Dynamic Modeling of Location Privacy Protection Mechanisms

Sophie Cerf, Sonia Ben Mokhtar, Sara Bouchenak, Nicolas Marchand, Bogdan Robu

Dynamic Modeling of Location Privacy Protection Mechanisms

Sophie Cerf[1], Sonia Ben Mokhtar[2], Sara Bouchenak[2],
Nicolas Marchand[1], and Bogdan Robu[1]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP[?], GIPSA-lab, 38000 Grenoble, France, firstname.lastname@gipsa-lab.fr

[2] INSA Lyon – CNRS – LIRIS, Distributed Systems Research Group
Lyon, France firstname.lastname@insa-lyon.fr

Abstract. Mobile applications tend to ask for users' location in order to improve the service they provide. However, aside from increasing their service utility, they may also store these data, analyze them or share them with external parties. These privacy threats for users are a hot topic of research, leading to the development of so called Location Privacy Protection Mechanisms. LPPMs often are configurable algorithms that enable the tuning of the privacy protection they provide and thus the leveraging of the service utility. However, they usually do not provide ways to measure the achieved privacy in practice for all users of mobile devices, and even less clues on how a given configuration will impact privacy of the data given the specificities of everyone's mobility. Moreover, as most Location Based Services require the user position in real time, these measures and predictions should be achieved in real time. In this paper we present a metric to evaluate privacy of obfuscated data based on users' points of interest as well as a predictive model of the impact of a LPPM on these measure; both working in a real time fashion. The evaluation of the paper's contributions is done using the state of the art LPPM Geo-I on synthetic mobility data generated to be representative of real-life users' movements. Results highlight the relevance of the metric to capture privacy, the fitting of the model to experimental data, and the feasibility of the on-line mechanisms due to their low computing complexity.

Keywords: Location Privacy, Control of Computing Systems, Modeling, Location Based Services, Points of Interest

Introduction

The democratization of mobile devices has fostered the development of services using the users' location data to provide or improve a service. Everyday examples of Location Based Services (LBS) are navigation applications, recommendation

[2] Institute of Engineering Univ. Grenoble Alpes

systems or fitness tracking apps. LBSs provide users with always more personalized and convenient services but at the cost of personal data publishing. Service providers, or any third party attackers, take advantage of these data to derive always more informations about users. These habits are threats against users privacy, as mobility information are highly sensitive data that can, once processed, lead to the inference of users living and working place [8], relatives [2], political or religious preferences [6], among many other. The foundation stone of advanced inferences are often users' Points Of Interest (POIs), that are places where the user stayed a significant amount of time. POIs, defined by a diameter (in meter) and a duration (in seconds), delimit a zone where and when the user were confined. The protection of POIs will be considered in the remaining of the article as the key challenge for privacy protection [17,8,3].

In order to provide ways to protect users' privacy, Location Privacy Protection Mechanisms (LPPMs) have been developed. This terminology gathers all algorithms that modifies location data in order to improve the users' privacy. There is a high diversity among LPPMs: some are at the user level, other require trusted servers; some are on-line mechanisms, others can only be applied on a dataset; etc. Most of them are configurable algorithms with parameters that enable to leverage their action, i.e. to enforce more or less privacy on the data. This property is highly valuable considering that privacy often comes at the cost of a reduction of the service utility. Hence, a configurable LPPM enable to tune the privacy to utility trade-off.

However, nowadays LPPMs face some limitations. On one hand, the notion of privacy is often addressed as high level, theoretical principles and might lack of practical meaning for average user of mobile devices. It is thus challenging to asses the impact of an LPPM on the privacy of data for a non expert user. On the other hand, the parametrization of LPPMs makes them tricky to use as the user is not able to predict what will be the impact of a given parametrization on her privacy. Moreover, location data are highly dynamic, meaning that a location record may be useful at a given time while it is of none interest few minutes later. Similarly, if a user start to obfuscate her data at a given point, it may take some time before she is actually protected, due to memory of the potential attacker. Thus the measures and predictions must be real time processes.

This paper presents a control-theoric approach to solve these challenges. Control theory is a mathematical framework that deals with dynamic systems and measures; and enables modeling and configuration (i.e. control) of systems. In this location privacy context, control methodology will be used to provide a on-line prediction algorithm that link the configuration of a LPPM to the privacy of a user taking into account her current mobility pattern. Evaluation of this approach is carried out using synthetic mobility data reflecting mobility data characteristics for a well known LPPM from the state of the art, Geo-

Indistinguishability [1].

The remaining of the article is organized as follows: first the location data and LPPMs are introduced and the problem is motivated in Section 2. Then the privacy metric is defined and illustrated in Section 3. Section 4 presents the modeling strategy with both a static and dynamic study. Evaluation of both the metric and the modeling ends the paper in Section 5, prior to conclusion and perspective of this work.

Background & Motivation

The mechanisms under study (LPPMs) manipulate location traces. First, mobility data will be presented, before reviewing the state of the art protection mechanisms and highlighting their limitations that motivate this paper contributions.

Mobility Data

A user location is a latitude and longitude couple sent at a given time to a service. The set of locations over time constitutes a mobility trace. Even though the raw information contained in a mobility trace is mathematically extremely simple, the amount of extractable information is almost limitless due to its semantics, especially if it is considered through its dynamic aspect. Indeed, a mobility trace reveals the transportation mean used, the places visited [8], the people encountered or even the name of the user when other sources of information are used for correlation such as maps or directories [7,13]. The analysis of a mobility trace can also lead to prediction of users' next moves based on their habits [22,9].

As it is highly complex and non-relevant to explore the entire properties of location data. In the following of this work, the mobility sets will be considered from the point of view of the user's speed and dispersion. This simple level of abstraction is particularly relevant for the location privacy formulation as the key notion of Points of Interest is linked with concentration of points in time and space, i.e. low speed and low dispersion. The movement will be first assumed unidirectional - this assumption will be discussed in the evaluation section. The variation of users' speed over time (values and frequencies of changes) will enable to represent the various mobility patterns a user could have, see Section 5.1 for further details.

Protection Mechanisms

Mobility data are considered as input for Location Privacy Protection Mechanisms. The aim of these algorithms is to provide obfuscated location data that, when sent to the service, improve the user's protection. The way to achieve privacy protection (i.e. the algorithm by itself) defines the various categories of LPPMs. A LPPM can work in real-time or require a dataset, it can work at the user level or

require a trusted server, etc. LPPMs realize various transformation to data: blurring [18,1], cloaking [10,21,15], merging [4,16], etc., see [20] for a complete review.

Another classification of LPPM is regarding the type of privacy it guaranties. A classic mechanism consist in hiding a user among $k-1$ others (called *k*anonymity) [19], see Never Walk Alone [16] for a location implementation. This LPPM merges the traces of users in order to make them anonymous within a set of users. The notion of *k*-anonymity has been extended by *l*-diversity by Machanavajjhala et al. [12] and *t*-closeness by Liu et al. [11].

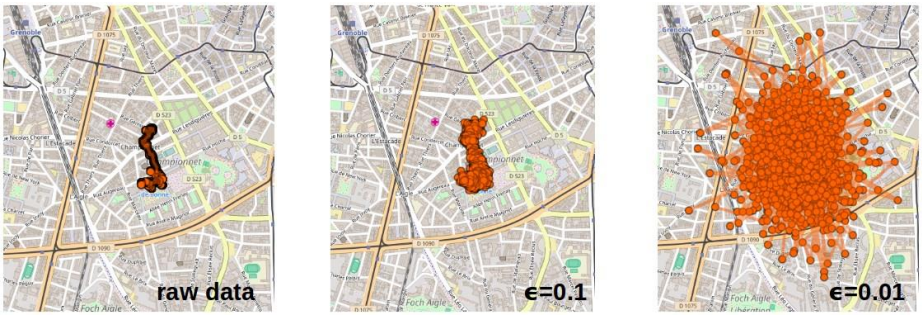


Fig.1. Application of Geo-I on a mobility trace for various configurations: (a) raw mobility data, (b) obfuscation with low noise ($\epsilon = 0.1\ m^{-1}$), (c) obfuscation with high noise ($\epsilon = 0.01\ m^{-1}$)

Another well known approach of privacy is -differential privacy [5], that quantifies the amount of data extractable from a dataset by . A location version of this algorithm has been developed by Andres et al. [1], called GeoIndistinguishability (Geo-I). It is an on-line mechanism that adds spatial noise at each location. An illustration of applying Geo-I to a mobility trace is given in Fig. 1. A key notion in Geo-I is the value of that quantifies the dispersion of the probabilistic distribution in which the value of the noise is drawn. Practically, the lower is, the more noise is added and thus the better the user is protected. Typical range of variation is$\epsilon \in [10^{-4};1]$, expressed in inverse of meter. The impact of the values of can also be seen in Fig. 1. The tuning of Geo-I parameter enables to leverage the privacy protection and also the utility of the data sent to the LBS. The noisier the data are, the less accurate the service will be.

In the following of the paper, when needed, the methodology will be applied to Geo-I, as it is one of the most used LPPM able to work in real time. Indeed, only a LPPM that obfuscate data on-line can be used to study the evolution of user's privacy over time. The methodology presented in the paper can apply for LPPMs satisfying the following requirements:

- being an on-line process, every location is individually obfuscated in real time,
- being tunable by a single parameter, such as the of Geo-I,
- being user centric: the obfuscation should not depend on other people's location or other properties such as the density of the area.

In a general way, most perturbation based mechanisms can be used, such as CloakDroid [14].

Motivation

It can be seen from Fig. 1 that the more noise is added to the mobility data, the more the user's privacy will be protected, as it decreases the accuracy of the attacker knowledge. However, in the same time, it will damage the quality of the service provided to the user. Even if these trends are intuitive, some quantification is missing regarding the protection of the user points of interest. Data of Fig. 1 (c) is less private that those of Fig. 1 (b), itself being more privacy preserving that Fig. 1 (a), but how to measure the differences between the levels of protection? There is a need for a privacy metric.

Second, the information on the level of privacy of a mobility trace may not be sufficient. Indeed, the end goal is to be able to use thoughtfully a protection mechanism, to be able to chose a LPPM among several and configure it in a way that ensure a user's expectations. In our applicative case with Geo-I, the idea is to get a mathematical relation between the configuration parameter and the privacy of the obfuscated data taking into account the user movement.

Moreover, due to the dynamics of a user mobility trace, the privacy protection of a user may also vary independently of the LPPM action. As privacy is POIrelated, if a user is moving fast for a long time (i.e. being in a train), he or she is protected as no POI can be extracted (or more precisely the smallest POI extractable is really large, and not containing much semantic information). Then as soon as the user stops, the threats on his or her privacy is increasing as the information about the stopping point is likely to be personal (i.e. home).

Measuring Privacy in Real-time

In this section, the problem of measuring POI-related privacy *in practice* is addressed. Privacy is defined as the radius of the smallest POI that can be extracted from the mobility trace over a past time window. Formal definitions, justifications and illustrations are given in this section.

This work takes as assumption that the objective of a user in terms of privacy is to prevent an attacker from retrieving her points of interest [8,17,3]. A point of interest is formally defined as a circular zone of a given diameter (in meters) where the user spent a significant amount of time. The ability to have one's POIs hidden is defined as being privacy. The POI diameter and minimal duration are parameters that allow to refine the POI definition to better fit a user's point of view

about her own privacy. For instance, if a user considers that work place and home are sensitive information but do not really care about other people knowing where she has lunch, the minimal duration of the user's POI should be set quite large. Moreover, if a user does not mind others to know the neighborhood where she lives but still want to keep the exact address private, the POI diameter can be set quite small. In the following, POIs are thus considered parametrized by users. Values will be picked for experimental validation but the developed method apply independently of the chosen values.

For the addressed problem, one should have an *on-line* measure of privacy. The privacy signal should represent how likely the user is to reveal a POI, i.e. if she is spending a significant time in a restricted area. Privacy is defined based on the dispersion of the obfuscated data over a past time window. Indeed a small dispersion represents a concentration in space and in time (due to the time window calculation) of location records, which matches with the definition of a POI. Formally, the privacy signal is defined as being the maximum distance between any location record of the time window to the centroid of these points. The current location record $l(k)$ is considered as being the vector of the user's coordinates at the surface of the earth at time k. Then, the centroid $l_c(k)$ of the locations over the past window of length T is defined by eq. (1):

$$l_c(k) = \frac{1}{T} \sum_{t=k-T}^{k} l(t \qquad) \tag{1}$$

and the privacy metric at time k is then:

$$priv(k) = \max_{t \in [k-T;k]} dist[l(t), l_c(k \qquad)] \tag{2}$$

with $dist[x,y]$ being the euclidean distance between two points x and y at the surface of the earth. The privacy signal is expressed in meters and is to be related with the POI diameter. The length of the time window T is again chosen by the user to fit her conception of privacy. Thus defined, the privacy measure at a given time is the radius of the smallest zone in which the user spent her last T seconds.

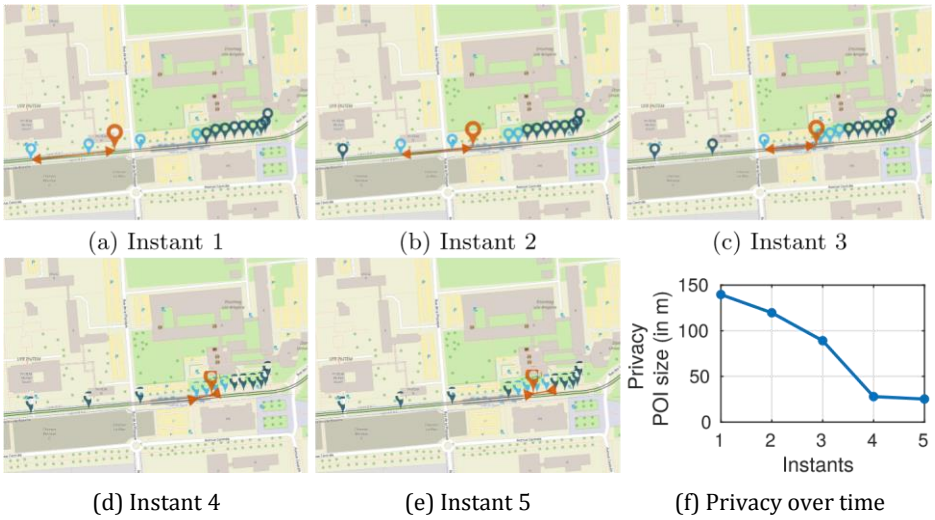


Fig.2. Privacy metric computation on a simple mobility trace.

An illustration of the metric computation is given in Fig. 2. Each subfigure (a) to (e) is the privacy computation at a given instant. Records of the user location are the small markers. The user was in a tram (most distant point revealing high speed) went out and started walking (points close one to another). The lighter points are the ones in the moving time window T. For instance on Fig. 2 (a), the privacy is computed for the fourth point (left to right), and the duration of the window T is four samples. Fig. 2 (b) illustrates the privacy at the instant of the fifth point, and so on. The centroid l_c is the large location position. The privacy metric is then the maximum distance between the centroid and any point of the time window (light points), drawn by the arrow. In this illustration, as the user is slowing down and is likely to arrive in an significant place for her, the privacy metric decreases, as reported in the graph Fig. 2 (f).

Dynamic Modeling of Location Privacy

This section presents a methodology to derive a predictive model of the impact of a LPPM on privacy, by taking Geo-I example. The objective is to have a mathematical equation that links the LPPM configuration () and the user movement (raw trace data) to obfuscated data privacy (*priv*), at each time instant.

Objectives and Methodology

The modeling requirements are the following:

- accurate fitting: the predicted privacy should be close to the actual one,
- light computation: the model is aimed at working on-line on a smartphone, – robustness: no matter the user's movement, the fitting should be accurate.

The model is derived in a two-step process: first through a static characterization and then by exploring its dynamic behavior over time. The general methodology is explained before detailing and applying it for Geo-I in the next two subsection.

Two parameters have been identified as influencing the privacy of a user trace: the LPPM parameter and the properties of the raw mobility trace by itself. In order to deal with this duality, we make the prior assumption that the two parameters are independently acting on privacy, and thus that the privacy function can be linearized. The limits of this assumption will be discussed in Section 5.3. Based on this assumption, the -to-privacy function will be studied for various trace speed: high (50 km/h), low (5 km/h) or null (the user is stopped).

The modeling is carried out in two steps: first a static characterization, where the LPPM is run with a constant configuration and the steady state privacy (the equilibrium value when the privacy has stabilized) is measured; then a dynamic study, where the LPPM parameter suddenly change in a step-wise way and the evolution of the privacy over time when reaching a new equilibrium value is analyzed.

In order to deal with the stochasticity of the LPPM, each simulation is run 100 times, only the means of the outputs are presented.

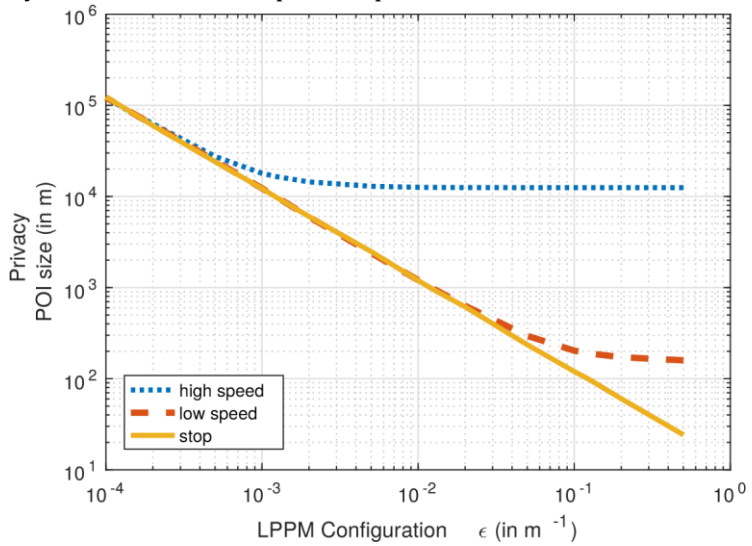


Fig.3. Static characteristic for various user's speed. Mean over 100 experiments.

Static Characterization

First, Geo-I is applied to several mobility traces, each one being the movement of a user with different speeds. Several experiments are launch per trace, each one with a different value of the parameter (values taken in its definition range); the steady state (i.e. converged) values of privacy are measured. Results - privacy over - are reported in Fig. 3, where each curve is a different trace. The following statements can be formulated: (i) the logarithm of privacy is linear with respect to the logarithm of Geo-I parameter for low values of (high noise) and (ii) for high values of (low noise) there is a saturation, and the level of this saturation depends on the user's speed.

The saturation reflects that there are some conditions, for instance if the user is moving fast, for which adding few noise has no impact on the privacy as the user is already protected (i.e. only POI with large diameters can be extracted from the raw trace). The linear part of the curve means that, at some point, the more noise is added to the data, the larger the diameter of the extracted POI is. The linear part of the static characteristic has the same equation in all cases:

$$log(priv) = a\ log(\epsilon) + b. \tag{3}$$

The saturation level corresponds to the privacy of the mobility data when $\epsilon \to +\infty$, i.e. no noise is added. It is then the privacy of the raw trace, that can be measured in real time thanks to equations (1) and (2). This value is denoted $priv_{raw}$.

The transition between the two zones is at $_0$, which corresponds to the intersection of the linear curve with the constant part:

$$\epsilon_0\ s.t.\ a\ log(\epsilon_0) + b = log(priv_{raw}). \tag{4}$$

Hence

$$\epsilon_0 = 10^{\frac{log(priv_{raw}) - b}{a}}. \tag{5}$$

Dynamic Study

Fig. 3 highlighted the zones in which the behavior from the LPPM parameter to the privacy measure is linear ($\epsilon < \epsilon_0$). Hence for the dynamic analysis, the step variation of Geo-I's parameter will be chosen as being part of this linearity zone, otherwise has no impact on privacy. The measures of privacy over time while changing suddenly the LPPM parameter are reported in Fig. 4.

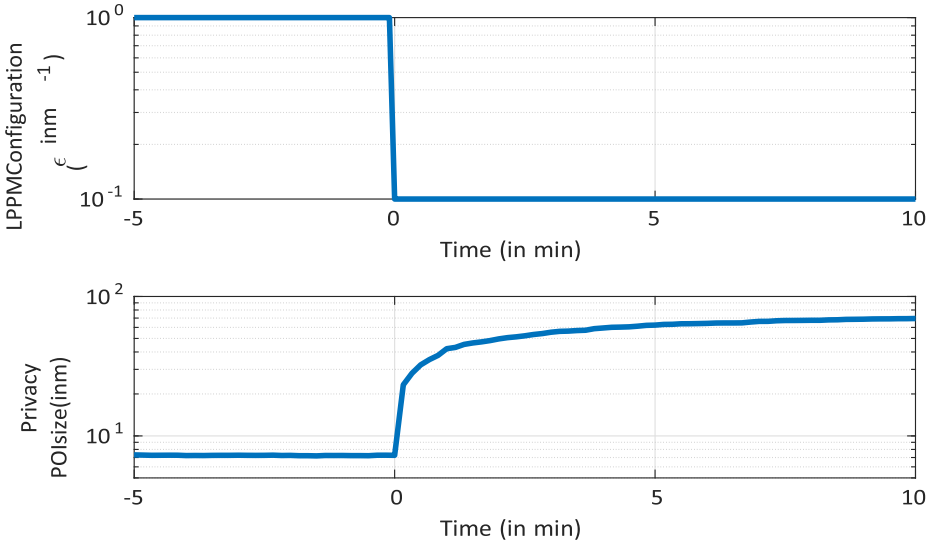


Fig.4. Privacy evolution over time with Geo-I parameter varying from$\epsilon = 1\ m^{-1}$ to $\epsilon = 10^{-1}\ m^{-1}$, during a stop. Mean over 100 experiments.

Several conclusions can be drawn from this figure: (i) in steady state, i.e. at equilibrium, the privacy amplification has approximately the same size as the amplification of . This consolidates the static characterization results. (ii) There is a dynamic change of privacy: it takes some time before the measure reaches its steady state value.

The shape of the privacy signal seems to be close to the inverse of an exponential, which makes us look for a recursive equation of the form:

$$log(priv(t)) = \alpha.log(priv(t-1)) + \beta\ .log(\epsilon(t)) + \gamma. \tag{6}$$

When time goes to infinity, eq. (6) should fit eq. (3) as it corresponds to the steady state value of privacy. This creates the following constraints:

$$a = \frac{\beta}{1-\alpha}, \quad b = \frac{\gamma}{1-\alpha.} \tag{7}$$

These two constraints let one degree of liberty in eq (6). This enables to tune the time dynamics of the response, i.e. the time the privacy signal takes to reach its steady state. Parameters *α,β* and *γ* can be found using simple regression tools.

The resulting model for privacy prediction, combining both static and dynamic studies is the following:

$$\text{if} \epsilon < \epsilon_0$$

$$priv(t) = \begin{cases} 10^{\alpha.log(priv(t-1))\beta .log(\epsilon(t))+\gamma} & priv_{raw}(t) \\ & \text{otherwise.} \end{cases}$$

This equation enables to predict, for each time instant, the value of privacy knowing the obfuscation level (), the past value of privacy ($priv(t-1)$) and the raw trace properties ($_0$ and $priv_{raw}$).

Evaluation

In this section, both the privacy metric and the prediction model are evaluated. Prior to this, the mobility scenario on which this evaluation is based is presented.

Evaluation Setup

The objective of the metric and model is to capture the privacy of users *no matter their mobility pattern*. This notion of disturbance being essential in this work, the contributions should be evaluated with the best representative mobility scenario. As explained in Section 2.1, two key properties of a mobility trace are the speed of the user and the frequency of variation of this speed. The main advantage of using a synthetic trace is that the moves are perfectly known, hence the trace is labeled at each instant with ”stop” or ”move”.

The synthetic trace is sampled every 10 seconds and has varying speeds (0, 5, 50, 150 km/h) representing various transportation means (stop, walk, car, train, etc.). The periods between two changes range from 30 seconds (e.g. stop at a traffic light) to one hour (e.g. medical visit), including middle values as 5 minutes (e.g. stop in a coffee shop). The synthetic mobility trace is illustrated in Fig. 5 by its speed over time. The total trace is 18 hours long. Other mobility properties are included, such as turnings (hours 1 to 2), acceleration and decelerations (hours 8 to 9) and local movements (i.e. the user’s speed is *almost* zero, between

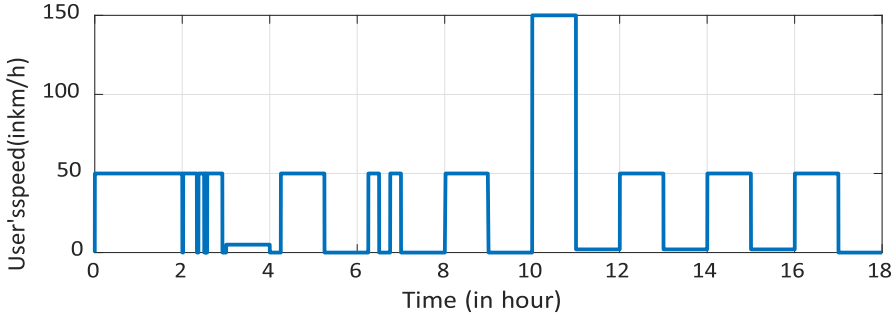


Fig.5. Mobility trace: changes of user’s speed over time

hours 10 and 18). However, due to space restrictions, these properties are hardly visible when plotting only speed over time as in Fig. 5.

The value of the time window on which the privacy metric is calculated is fixed at the medium value of T = 15 *min* (i.e. stops of more than 15 min must be protected). The parameters of the models of eq. (3) and eq. (6) have been found using Matlab [R] identification tool, that carries out regressions for modeling. The values of the parameters are $a = -1$, $b = 0.85$, $\alpha = 0.9474$, $\beta = 0.0526$ and $\gamma = 0.0447$. Indeed, if the duration of the time window T is changed, the previous parameters would change too. Only the regression mechanism would need to be run again. Without loss of generality, the evaluation will be presented only for $T = 15$ *min*.

Privacy Metric Evaluation

The privacy sensor is applied to the mobility trace described just before without any obfuscation. Results are illustrated in Fig. 6, which plots privacy over time, where dark dots are during the user movement and light ones during a stop. The privacy signal reflects the user's stop by having decreasing values. Privacy tends to zero with some dynamics which is due to the time-window calculation of the metric (T = 15 min for this plot).

If one takes small values of the privacy as being a stop indicator, each detection corresponds to a stop, precision is then of 100%. False negative can be found around hour 2 to 3, leading to a recall of 70%. These false negative correspond to short stops of less than T =15 minutes. They can be identified by considering the privacy derivative sign (that should thus be negative). However, it would lead to reduction of the metric precision, due to the presence of turnings at hour 1 to 2 that also generate decreasing privacy. However, if one goes back to the definition of privacy, stops shorter than T do not define POIs and thus are not a threat on users' privacy.

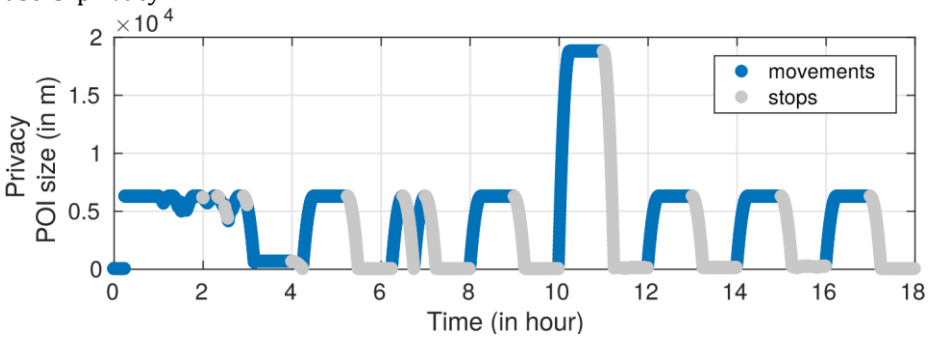


Fig.6. Privacy over time of the raw mobility trace.

As a conclusion the presented metric, based on the radius of the smallest POI that can be extracted from a past time window, successfully reflects users' privacy: the smallest it is, the more sensitive the mobility trace is.

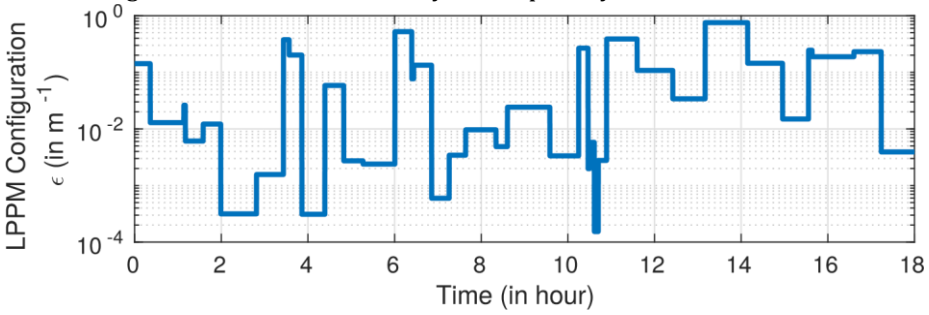
Prediction Model Evaluation

The accuracy of the model presented in Section 4 is now investigated, using the mobility trace of Fig. 5. The model input scenario is illustrated in Fig. 7, top plot: is taken to vary in its whole range of values with changes at various frequencies (randomly chosen between 10 seconds and one hour). The comparison of the measured data and the model predictions are in Fig. 7, bottom plot.

The two curves are almost identical, indicating a good model accuracy most of the time. At some instants (around 3h, 6h, etc.) the model fails to perfectly match the reality. These moments corresponds to situations where the LPPM configuration raises with a large amplitude and for a long time. In these cases, the model predicts a decrease of privacy which is faster than the reality. However, the steady state value achieved is correct. Note that the modeling is always underestimating the privacy, which is more valuable that overestimating. The model accuracy could be improved by modeling this non-linear behavior. However, it will be with a cost in complexity, which would not be necessary beneficial considering the intended implementation of this algorithm on a smartphone. An extended analysis of this point will be done in a future work.

The computing complexity of the algorithm of eq. (6) is $O(1)$, as it consists only of scalar products and sums. This makes the modeling algorithm suitable for a real-time usage.

To conclude, the model is able to successfully capture the influence of the LPPM Geo-I configuration and user's mobility on the privacy metric in an online fashion.



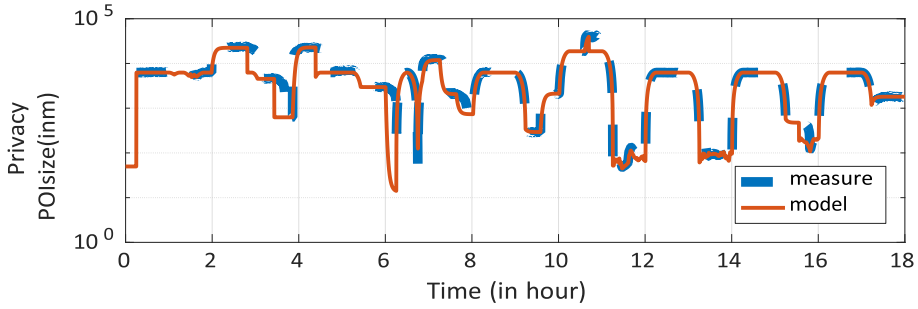


Fig.7. Model evaluation. Top plot: LPPM parameter variations over time. Bottom plot: Comparison of measured privacy and predicted one (mean over 100 experiments).

Conclusions

The democratization of Location-Based Services has increased the threats on users' privacy. Location Privacy Protection Mechanisms (LPPMs) have been developed to tackle this issue. Yet, the existing algorithms often lack of applicability for mobile devices users as they do not provide practical ways neither to evaluate nor to predict the gain in privacy. In this paper a model-based approach is presented, that enables users to predict their privacy when using such protection mechanisms, regardless of their mobility behavior. Contributions are on the definition of real-time Points of Interest oriented privacy metric and on the modeling of the impact of a state-of-the-art LPPM on users' privacy. Evaluation carried out in simulation highlight the relevance of the model formulation and the efficiency of the prediction to fit the real data. The future of this work will be its evaluation using data collected from real users, as well as the development of strategies to configure LPPMs to ensure privacy objectives.

References

1. Miguel E. Andr´es, Nicola´s E. Bordenabe, Konstantinos Chatzikokolakis, andCatuscia Palamidessi. Geo-indistinguishability: Differential Privacy for Locationbased Systems. In *CCS*, pages 901–914, 2013.
2. Igor Bilogrevic, K´evin Huguenin, Murtuza Jadliwala, Florent Lopez, Jean-PierreHubaux, Philip Ginzboorg, and Valtteri Niemi. Inferring Social Ties in Academic Networks Using Short-Range Wireless Communications. *Wpes*, pages 179–188, 2013.
3. Sophie Cerf, Vincent Primault, Antoine Boutet, Sonia Ben Mokhtar, Robert Birke,Sara Bouchenak, Lydia Y Chen, Nicolas Marchand, and Bogdan Robu. Pulp:
   Achieving privacy and utility trade-off in user mobility data. In *Reliable Distributed Systems (SRDS), 2017 IEEE 36th Symposium on*, pages 164–173. IEEE, 2017.

4. Kai Dong, Tao Gu, Xianping Tao, and Jian Lu. Complete bipartite anonymity: Confusing anonymous mobility traces for location privacy. In *Parallel and Distributed Systems (ICPADS), 2012 IEEE 18th International Conference on*, pages 205–212. IEEE, 2012.
5. Cynthia Dwork. Differential Privacy. In *Automata, Languages and Programming*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2006.
6. Lorenzo Franceschi-Bicchierai. Redditor cracks anonymous data trove to pinpoint muslim cab drivers. http://mashable.com/2015/01/28/ redditor-muslim-cab-drivers/, January 2015.
7. Sebastien Gambs, Marc-Olivier Killijian, and Miguel Nunez del Prado Cortez.De-anonymization Attack on Geolocated Data. *2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pages 789–797, 2013.
8. S´ebastien Gambs, Marc-Olivier Killijian, and Miguel Nu´n˜ez del Prado Cortez.Show Me How You Move and I Will Tell You Who You Are. *Transactions on Data Privacy*, 4(2):103–126, August 2011.
9. S´ebastien Gambs, Marc-Olivier Killijian, and Miguel Nu´n˜ez del Prado Cortez.Next place prediction using mobility markov chains. In *Proceedings of the First*
   Workshop on Measurement, Privacy, and Mobility, page 3. ACM, 2012.
10. Bugra Gedik and Ling Liu. A customizable k-anonymity model for protecting location privacy. Technical report, Georgia Institute of Technology, 2004.
11. Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacybeyond k-anonymity and l-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 106–115. IEEE, 2007.
12. Ashwin Machanavajjhala, Johannes Gehrke, Daniel Kifer, and MuthuramakrishnanVenkitasubramaniam. l-diversity: Privacy beyond k-anonymity. In *Data Engineering, 2006. ICDE'06. Proceedings of the 22nd International Conference on*, pages 24–24. IEEE, 2006.
13. Mohamed Maouche, Sonia Ben Mokhtar, and Sara Bouchenak. Ap-attack: A noveluser re-identification attack on mobility datasets. In *MobiQuitous*. ACM, 2017.
14. Kristopher Micinski, Philip Phelps, and Jeffrey S Foster. An empirical study oflocation truncation on android. *Weather*, 2:21, 2013.
15. Mohamed F Mokbel, Chi-Yin Chow, and Walid G Aref. The new casper: Queryprocessing for location services without compromising privacy. In *Proceedings of the 32nd international conference on Very large data bases*, pages 763–774. VLDB Endowment, 2006.
16. Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 261–268. IEEE, 2009.
17. Vincent Primault, Sonia Ben Mokhtar, C´edric Lauradoux, and Lionel Brunie. Differentially Private Location Privacy in Practice. In *MoST'14*, San Jose, United States, 2014.

18. Vincent Primault, Sonia Ben Mokhtar, C´edric Lauradoux, and Lionel Brunie. Timedistortion anonymization for the publication of mobility data with high utility. In *TrustCom*, pages 539–546, 2015.
19. Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.
20. Marius Wernke, Pavel Skvortsov, Frank Du¨rr, and Kurt Rothermel. A classificationof location privacy attacks and approaches. *Personal and ubiquitous computing*, 18(1):163–175, 2014.
21. Yi-Chin Wu, Karthik Abinav Sankararaman, and St´ephane Lafortune. Ensuringprivacy in location-based services: An approach based on opacity enforcement. *IFAC Proceedings Volumes*, 47(2):33–38, 2014.
22. G¨okhan Yava¸s, Dimitrios Katsaros, Ozgu¨r Ulusoy, and Yannis Manolopoulos. A¨ data mining approach for location prediction in mobile environments. *Data & Knowledge Engineering*, 54(2):121–146, 2005.