The Same-Origin Attack against Location Privacy

George Theodorakopoulos
School of Computer Science & Informatics
Cardiff University
5 The Parade, Cardiff CF24 3AA, UK
TheodorakopoulosG@cardiff.ac.uk

ABSTRACT

A plethora of applications benefit from location context, but a person's whereabouts can be linked to her personal sensitive information. Hence, protection mechanisms have been proposed that add systematic noise to a user's location before sending it out of the user's device. We describe the same-origin attack, to which a group of such mechanisms are vulnerable, we evaluate it against two mechanisms (spatial cloaking and geo-indistinguishability), and we propose our own mechanism, inspired by the maximum entropy principle. We find that spatial cloaking is much worse than the other two, and the maximum-entropy mechanism performs slightly better than geo-indistinguishability. Designing an optimal mechanism remains an open problem.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*Security and protection*; G.3 [**Mathematics of Computing**]: Probability and Statistics; K.4.1 [**Computers and Society**]: Public Policy Issues—*Privacy*

General Terms

Security

Keywords

Location Privacy; Mobile Networks; Maximum Likelihood Estimation

1. INTRODUCTION

Location information can provide context to applications, thus enabling location-based smartphone apps (e.g. of the type "find the nearest Point-of-Interest") and participatory sensing platforms (e.g. for collecting traffic data). Undeniably useful, location information is also very sensitive, as a person's whereabouts can reveal her identity and other personal information such as religion and political affiliation.

WPES'15, October 12, 2015, Denver, Colorado, USA.

DOI: http://dx.doi.org/10.1145/2808138.2808150.

Of the mechanisms that have been proposed to protect privacy, *obfuscation* is a straightforward and intuitively appealing one: Obfuscation merely adds some form of noise before transmitting the user's location out of the device. Hence, it is easy for a user to comprehend, and it does not require any changes to the infrastructure (e.g. there is no need for a trusted-third party, nor for the server that provides the location-based service to implement its part of a privacy-preserving protocol).

There are two approaches to obfuscation in the literature: *with-prior* and *without-prior*. The former assumes that the adversary (e.g. the server) has some prior information in the form of a probability distribution about the user's location. This assumption is then used to generate a protection mechanism that is customised for that prior probability distribution. Shokri et al. have quantified privacy as the adversary's error in estimating the user's location [7], and they have proposed optimal with-prior mechanisms [8]. The without-prior mechanisms make no such assumption, and they are thus independent of any prior knowledge that the adversary may or may not have. A generic mechanism in this group is spatial cloaking, the main idea of which is to obfuscate a location into a larger area that contains it. Geo-indistinguishability [1] is a principled representative of the without-prior group, extending differential privacy [5] into domains with distance metrics (such as the Euclidean distance for locations) [2].

In this paper, we focus on the without-prior group of privacy mechanisms and, in particular, we scrutinize the fact that the added noise depends on the user's true location, forming in effect a signature for that location. So, when the user sends repeated queries from the same location, the attacker can match the observed noise pattern to a location that most likely produced it. Our first contribution is a maximum likelihood estimation technique that characterizes the *speed* with which the attacker can recover the true location. We then apply it to three protection mechanisms: The aforementioned spatial k-cloaking and geo-indistinguishability, and a new one that we design based on the maximum-entropy principle. For each, we find how many queries a user can send from the same location before the attacker can localize her with some particular accuracy (measured by distance to the true location or by the probability of finding the true location). Comparing the three mechanisms for similar amounts of added noise, spatial cloaking is by far the worst, whereas the other two perform approximately the same, with a slight advantage for the maximum-entropy one. The question of designing an optimal mechanism is an open problem.

2. THE SAME-ORIGIN ATTACK

We consider a user sending queries that contain her location to a Location-Based Server (LBS). A protection mechanism obfuscates the location in each query and sends it along to the LBS. The attacker (possibly the LBS itself) intercepts the queries and tries to infer the user's true location.

The basic premise of the same-origin attack is that the attacker has external reasons to believe that the user sends all the obfuscated queries from the same location. For example, these locations could all be sent at 9pm each day, so the user is probably at her home, or the user is currently known to be sending her queries from a secret meeting, but the location of the meeting is unknown.

The protection mechanism generically works as follows: When the user sends a location query from location i, it produces a location j with probability $p_i(j)$, for each i, j in a region of interest R, assumed to be discretized as a grid. Every time the user sends another query, the mechanism draws another obfuscated location from the appropriate $p_i()$, *independently* of all previously reported locations.

The attacker knows all the probability mass functions (pmfs) $p_i(\cdot)$ that the mechanism employs and he collects a sequence of reported locations $x_{1:t} = x_1, x_2, \ldots, x_t$, all of which are assumed to come from the same (unknown) location. His objective is to find the true location of the user. To do this he computes the probability of generating the sequence $x_{1:t}$ from location i, i.e. with pmf $p_i(\cdot)$, and selects the location that is most likely to have produced the observed sequence.

Sequential Hypothesis Testing. Having observed t locations, the attacker computes the likelihood of the hypothesis that the user's true location is $i \in R$:

$$L(i|x_{1:t}) = p_i(x_{1:t}) = \prod_{s=1}^{t} p_i(x_s), \tag{1}$$

where the second equality holds because the protection mechanism draws locations independently with each query. The attacker then decides in favor of the location with the highest likelihood. Ties are broken randomly.

Note that the attack only requires that the noise added at each location i depend only on i and not, e.g., on the time t or on previously reported locations. In a sense, $p_i(\cdot)$ is used as the signature of location i, and the attack tries to match the observed sequence with one of the signatures.

3. EVALUATION AGAINST PROTECTION MECHANISMS

We now test the above attack against three protection mechanisms, to be described below: spatial k-cloaking, geo-indistinguishability, and a maximum-entropy method that we design in this paper. To be fair among the three, we assume that each adds the same average distance (noise) to the user's true location. For simplicity, we assume that the true location is the origin $(0, 0)$ of the grid that the region of interest is discretized into.

The attack's success is quantified in two ways: First, as the probability of identifying the correct location after t observations (*success probability*). Knowing this curve, the user can decide how many queries to send before the attacker can localize her with a probability that the user deems dangerous. Conversely, note that the attacker can also compute his probability of success after t observations. So, if that probability is too low, he can postpone taking a decision and hope that he will observe more locations.

The second quantification of the attack's success is via the average distance between the true location and the location that the attacker chooses after t observations (*distance error*). The motivation of the second quantification is that, even though the attacker might not identify the correct location with certainty, he will in general be able to come closer and closer with more observations, so the user can see how close he can come after t observations.

Note that the success probability is a scalar number, so the simulation-based estimates shown below are point estimates with a 95% confidence interval (as error bar). In contrast, the distance error is a random variable, so the figures below show its mean and standard deviation (as error bar). For each scenario tested, 200 simulations were done.

Spatial k-cloaking. Let $K(i)$ be the set of locations in a $(2k+1)\times(2k+1)$ square centered on location i, e.g. $K(0,0)$ is the square with endpoints $(-k,-k), (k,-k), (-k,k), (k,k)$. For a query from location i, spatial k-cloaking selects a random location in $K(i)$ with equal probability:

$$p_i^{\text{K-CLOAK}}(j) = \begin{cases} \frac{1}{(2k+1)^2} & \text{if } j \in K(i) \\ 0 & \text{if } j \notin K(i) \end{cases} \tag{2}$$

The average magnitude of the noise that K-CLOAK introduces for queries coming from $(0, 0)$ is $\bar{r}_{\text{K-CLOAK}} = \sum_{m=-k}^{k} \sum_{n=-k}^{k} \frac{1}{(2k+1)^2} \sqrt{m^2+n^2}$.

The likelihood $L(i|x_1) = p_i(x_1)$ of location i for a single observation x_1 is equal to $\frac{1}{(2k+1)^2}$ if x_1 is in $K(i)$, and zero otherwise. Similarly, for a sequence $x_{1:t}$, the likelihood $L(i|x_{1:t}) = p_i(x_{1:t})$ is $\frac{1}{(2k+1)^{2t}}$ if all observations $x_{1:t}$ are in $K(i)$, and zero if even one observation is outside $K(i)$.

The implication for the attacker is that a location i is still a candidate to be the true location after t observations if and only if all t observations could have been produced from i. Equivalently, the surviving candidates are those in the intersection of all the $K(x_s), s = 1, \ldots, t$. As the likelihood of i does not depend explicitly on i nor on the observations $x_{1:t}$ (it is either $\frac{1}{(2k+1)^{2t}}$ or zero), all surviving candidates are equally likely, so the attacker picks one at random.

Figures 1 and 2 show the success probabilities (resp. distance errors) of the same-origin attack against the K-CLOAK mechanism for $k = 2, 5, 10$. We see that, even though the mechanism promises $\frac{1}{(2k+1)^2}$ probability of correct localization, this probability quickly increases for more observations, increasing by about a factor of 10 after 4 observations. Similarly, the distance error drops by about a factor of 2 after 3 observations.

Geo-indistinguishability. Geo-indistinguishability is an extension of differential privacy to location privacy [1]. Its main idea is to add noise in such a way as to bound the likelihood ratio of nearby locations given any particular observation. A mechanism that satisfies geo-indistinguishability is the planar Laplacian centered at i:

$$p_i^{\text{GEO-IND}}(j) = \frac{\epsilon^2}{2\pi} e^{-\epsilon d(i,j)}, \tag{3}$$

where $d(i, j)$ is the Euclidean distance between locations i and j. To prevent the discretization from interfering with the geo-indistinguishability guarantees, we first select continuous (x, y) coordinates, and then map them to the nearest grid point as the inventors of the mechanism suggest [1].

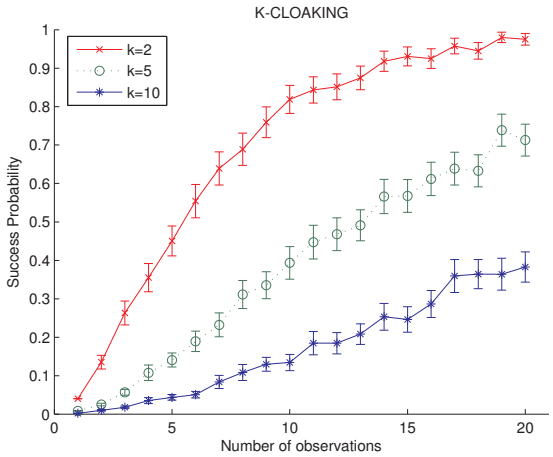

Figure 1: Success probabilities with 95% confidence intervals after t observations for the spatial k-cloaking mechanism, $k = 2, 5, 10$.

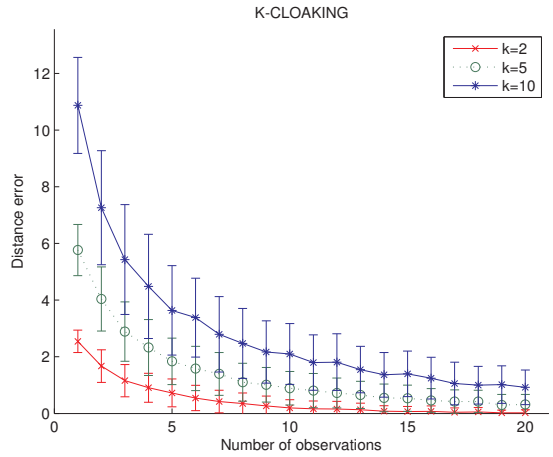

Figure 2: Distance errors after t observations for the spatial k-cloaking mechanism, $k = 2, 5, 10$. Error bars indicate ± one standard deviation from the mean.

The average magnitude of the noise that the GEO-IND mechanism introduces is $\bar{r}_{\text{GEO-IND}} = \frac{2}{\epsilon}$.

The likelihood $L(i|x_1) = p_i(x_1)$ of location i for a single observation x_1 is equal to $p_i^{\text{GEO-IND}}(x_1)$. After a sequence of t observations $x_{1:t}$, the likelihood of location i is

$$L(i|x_{1:t}) = \prod_{s=1}^{t} p_i^{\text{GEO-IND}}(x_s) = \left(\frac{\epsilon^2}{2\pi}\right)^t e^{-\epsilon \sum_{s=1}^{t} d(i,x_s)}. \quad (4)$$

The likelihood-maximizing location, which the attacker picks, is the location that minimizes the sum of distances to the t observations, also known as the *geometric median* of the t observations. The geometric median of 2-dimensional points is unique in general. An exception is when $t = 2$, in which case all points in the linear segment $x_1 - x_2$ have an equal sum of distances to x_1 and x_2. In that case, the attacker selects a point at random.

Figures 3 and 4 display the success probabilities and the distance errors of the GEO-IND mechanism for $\epsilon = 0.2, 0.5, 1$.

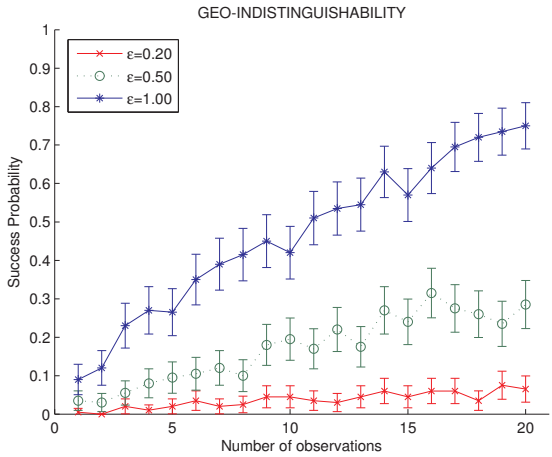

Figure 3: Success probabilities with 95% confidence intervals after t observations for the Geo-Ind mechanism $\epsilon = 0.2, 0.5, 1$.

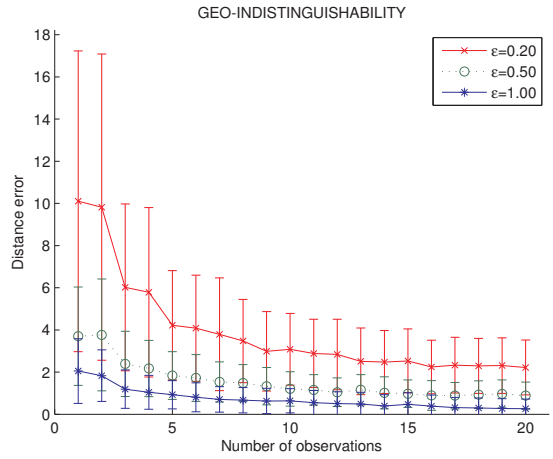

Figure 4: Distance errors after t observations for the Geo-Ind mechanism, $\epsilon = 0.2, 0.5, 1$. Error bars indicate ± one standard deviation from the mean.

Although there is again a significant deterioration of privacy with more observations, it takes the full 20 observations for the success probability to increase 10-fold, when only 4 observations were sufficient for K-CLOAK. This holds across all three ϵ values. The distance error seems to be dropping slightly less rapidly than K-CLOAK, by about a factor of 2 after 4-5 observations.

Maximum entropy. The two previous mechanisms are plausible ways to add noise. In search of a more principled mechanism, we turn to the *maximum-entropy* principle, which we motivate as follows: The amount of noise that a mechanism adds must be constrained in some way, otherwise privacy would be perfectly preserved, but application utility would suffer. However, other than satisfying this constraint, we would like the noise to be, intuitively, as "random" as possible. The maximum-entropy principle provides a way to make this "maximum randomness" intuition concrete, by

selecting a mechanism that has maximum entropy among all those that satisfy the constraint.

The one-dimensional Gaussian distribution, $\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, has the largest entropy among all distributions on the interval $(-\infty, \infty)$ with a given variance σ^2 [4][p. 254]. The two-dimensional analog of the variance, for two random variables x and y, is the covariance matrix

$$\begin{pmatrix} E[(x-\mu_x)^2] & E[(x-\mu_x)(y-\mu_y)] \\ E[(x-\mu_x)(y-\mu_y)] & E[(y-\mu_y)^2] \end{pmatrix},$$

where μ_x is the mean of x, $E[(x-\mu_x)^2]$ is the variance of x, and similarly for y. The quantity $E[(x-\mu_x)(y-\mu_y)]$ is the covariance of x and y. Completely analogously to the one-dimensional case it can be proven, e.g. using the method of Lagrange multipliers, that the two-dimensional Gaussian distribution has the largest entropy among all two-dimensional distributions with a given covariance matrix.

Two Gaussian random variables x and y, each with a given variance σ_x^2, σ_y^2, and constrained by a "variance budget" $\sigma_x^2 + \sigma_y^2 \leq 2\sigma^2$, produce the highest-entropy two-dimensional Gaussian when their variances are equal $\sigma_x^2 = \sigma_y^2 = \sigma^2$ and their covariance matrix is diagonal, i.e. when they are independent, or equivalently when their covariance is zero.

These considerations suggest the following MAX-ENT mechanism that adds noise equal to the aforementioned independent Gaussians x and y to the two coordinates of location i:

$$p_i^{\text{MAX-ENT}}(j) = \frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}d^2(i,j)}, \tag{5}$$

where, as before, $d(i,j)$ is the Euclidean distance between i and j.

The average magnitude of the noise that MAX-ENT introduces is $\bar{r}_{\text{MAX-ENT}} = \sigma\sqrt{\frac{\pi}{2}}$.

As before, for a sequence $x_{1:t}$, the likelihood of location i is

$$L(i|x_{1:t}) = \prod_{s=1}^{t} p_i^{\text{MAX-ENT}}(x_s) = \left(\frac{1}{2\pi\sigma^2}\right)^t e^{-\frac{1}{2\sigma^2}\sum_{s=1}^{t} d^2(i,x_s)}. \tag{6}$$

The likelihood-maximizing location is now the one that minimizes the sum of *squared* distances to the t observations (the center of mass of the observed locations).

Figures 5 and 6 display the success probabilities and distance errors of the MAX-ENT mechanism for $\sigma = 2, 5, 10$. The trends here are similar to GEO-IND.

Comparison of the K-Cloak, Geo-Ind, Max-Ent mechanisms. In Figures 7 and 8 we show results for equal average magnitudes of noise across mechanisms, $\bar{r}_{\text{K-CLOAK}} = \bar{r}_{\text{GEO-IND}} = \bar{r}_{\text{MAX-ENT}} = 4.2$, which correspond to $k = 5$, $\epsilon = 0.48$ and $\sigma = 3.35$. The main conclusion is that K-CLOAK is much worse than the other two, and MAX-ENT is slightly better than GEO-IND in terms of success probability. However, GEO-IND was not specifically engineered against the same-origin attack, and there is no proof that MAX-ENT, although intuitively appealing as a maximum-entropy construction, optimally delays the attacker under a constraint on the magnitude of the noise. Such a proof, or a different optimal construction, is the subject of future work.

Note also that the figures quantify the marginal contribution of each successive observation towards the attacker's success, in terms of an increase in the success probability or a decrease in the distance error. A general conclusion

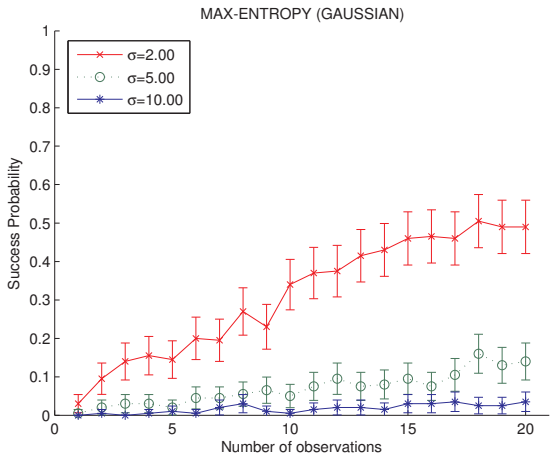

Figure 5: Success probabilities with 95% confidence intervals after t observations for the Max-Ent mechanism, $\sigma = 2, 5, 10$.

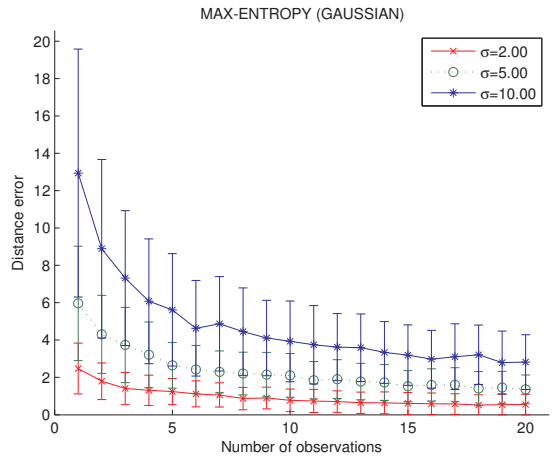

Figure 6: Distance errors after t observations for the Max-Ent mechanism, $\sigma = 2, 5, 10$. Error bars indicate $\pm$ one standard deviation from the mean.

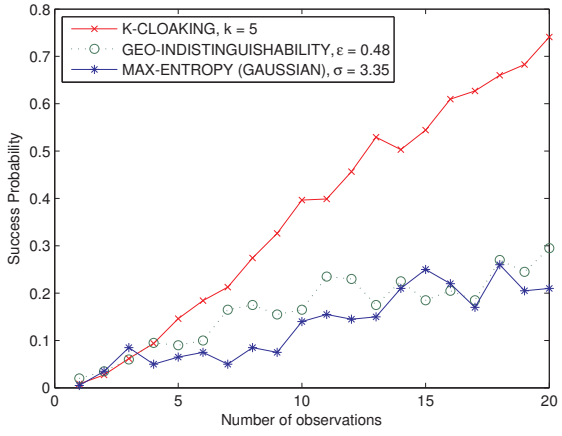

Figure 7: Success probabilities after t observations for all three mechanisms.

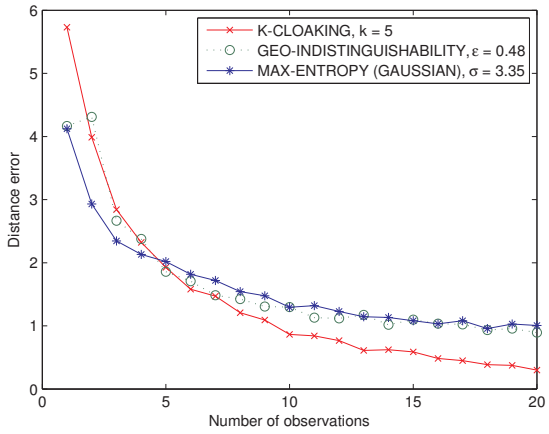

Figure 8: Distance errors after t observations for all three mechanisms.

is that early observations contribute more than later ones, i.e. we observe a pattern of diminishing returns with each additional observation that the attacker makes.

4. DISCUSSION OF RELATED WORK AND OF OTHER DEFENSES

Sending many queries from the same location (or, more generally, from nearby locations) is known in the location privacy community to cause problems for privacy (e.g. [3, 9]). However, it is a novel contribution of this paper to quantify precisely the effect of a given number of queries on privacy.

Regarding proposed defenses, this paper only considers privacy mechanisms that add independent random noise at each query. In future work, we aim to consider more types of defenses, and in particular the following: The privacy mechanism adds noise just once, thus selecting and reporting a single obfuscated location, and then keeps sending that obfuscated location in any future query that originates at the same true location. This defense is e.g. used by Fawaz and Shin [6], where the initial obfuscation is done with the mechanism proposed by Andrés et al. [1]. In this type of defense, there is obviously no point in estimating the attacker's success after $t > 1$ observations, because all observations after the first are identical, so no more information is disclosed. However, the probability distribution for the original obfuscation must still be designed in an optimal way, in order to minimize the attacker's success probability and maximize his distance error.

5. CONCLUSION

This paper presents an attack against location privacy that applies when the user sends repeated queries from the same location and the noise added by the obfuscation mechanism is a function only of the user's true location, which is a quite intuitive and pervasive assumption. The attack forms the maximum likelihood estimate of the user's location, and we find that this estimate improves very rapidly with just a few queries. The spatial k-cloaking mechanism does much worse than either Geo-Ind [1] or the new maximum-entropy mechanism Max-Ent that we propose in this paper, while Geo-Ind and Max-Ent perform approximately equally, with perhaps a small advantage for the latter. In future work, we plan to design a mechanism that optimally defends against this attack for a given noise budget. We also plan to design optimally the "one off" defense that generates a single obfuscated location and reports it in every future query that comes from the same true location.

6. REFERENCES

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

[2] K. Chatzikokolakis, M. E. Andrés, N. E. Bordenabe, and C. Palamidessi. Broadening the scope of differential privacy using metrics. In *PETS 2013*.

[3] K. Chatzikokolakis, C. Palamidessi, and M. Stronati. A predictive differentially-private mechanism for mobility traces. In *PETS 2014*.

[4] T. Cover, and J. Thomas. Elements of Information Theory. John Wiley and Sons, 2006.

[5] C. Dwork. Differential privacy. In *LNCS 4052*, 2006.

[6] K. Fawaz, and K. Shin. Location privacy protection for smartphone users. In *CCS 2014*.

[7] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux. Quantifying location privacy. In *SP 2011*.

[8] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec. Protecting location privacy: optimal strategy against localization attacks. In *CCS 2012*.

[9] G. Theodorakopoulos, R. Shokri, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec. Prolonging the hide-and-seek game: Optimal trajectory privacy for location-based services. In *WPES 2014*.