_{3,\hat{T}_2}$ on the y-axis. The analytical results computed on these example scenarios are reported in Figure 3. The results confirm the claim of Theorem 1—the actual level of entropy for ADV^T can be smaller or greater than the one expected for ADV.

In Figure 3, three curves are plotted for ADV^T, setting, respectively $u_{1,\hat{T}_2} = 4$, $u_{1,\hat{T}_2} = 8$, and $u_{1,\hat{T}_2} = 16$. Consistently with Theorem 1, varying footprints distributions may result in ADV^T entropy values (thus adversary effort) much lower than the one calculated for ADV. This is the case for the two curves in Figure 3 obtained with $u_{1,\hat{T}_2} = 4$ and $u_{1,\hat{T}_2} = 16$. ADV^T entropy values greater than 1.48 (see Figure 3, ADV^T curve $u_{1,\hat{T}_2} = 8$) raise another issue. Indeed, on the one hand, a greater entropy for ADV^T (compared to the one for ADV) might imply a privacy level higher than the one requested; on the other hand, this implies a loss in the service accuracy—cloaking the user in an area bigger than necessary. While we plotted only the results for the entropy, the curves we computed for the popularity reflect a shape similar to the ones for entropy—popularity curves have the maximum value of 3 for the uniform distribution obtained setting $u_{1,\hat{T}_2} = 8$, $u_{1,\hat{T}_2} = 8$, and $u_{1,\hat{T}_2} = 8$.

Theorem 1 proves that the problem related to considering time in designing privacy solutions is relevant. However, one might ask how much likely is that the distributions of footprints falls in the case of Theorem 1. In fact, if the chances to fall into such a scenario were very small, this could be considered not a big concern. In the sequel, we show that this is not the case, that is, the chances to match the conditions for which Theorem 1 holds are not negligible.

To investigate this aspect, we considered the following example. In a scenario with two users, we set the number of footprints for the two users, respectively to $u_1 = 5$ and $u_1 = 8$. We vary all the possible distributions of the user footprints split into two time frames $\hat{T}_1$ = *morning* and $\hat{T}_2$ = *afternoon*. For each possible distribution we calculate the corresponding entropy. Assuming each distribution to be equally probable, we thus calculate the ratio between the number of occurrences of each entropy value obtained and the total number of possible distributions, 54 in our example. The resulting probability density function is shown in Figure 4. In particular, Figure 4 reports on the probability density of the observed entropy. The entropy calculated for the total number of user footprints is 0.96. It is represented as a vertical line to highlight the points closest to this value. Small squares represent the relation between entropy values (x-axis) and their corresponding probability density (y-axis). We can also observe that the highest probability (0.26) is reached for the entropy value zero obtained for all the distributions, in which at least one of the two users has zero footprints—14 cases in our example.

Figure 5 reports the entropy values obtained for each footprints distribution considered at time frame $\hat{T}_1$ = *morning*. On the x-axis we, vary the footprints value for user $u_{1,\hat{T}_1}$, on the y-axis the ones for user $u_{2,\hat{T}_1}$, and on the z-axis we show the resulting entropy. We notice that the values for $u_{2,\hat{T}_2}$ and $u_{2,\hat{T}_2}$ can be derived, once determined the value

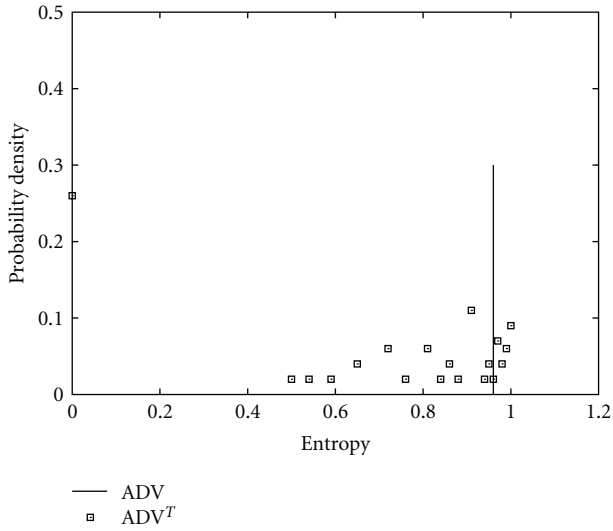


FIGURE 4: ADV^T entropy: probability density function ($u_1 = 5, u_2 = 8$).

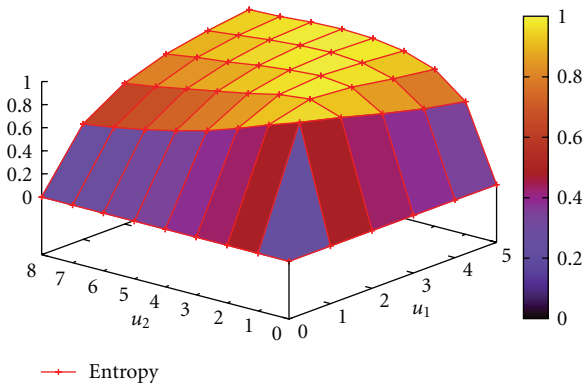


FIGURE 5: ADV^T entropy: $\hat{T}_1$ (*morning*) footprints distributions ($u_1 = 5, u_2 = 8$).

for $u_{1,\hat{T}_1}$ and $u_{2,\hat{T}_1}$, leveraging the above assumptions on the total number of footprints per user. From Figure 5, we can observe that the maximum entropy is obtained, as expected, when the numbers of footprints for user u_1 and user u_2 are the same. We can observe this in the diagonal that goes from point $\langle u_1 = 0,\ u_2 = 0\rangle$ to the point $\langle u_1 = 5, u_2 = 5\rangle$. From this diagonal, when the values for u_2 remains in the high range (e.g., $u_2 = 8$), the entropy remains high. However, when one of the two values decreases, the entropy decreases accordingly. In particular, as already noticed, when one of the two values is equal to zero, the entropy also goes to zero.

5. Comparisons and Discussions

The aim of this section is to discuss the results from the analysis of an existing data set of footprints information. The series of experiments using real data confirms the observation that the feeling-based model, and in particular the solution proposed in [4, 5], while promising in terms of

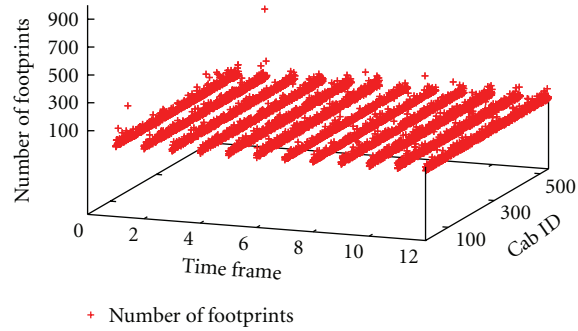


FIGURE 6: R_g global dataset view: cabs footprints per time frames.

user capability to specify the anonymity level, has a problem in dealing with a realistic adversary such as ADV^T.

5.1. Experimenting with Real Data. The *San Francisco Cabs data set* is provided by the Crawdad project [36] and contains traces of 536 cabs vehicles, collected over approximately 30 days in the San Francisco Bay Area (USA). Cab mobility traces are provided by the cabspotting project [37]. Each record in the data set takes the form (*id*, *p*, *t*, *fare*), where $p = (x, y)$ is the location of the vehicle identified by *id* at time *t* and fare formalizes whether the cab itself is busy or not at time *t*. We transform the latitude and longitude coordinates (x, y) provided by the data set in the UTM (Universal Transverse of Mercator) system obtaining a grid-based representation for locations.

We consider for the simulations the region that delimits the Golden Gate Park in San Francisco (referred as R_g). Figure 6 reports an overall view of the footprints in the data set for this region. In particular, on the *x*-axis we vary the time frames starting from $\hat{T}_1$—indicating the 00AM : 02AM time interval—to $\hat{T}_{12}$—indicating the 10P M : 12 PM one. On the y-axis we represents the cab *id*, and, on the *z*-axis we show the corresponding number of footprints for each cab in each time frame.

Among the 536 cabs, we select for the simulation the four cabs (54, 293, 404, 475) with the highest variance as for the number of footprints, with respect to time frames. The footprints trend for these cabs is depicted in Figure 7. We can observe that the footprints of the cab 404 show the highest variation in the time frame $\hat{T}_3$ (04 AM : 06 AM), with 632 footprints; at the same time, they show the same value, 0, in three time frames ($\hat{T}_7$, $\hat{T}_9$, $\hat{T}_{10}$). This means that data for cab 404 vary in a large range but do not vary so much between time frames. On the contrary, the other three cabs vary in a smaller range: from 0 to approximately 250. Thus, they present a higher variation with respect to time frames (see cab 475). Table 3 reports footprints data for the four selected cabs in the region R_g. More specifically, Table 3(a) contains the total number of footprints in the data set while Table 3(b) shows the same data split into 12 time frames.

Similarly to Figure 3, we plotted the results for the entropy corresponding to the footprints distribution in Table 3(b). In Figure 8, the points represent the entropy values calculated for each time frame; the straight line

Table 3: ADV and ADV^T cabs data.

(a) ADV.

R_g	daily
$id = 54$	989
$id = 293$	808
$id = 404$	858
$id = 475$	565
$E(R_g)$	5.62
$P(R_g)$	49.20

(b) ADV^T : $\widehat{T}_j$, $1 \le j \le 12$ (2 hours)

R_g	$\widehat{T}_1$	$\widehat{T}_2$	$\widehat{T}_3$	$\widehat{T}_4$	$\widehat{T}_5$	$\widehat{T}_6$	$\widehat{T}_7$	$\widehat{T}_8$	$\widehat{T}_9$	$\widehat{T}_{10}$	$\widehat{T}_{11}$	$\widehat{T}_{12}$
$id = 54$	244	118	52	52	13	85	207	35	72	45	29	37
$id = 293$	14	48	9	10	14	50	26	67	229	80	160	101
$id = 404$	12	33	632	111	14	17	0	6	0	0	27	6
$id = 475$	27	19	23	81	81	159	47	4	13	34	39	38
$E(R_g, \widehat{T}_j)$	1.94	2.43	1.33	2.50	2.65	2.67	1.92	1.77	2.29	2.67	2.63	2.61
$P(R_g, \widehat{T}_j)$	3.85	5.39	2.51	5.65	6.30	6.36	3.79	3.43	4.90	6.35	6.22	6.09

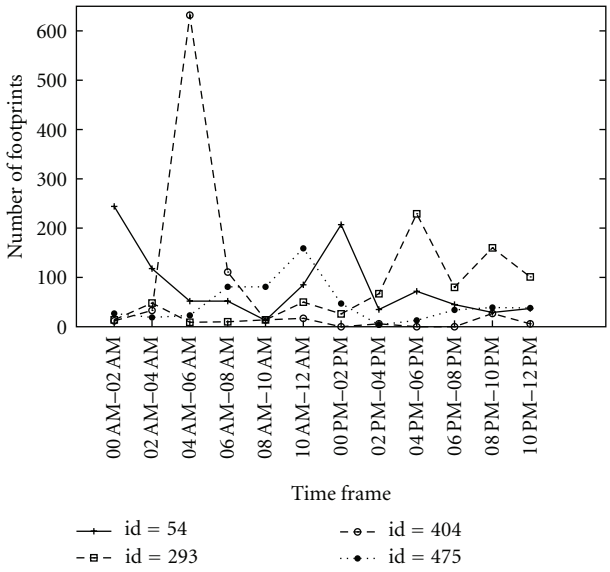


Figure 7: Footprints distribution of four sample cabs, $\widehat{T}_j$, $1 \le j \le 12$ (2 hours).

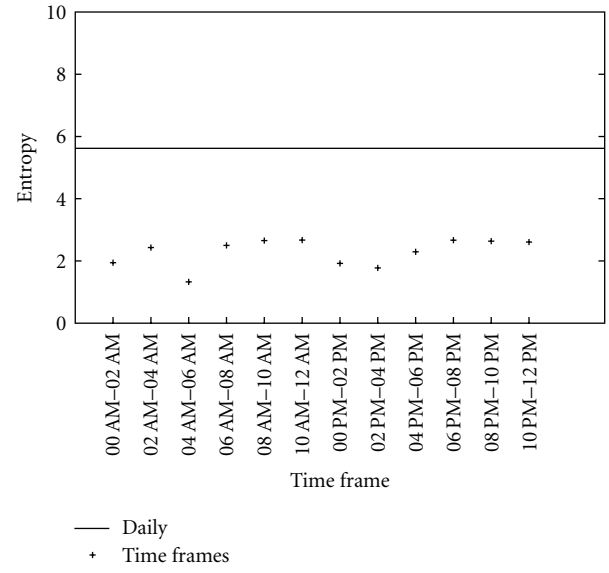
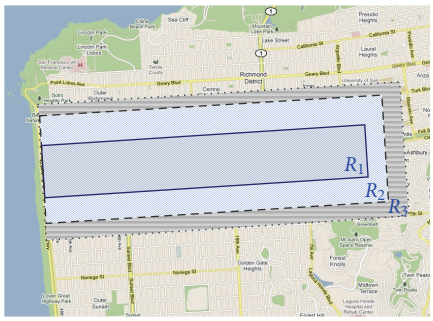


Figure 8: Comparing entropy between ADV and ADV^T for cabs: daily and time frames.

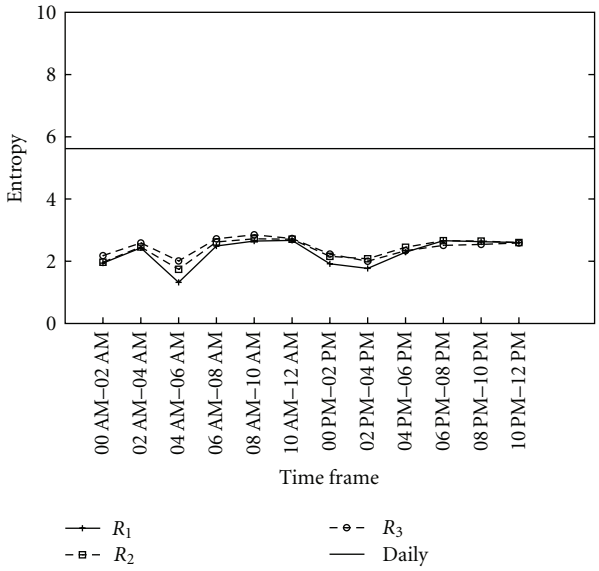
represents the entropy value obtained from the data set considered as a whole. Consistently, with the example in Section 4.1 and with Theorem 1, Figure 8 shows that considering time frames may result in ADV^T entropy values (e.g., 2.65 for time frame $\widehat{T}_5$) much lower than the one calculated for ADV (i.e., 5.62).

We also notice that all the entropy values depicted in Figure 8 are lower with respect to the ADV value of 5.62; also the maximum gap between ADV and ADV^T is quite significant (i.e., 4.30 for time frame $\widehat{T}_3$ (04 AM : 06 AM)). This is due to the choice of cabs with a number of footprints with high variance between time frames. In fact, this choice implies obtaining the minimum entropy values and thus the worst case in the lack of privacy guarantees.

As this is exactly what we expected, we have further confirmed our preliminary findings through real data analysis. Figure 9 shows how much the service quality can be influenced by considering time frames. In particular, we considered three regions: $R_1 = R_g$, R_2, R_3 of 5.83, 10.22, and 13.98 km^2 area, respectively. The region $R_1 = R_g$ matches the area of the Golden Gate Park, while R_2 and R_3 expand the park area. Figure 9(a) reports R_1, R_2, and R_3 on the map of San Francisco. While one might expect that significantly increasing the area of the region (from 5.83 to 13.98, in our



(a) R_1, R_2, R_3 on the map of San Francisco



(b) Comparing entropy of R_1, R_2, R_3

Figure 9: San Francisco cabs: comparing service quality (region area) and privacy for ADV^T

case) will significantly increase the entropy (e.g., being closer to the ADV entropy value of 5.62), there are cases where this does not happen. In fact, as we can notice from Figure 9(b), even significantly increasing the size of the considered region, the gain in the entropy values is negligible. In particular, varying time frames (x-axis) and the area of the regions (R_1, R_2, R_3), the resulting entropy values (y-axis) are very similar. This highlights how, considering a time-aware adversary, affects both the privacy and the quality of the service.

6. Revisiting Assumptions and Approaches

We observed how the feeling-based model [4, 5], while promising in terms of user capability to specify the anonymity level, has a problem in dealing with a realistic adversary such as ADV^T.

Conducting our experiments on a data set of real vehicles traces strengthens the validity of assertions, with respect to experiments performed on synthetically generated datasets only. Our results also show how the threats to user privacy in LBSs are realistic and motivates further investigations. In fact, the results suggest that the problem of protecting location privacy requires to tackle the assumption that the adversarial knowledge is unknown to the anonimyzer. We believe this scenario to be the most challenging and realistic to consider. In fact, depending on the knowledge that the anonymizer has about the adversary, the following scenarios are possible.

(i) The anonymizer knows that the adversary has traces information structured in time slices of equal size. In particular, the anonymizer also knows the size of such time interval.

(ii) The anonymizer knows that the adversary has traces information structured in time slices of different size. In particular, the anonymizer also knows the size of the smaller time slice.

(iii) The anonymizer does not have any information about the adversary knowledge over the traces.

Let us consider (i) and (ii), that is the hypothesis in which the anonymizer has some knowledge of the adversary. Under these hypothesis, a possible direction could be extending the protocol in [4, 5] in order to handle time in a finer manner, so as to thwart ADV^T. For example, it could be argued that for each time frame (e.g., $\widehat{T}_2$), the LBSs requests in that time frame should be anonymized-based on the footprints of that time frame. We assume the anonymizer being able to restructure the possessed traces over any possible time frame. The point is that the anonymization should be computed considering time frame with time slices that are so small as the ones considered by the adversary, that is, the LDS anonymizes against the worst case scenario. This could be an acceptable solution (that assure the level of privacy promised to the user), even if at a cost of a worse service than the one that might be required (i.e., LBS referred to a broader region). In particular, in case of scenario (i), the anonymizer needs to anonymize the request-based on the traces in the same time frame. For scenario (ii), the anonymizer can do anonymization considering always the smallest time slice used by the adversary. As an example, if the adversary has information structured on time slices of three, two and one hours, the anonymizer should always use time frame of one hour to compute the anonimization region. Furthermore, we assume that time slices can only start at multiples of the smaller time slice. For instance, in the cited example the time slice can only start at the beginning of an hour—8.00 AM, 9.00 AM, and so on. Anonymizing the user in a consistent way (i.e., assuring her always the promised level of privacy) using the footprint model and making no assumption about the knowledge of the adversary, that is, scenario (iii) still an open issue that calls for further investigations.

7. Conclusion

We showed that an adversary that has a time-related knowledge different from the one used by the anonymizer poses a serious threat to the privacy of users of Location-based Services. We specifically considered a recently proposed

footprints privacy model. We showed that, once the time is taken into consideration, the claimed privacy assurance does not hold anymore, even when the adversary knowledge about footprints is partial compared to the one of the anonymizer. We supported our claim with both analysis and a concrete example. In particular, we considered real mobility traces of cabs of San Francisco. The analysis of this data set not only confirmed our claim on a real vehicular network scenario. It also showed that the size of the highlighted problem is all but negligible. In practical scenarios, the distance between the expected (claimed) privacy level is far away from the one actually granted by the system. We concluded the paper highlighting further research directions.

Acknowledgments

The work of this paper is partly supported by the project S-MOBILE, contract no. VIT.7627 funded by STW Sentinels (The Netherlands), and by the grant no. HPC-2010 from CASPUR (Italy).

References

[1] M. Conti, V. T. N. Nguyen, and B. Crispo, "CRePE: context-related policy enforcement for android," in *Proceedings of the 13th Information Security Conference (ISC '10)*, 2010.

[2] "Electronic toll collection california (USA)," http://www.bayareafastrak.org/.

[3] "London congestion charge," http://www.tfl.gov.uk/road-users/congestioncharging/.

[4] T. Xu and Y. Cai, "Exploring historical location data for anonimity preservation in location-based services," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 547–555, Phoenix, Ariz, USA, April 2008.

[5] T. Xu and Y. Cai, "Feeling-based location privacy protection for location-based services," in *Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS '09)*, pp. 348–357, November 2009.

[6] J. Krumm, "A survey of computational location privacy," *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 391–399, 2009.

[7] K. LeFevre, D. J. DeWitt, and R. Ramakrishnan, "Incognito: efficient full-domain K-anonymity," in *ACM SIGMOD International Conference on Management of Data (SIGMOD '05)*, pp. 49–60, June 2005.

[8] A. Solanas and R. Di Pietro, "A linear-time multivariate micro-aggregation for privacy protection in uniform very large data sets," in *Proceedings of the 5th International Conference on Modeling Decisions for Artificial Intelligence (MDAI '08)*, vol. 5285 of *Lecture Notes in Computer Science*, pp. 203–214, 2008.

[9] L. Sweeney, "k-anonymity: a model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[10] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services (MobiSys '03)*, pp. 31–42, 2003.

[11] B. Gedik and L. Liu, "A customizable k-anonymity model for protecting location privacy," in *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '05)*, pp. 620–629, 2005.

[12] T. Xu and Y. Cai, "Location anonymity in continuous location-based services," in *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems (GIS '07)*, pp. 1–8, 2007.

[13] M. F. Mokbel, C. Chow, and W. G. Aref, "The new casper: query processing for location services without compromising privacy," in *Proceedings of the 32nd International Conference on Very Large Data (VLDB '06)*, pp. 763–774, 2006.

[14] C. Y. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based service," in *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems (GIS '06)*, pp. 171–178, November 2006.

[15] C.-Y. Chow and M. F. Mokbel, "Enabling private continuous queries for revealed user locations," in *Proceedings of the 10th International Symposium on Advances in Spatial and Temporal Databases (SSTD '07)*, vol. 4605 of *Lecture Notes in Computer Science*, pp. 258–275, 2007.

[16] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 46–55, 2003.

[17] B. Hoh and M. Gruteser, "Protecting location privacy through path confusion," in *Proceedings of the 1st International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm '05)*, pp. 194–205, September 2005.

[18] J. Freudiger, M. H. Manshaei, J. P. Hubaux, and D. C. Parkes, "On non-cooperative location privacy: a game-theoretic analysis," in *Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS '09)*, pp. 324–337, November 2009.

[19] J. Freudiger, M. H. Manshaei, J. Le Boudec, and J. Hubaux, "On the age of pseudonyms in mobile ad hoc networks," in *Proceedings of the 29th Conference on Information Communications (INFOCOM '10)*, pp. 1577–1585, San Diego, Calif, USA, March 2010.

[20] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Preserving privacy in GPS traces via uncertainty-aware path cloaking," in *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS '07)*, pp. 161–171, November 2007.

[21] J. Freudiger, M. Raya, M. Felegyhazi, P. Papadimitratos, and J. Hubaux, "Mix-zones for location privacy in vehicular networks," in *Proceedings of the 1st International Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS '07)*, 2007.

[22] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Proceedings of the 2nd International Conference on Pervasive Services (ICPS '05)*, pp. 88–97, July 2005.

[23] C. Bettini, X. S. Wang, and S. Jajodia, "Protecting privacy against location-based personal identification," in *Proceedings of the 2nd VLDB Workshop on Secure Data Management*, pp. 185–199, 2005.

[24] K. Sampigethaya, L. Huang, M. Li, R. Poovendran, K. Matsuura, and K. Sezaki, "Caravan: providing location privacy for vanet," in *Proceedings of Embedded Security in Cars (ESCAR '05)*, 2005.

[25] P. Cencioni and R. Di Pietro, "A mechanism to enforce privacy in vehicle-to-infrastructure communication," *Computer Communications*, vol. 31, no. 12, pp. 2790–2802, 2008.

[26] M. Gruteser and X. Liu, "Protecting privacy in continuous location-tracking applications," *IEEE Security and Privacy*, vol. 2, no. 2, pp. 28–34, 2004.

[27] M. L. Damiani, E. Bertino, and C. Silvestri, "The PROBE framework for the personalized cloaking of private locations," *Transactions on Data Privacy*, vol. 3, no. 2, pp. 123–148, 2010.

[28] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K. L. Tan, "Private queries in location based services: anonymizers are not necessary," in *ACM SIGMOD International Conference on Management of Data (SIGMOD '08)*, pp. 121–132, June 2008.

[29] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Proceedings of the International Symposium on Privacy Enhancing Technologies (PETS '02)*, pp. 41–53, 2002.

[30] R. Di Pietro and A. Viejo, "Location privacy and resilience in wireless sensor networks querying," *Computer Communications*, vol. 34, no. 3, pp. 515–523, 2011.

[31] R. Di Pietro, P. Michiardi, and R. Molva, "Confidentiality and integrity for data aggregation in WSN using peer monitoring," *Security and Communication Networks*, vol. 2, no. 2, pp. 181–194, 2009.

[32] S. Ortolani, M. Conti, B. Crispo, and R. Di Pietro, "Event handoff unobservability in WSN," in *Proceedings of the IFIP Open Research Problems in Network Security Conference (iNetSec '10)*, 2010.

[33] O. Abul, F. Bonchi, and M. Nanni, "Never walk alone: uncertainty for anonymity in moving objects databases," in *Proceedings of the 24th IEEE International Conference on Data Engineering*, pp. 376–385, IEEE Computer Society, Washington, DC, USA, 2008.

[34] L. Marconi, R. Di Pietro, B. Crispo, and M. Conti, "Time warp: how time affects privacy in LBSs," in *Proceedings of the 12th International Conference on Information and Communications Security (ICICS '10)*, pp. 325–339, 2010.

[35] R. Shokri, J. Freudiger, and J. Hubaux, "Unified framework for location privacy," in *Proceedings of the 9th International Symposium on Privacy Enhancing Technologies (PETS '10)*, pp. 203–214, 2010.

[36] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser, "CRAWDAD data set epfl/mobility (v. 2009-02-24)," February 2009, http://crawdad.cs.dartmouth.edu/epfl/mobility.

[37] "San francisco exploratorium—cabspotting project," http://cabspotting.org/.