J. Cuellar, M. Ochoa, and R. Rios, "Indistinguishable Regions in Geographic Privacy", *Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC 2012)*, pp. 1463-1469, 2012.
http://doi.org/10.1145/2245276.2232010
NICS Lab. Publications: https://www.nics.uma.es/publications

Indistinguishable regions in Geographic Privacy

Jorge Cuellar
Siemens AG
Munich
jorge.cuellar@siemens.com

Martín Ochoa
Siemens AG
TU Dortmund
martin.ochoa@cs.tu-dortmund.de

Ruben Rios
Computer Science Department
University of Malaga
ruben@lcc.uma.es

Abstract

The ubiquity of positioning devices poses a natural security challenge: users want to take advantage of location-related services as well as social sharing of their position but at the same time have security concerns about how much information should be shared about their exact position. This paper discusses different location-privacy problems, their formalization and the novel notion of indistinguishability regions that allows one to proof that a given obfuscation function provides a good trade-off between location sharing and privacy.

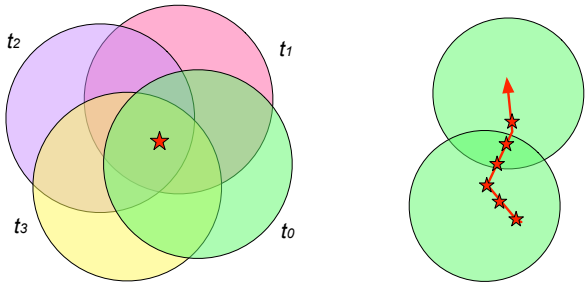
Keywords: Location privacy, obfuscation, indistinguishability

1 Introduction

The increasing growth in the number of positioning devices, readily present in low-end smartphones, introduces the opportunity to create countless services that take advantage of location information in order to provide a value-added service to the user. Clearly, these location-based services are tremendously beneficial not only to individuals but also to companies since their adoption imply a significant increase in revenue [6]. Also, the integration of these devices with social networks and location-sharing applications has seen rapid growth in the last few years giving rise to a multitude of geo-social networks [13] such as Foursquare, Twitter, or Facebook Places.

However, the location information of individuals is a very sensitive information that must be carefully protected. In the long term, sharing precise location information might reveal very diverse personal information, such as habits, home address, workplace, and even the health condition of a subject. As as matter of fact, users might want to have access to the attractive services offered by service providers but they might be reluctant to disclose precise location information. This leads to a potential definition of location privacy as the desire of individuals to determine when, how, and to what extend their location information is transferred to other entities [14].

Several countermeasures have been proposed to protect the privacy of individuals while accessing location-based services [1,3,12]. When the provision of the service requires the identity of the user, the common solution is to obfuscate his location, i.e. reduce its quality, by providing a region containing the actual location. Most of these solutions have concentrated on maximizing the utility of the location information while providing an adequate privacy protection level. Nevertheless, there are some challenging problems in obfuscation-based techniques that still need to be solved. In Fig. 1a we show how an individual eventually reveals its actual location, i.e. the star, although obfuscation regions are being reported. Suppose that the fixed position is the home address of the user, since the user will be visiting this location repeatedly ($t_0, \ldots, t_3$), the adversary will eventually obtain the exact or a very good approximation of the actual



(a) Fixed position (b) Mobility

Fig. 1: Typical problems in location obfuscation algorithms.

location by intersecting the reported regions. Additionally, users might need to report their location while they move from one location to another and, during the movement, their location information must be updated. However, simply providing a new obfuscation region when the previous one is no longer valid, i.e. at bordering areas, will provide the attacker a great amount of information (see Fig. 1b). This problem is magnified when the user follows the same route regularly.

The goal of this work is to formalize some natural notions that arise in the context of the described problem and that have been so far, to the best of our knowledge, not treated rigorously. In particular we consider different scenarios where an attacker could gain more information than initially intended from his observations of a shared position over time, should the obfuscation function not be carefully designed. Our main results are proofs about the existence/non-existence of certain obfuscation functions under different attack scenarios and attacker models.

The following sections are structured as follows: Section 2 defines what we mean by *obfuscation function*, Sec. 3 discusses several games where an adversary tries to guess the position of a user with high precision whereas Sec. 4 deals with indistinguishability notions due to the user *displacement*. Section 5 discusses related work and we conclude with Sec. 6.

2 Obfuscation functions

In this section we introduce the concept of obfuscation function in a precise way, which is central to our discussions in the following sections. Prior to the definition of this important concept we will present an assumption on the precision of the location measurement.

Positioning devices are not accurate enough to provide an exact geodesic point, instead they usually provide an imprecise region where the actual location of the device is contained. We assume that the space can be divided into areas given by the maximum precision of the positioning device (see the dotted grid in Fig. 2). That is, we assume there is a fixed *measurement grid*. Therefore, any geodesic point is mapped to a single minimal area but a minimal area contains an infinite number of points. The position of the device is given by the area which contains its precise location and, consequently, retrieving this area is equivalent to obtaining the precise location. In the rest of this paper when we refer to the exact or real position of the user we are actually referring to the measurement position. More precisely, when we use the coordinates (x, y) we mean the measurement area containing (x, y) which is unique and independent of time.

Given the previous assumption on the measurement grid we can now define the concept of obfuscation function. Formally:

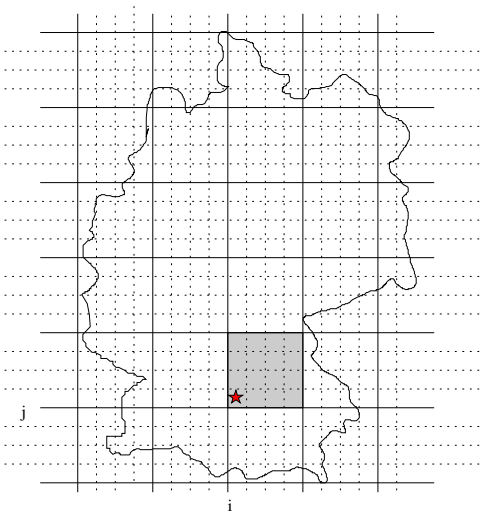


Fig. 2: Example of obfuscation function based on a lattice.

Definition 1. *Let $M = \{M_0, M_1, \dots\}$ be a partition of $\mathbb{R}^2$ (that is $\bigcup_i M_i = \mathbb{R}^2$). Let $d \in P \subset \mathbb{R}$ be a degree of accuracy the subject would like to have for the disclosed information about his location*[1]*. We say f is an obfuscation function:*

$$f : M \times P \times T \to \mathcal{G}$$

where $\mathcal{G}$ is the set of connected and convex sets of the obfuscated coordinates, such that $\bigcup_{G \in \mathcal{G}} = \mathbb{R}^2$ and T is a time coordinate.

We now give a very natural example of obfuscation function based on a *grid*.

Example 1. *Let M be a measurement grid. Let $\mathcal{G}$ be the sets resulting of a tessellation of the plane by squares of side $2d$ for $d \in \mathbb{N}$ that we call obfuscation squares or regions. Then we define f as the function returning the obfuscation square containing the original position (x, y), as depicted in Fig. 2. Formally:*

$$f((x, y), d, t) = (i, j)$$

such that $i \cdot 2d + r = x$ with $|r| < 2d$ and $j \cdot 2d + r' = y$ with $|r'| < 2d$.

Each $(i, j) \in \mathbb{Z}^2$ corresponds to a square of the tessellation. Note that f is independent of time.

Now recall the example on Fig. 1a where, for different requests in time, we get an obfuscated region containing our exact position but with a random center every time. This method was discussed in the IETF[2] geopriv mailing list. It can be formalized as follows:

Example 2. *Given the original position (x, y), the obfuscation function g returns, for every request, a disc of radius d which is displaced by a random vector no larger than d such that the original position is contained in the disc B_d. Formally:*

$$g((x, y), d, t) = B_d(r(t))$$

where $r(t)$ is a random function of time such that $(x, y) \in B_d(r)$.

[1] Although in principle d could be any reasonable parameter defining an area, for example a number of habitants, in this work we think of d as a physical distance typically defining an obfuscation area.

[2] The Internet Engineering Task Force `http://www.ietf.org/`

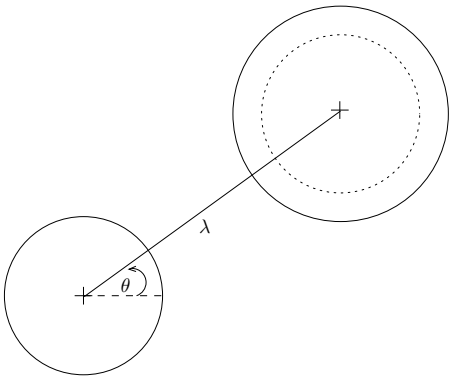


Fig. 3: Composition of shift and enlarge

A more sophisticated alternative was finally proposed as an IETF geopriv draft [12]. In this case, also random discs are returned but the main difference is that the randomness is actually dependent on the actual position and a secret hold by the user. Therefore, the same disc is provided for a given location regardless of the moment in time. We can formalize it as follows:

Example 3. *The obfuscation function h returns, a disc of radius d which is displaced by a vector dependent on the actual position (x, y) and a secret key K such that the original position is contained in the disc B_d. Formally:*

$$h((x, y), d, t) = B_d(H_K(x, y))$$

where $H_K(x, y)$ is a deterministic function of the position and the key K such that $(x, y) \in B_d(H_K)$.

Recently, the composition of obfuscation functions was proposed in [1]. The operators under consideration are the shifting of the measurement center and the distance enlargement and reduction of the radius. Formally:

Example 4. *Let* $\mathsf{shift}((x, y), d, t)$ *the obfuscation function that shifts the measurement region by a distance λ, which depends on the obfuscation parameter d and the measurement precision, and a random angle θ depending on the time. The* enlarge *and* reduce *operators simply enlarge and reduce respectively the accuracy of the original measurement precision. The composition of two obfuscation functions has the form:*

$$c : M \to M' \to \mathcal{G}$$

where M is an original measurement region, M' is an intermediate obfuscation region and $\mathcal{G}$ is the resulting obfuscation region. Formally:

$$c = f \circ g = f(g((x, y), d, t), d, t)$$

For example, Fig. 3 depicts a composition of shift *and* enlarge.

3 Indistinguishability

In the previous section we have defined obfuscation functions in general, now we introduce some privacy notions for different models of users and adversaries. They are all based in the concept

of *indistinguishability* that captures the intuition that the real position of the user should be, in the eyes of an adversary, indistinguishable from a set of possible positions (this set depending on the obfuscation accuracy parameter). Nevertheless, in order to take advantage of location related services and to share our position with friends and relatives in a 'honest' way, we require this obfuscated set to contain the real position of the user.

3.1 Indistinguishability properties

To be able to be more accurate in our discussion about the security implications of sharing the location to third parties, we consider several attack scenarios. These scenarios encompass situations where the user gives away sensitive information about himself by reporting his current location. We pay special attention to frequently visited places because this provides personal information such as political affiliation, religious condition, and even the work or home address.

The first scenario is as follows: assume an attacker queries a user position at regular intervals (for example, everyday when he is at home, or every Friday night, etc.). From this collection of observations he should not be able to increase the precision of his knowledge on the real position up to some defined and fixed parameter chosen by the user. That is, if the user decides he wants to obfuscate his position up to a radius of 5 Kms, then an adversary should not be able to reduce this area any further by making multiple location requests.

Definition 2 (Fixed-position-indistinguishability)**.** *Consider the following game between a user $\mathcal{U}$ and an adversary $\mathcal{A}$:*

- *$\mathcal{U}$ is in a fixed position $\bar{x}$. $\mathcal{A}$ sends n position requests to $\mathcal{U}$.*
- *To each position request i, $\mathcal{U}$ sends back the coordinates G_i such that $G_i = f(\bar{x}, d, i)$.*
- *For each request $\mathcal{A}$ guesses the original value of $\bar{x}$. This guess is a coordinate y_i.*

We say that f satisfies fixed-position-indistinguishability if the probability of $\mathcal{A}$ guessing the actual position $\bar{x}$ after n position requests is bounded by a constant depending on the accuracy parameter d:

$$\mathbb{P}(y_n = \bar{x}) \leq C(d)$$

The second scenario is the case when the adversary knows that a user has departed from a certain position and he tries to conclude his final destination. It is desirable that this final conclusion is not precise, again up to some fixed parameter. In other words, an adversary should not be capable of determining the destination of a user from his original position and his location updates on the route.

Definition 3 (Destination-indistinguishability)**.** *Consider the following game between a user $\mathcal{U}$ and an adversary $\mathcal{A}$:*

- *$\mathcal{U}$ begins a journey at point x_0 and wants to reach a goal destination $\bar{x}$. $\mathcal{A}$ sends n position requests to $\mathcal{U}$ over a lapse of time.*
- *To each position request i, $\mathcal{U}$ sends back the obfuscated coordinates such that $G_i = f(x_i, d, i)$ where x_i is his real position at time i.*
- *After each request $\mathcal{A}$ guesses the actual value x_i. This guess is the coordinate y_i.*

We say that f satisfies destination-indistinguishability if:

$$\mathbb{P}(y_n = \bar{x}) \leq C(d)$$

Now consider the following scenario. An attacker is able to obtain (by physically watching or other means) the real position of a user at a given point in time. Afterwards, the adversary only observes the obfuscated position disclosed by the user. It is desirable that the adversary is not able to deduce that the user has returned to the initial position just by observing the obfuscated coordinates. For example, if there is a one-to-one correspondence between the real position and the obfuscated coordinates, then the adversary could immediately conclude the real position of a user if he has ever gain possession of the pair real position - obfuscated position.

Definition 4 (Known-position-indistinguishability)**.** *Consider the following game between a user $\mathcal{U}$ and an adversary $\mathcal{A}$. An oracle $O(x, d, t)$ gives back the value of $f(x, d, t)$ to an adversary.*

- *The adversary $\mathcal{A}$ runs m queries on an oracle O at fixed time t by choosing m locations $\bar{x_j}$. That is: $\bar{G_i} = O(\bar{x_i}, d, t) = f(\bar{x_i}, d, t)$ for $1 \leq i \leq m$.*
- *$\mathcal{A}$ makes arbitrary location requests to $\mathcal{U}$, with time always greater than t. We denote the real position of the user at time t' by $x_{t'}$.*

We say that f satisfies known-position-indistinguishability if:

$$\forall t' > t \ \mathbb{P}(x_{t'} = \bar{x_j} | f(x_{t'}, d, t') = \bar{G_j}) \leq C(d).$$

3.2 Indistinguishable regions and honesty

So far, we have made no assumptions on the obfuscation functions beyond Def. 1. In the following we give two additional characterizations of obfuscation functions that will be useful in the following. As already mentioned in the introduction of this section, the first is a reasonable requirement on a function in order to take advantage of location related services and to share our position with friends and relatives:

Definition 5. *The obfuscation function f is honest if:*

$$x \in f(x, d, t)$$

The second and central definition formalizes an implicit necessary condition for an obfuscation function to pass any of our indistinguishability definitions so far: the obfuscated sets returned to an adversary must be constant for all points contained within those sets. Formally:

Definition 6. *Given an obfuscation function f, we say a region N of the plane is an indistinguishability region if:*

$$\forall \ t \in T \ \forall \ x, y \in N \ \ f(x, d, t) = f(y, d, t)$$

We will compare the examples of obfuscation functions introduced in Sec. 2 based on this two definitions in the next subsection.

	Fixed-Position	Destination	Known-Position
Example 1	$\surd$	$\surd$	$\surd$
Example 2	$\times$	$\times$	$\surd$
Example 3	$\surd$	$\surd$	$\times$
Example 4	$\times$	$\times$	$\surd$

Table 1: Indistinguishability properties on Examples.

3.3 Satisfability

In this subsection we compare some of the examples of obfuscation functions introduced in the previous section with respect to the indistinguishability properties we have defined so far.

Theorem 1. *If all possible positions of x are equally distributed over $\mathbb{R}^2$ and $\mathcal{G}$ are indistinguishable regions for a honest obfuscation function f, then f is both destination and fixed-position indistinguishable*

Proof. We assume that all points in $\mathbb{R}^2$ are equally probable to be visited by $\mathcal{U}$. Fixed-position follows immediately, since x_n is equally likely to be any point of $G_n = f(x, d, n)$. Therefore, an adversary $\mathcal{A}$ observing the coordinates G_n returned by the user $\mathcal{U}$ gains information about the location of the user with no further precision than d. The same holds for destination-indistinguishability, since once the user enters the region corresponding to destination $\bar{x}$ he is likely to be in any point of this region. □

It is easy to see that Example 1 is a honest function with $\mathcal{G}$ indistinguishable, therefore it is straight-forward that all the properties hold. More precisely, for a given position the function returns always the same value G regardless of time. Similarly, when the user reaches his destination, since he continually reports the same obfuscation region, the adversary is not able to determine in which precise location within the obfuscated region the user is finally located. The known-position-indistinguishability property also holds because the next time the user visits that location the adversary has no certainty about whether he came back to the same location or he is anywhere else within that region.

Besides, Example 2 and 4 violate both fixed- and destination-indistinguishability. In the first case, the reason is that by performing enough observations the position can be arbitrarily reduced by performing the intersections of the reported locations. In the latter case, it is possible to approximate the original position also with a sufficient number of observations. If the shifting displaces the measurement a constant λ, the expected value of the original position is easy to calculate. Also, if the value of λ is not constant it would be possible to determine its maximum value by using the centers and, consequently, the original position. In fact, the next operator after the shift provides no more uncertainty because to obtain the expected value it is sufficient to know the centers of the final obfuscation region. On the contrary, both Example 2 and 4 hold the known-position-indistinguishability property since they provide different obfuscation regions in time.

Finally, Example 3 holds both fixed- and destination-indistinguishability because the obfuscation region reported is independent of time and it is based on the actual location and a secret key. However, this approach violates the known-position-indistinguishability property since once an obfuscation region is known for a given position, if the adversary observes the same region, the adversary will know with no uncertainty where the user is located.

4 Displacement obfuscation

In this section we discuss the location privacy problem in the presence of user displacement. While moving from one location to another, the user reports the regions he traverses. For some scenarios, a user might not only want to prevent the adversary from determining the final destination (as already captured by Def. 3) but also to cover the location information on the route. For example, the non-disclosure of precise location information during user displacement might prevent physical harassment. We formalize this as follows:

Definition 7 (Displacement-indistinguishability)**.** *Consider the following game between a user $\mathcal{U}$ and an adversary $\mathcal{A}$:*

- *$\mathcal{U}$ begins a journey at public point x_0 and wants to reach a goal destination $\bar{x}$. $\mathcal{A}$ sends n position requests to $\mathcal{U}$ over a lapse of time.*
- *To each position request i, $\mathcal{U}$ sends back the obfuscated coordinates such that $G_i = f(x_i, d, i)$, where x_i is his real position at time i.*
- *After each request $\mathcal{A}$ guesses the actual value x_i. We say this guess is a coordinate y_i.*

We say that f satisfies displacement-indistinguishability if:

$$\forall\ i\ 1 \leq i \leq n\ \mathbb{P}(y_i = x_i) \leq C(d)$$

In other words, the obfuscation function should not allow an adversary to determine the location of the user with probability higher than $C(d)$ for any given position response G_i.

Definition 8 (Past-disp-ind)**.** *Consider the same game as in displacement-indistinguishability, we say that f satisfies past-disp-ind if:*

$$\forall\ i\ 1 \leq i \leq n\ \forall\ j \leq i\ \mathbb{P}(y_j^i = x_j) \leq C(d)$$

where by y_j^i we indicate the guess the adversary makes of x_j at time i.

This is a stronger property since it does not only prevent the adversary from determining present but also previous coordinates given the current knowledge. Such an obfuscation function reveals no information even if the adversary, at step i, attempts to recompute the values of x_j for $1 \leq j \leq i$.

4.1 Attacker and User models

In this scenario where the user is willing to move while a certain degree of uncertainty about his precise location is retained, it is necessary to clearly define what are the assumptions made on both the user and the adversary side. In particular, it is important to define the capabilities of the user in terms of the movement and the adversary in terms of the number of location responses he might obtain.

We consider that the user is able to move at various speeds. In the first place we will discuss about a user with the ability to travel at infinite speed. Additionally, we present a user whose speed is limited to a maximum value although he might decide not to reach it. Moreover, we consider that the user might travel in any direction and make a turn at any moment.

Besides, in terms of the frequency of location requests we discuss two potential cases. In the first case, the attacker might choose the number and frequency of location requests performed and the user will honestly reply to all of them. Secondly, we consider a case in which the user chooses the frequency of location replies to be sent to the attacker. This is equivalent to allowing the user decide when to inform the attacker about his location.

4.2 Satisfiability

Previous considerations on the capabilities of the attacker and the user lead to several application cases that we analyze for certain obfuscation functions. We will concentrate on tessellations of the plane (e.g., Example 1) since they preserve all the properties defined in Sec. 3.1.

Lemma 1. *Consider the case where the attacker $\mathcal{A}$ and the user $\mathcal{U}$ have the following properties:*

- *$\mathcal{A}$ chooses a requesting frequency $freq = \varepsilon > 0$*
- *$\mathcal{U}$ can travel at speed $v = \infty$*

We say that the attacker $\mathcal{A}$ is unable to determine the location of the user with no further precision than $C(d)$.

Proof. Since all points over M and $\mathcal{G}$ are equally distributed and the user $\mathcal{U}$ is able to reach any location in the plane in a time $\varepsilon > 0$, the adversary $\mathcal{A}$ gains no further information than the region G_i returned by the user $\mathcal{U}$. No matter how close is ε to 0, the user might be located at any position.

□

However, we admit that considering a user moving at an infinite speed is quite unrealistic. Therefore, the following cases present scenarios where the speed of the user is upper bounded.

Lemma 2. *Consider the case where the attacker $\mathcal{A}$ and the user $\mathcal{U}$ have the following properties:*

- *$\mathcal{A}$ chooses a requesting frequency $freq = \varepsilon > 0$*
- *$\mathcal{U}$ travels at speed $v \leq v_M$*

We say that f does not satisfy the displacement indistinguishability property because $\mathcal{A}$ is able to reduce its uncertainty about the location of $\mathcal{U}$ for some location requests.

Proof. It is sufficient to demonstrate that $\exists\ i\ \mathbb{P}(y_i = x_i) > C(d)$. This is equivalent to say that the adversary is capable of reducing the obfuscation area G_i provided by the user at some point. Consider an obfuscation function that tessellates the plane by squares of side d for $d \in \mathbb{N}$ and whose $C(d) = 1/d^2$. Provided that the speed of $\mathcal{U}$ is bounded by v_M, the maximum area covered by the user is given by $\pi \cdot d_M$, where $d_M = v_M \cdot freq$. Consequently, the attacker is able to reduce the location uncertainty $C(d)$ when the user enters a new obfuscation region. Without loss of generality, the adversary can choose a frequency $0 < freq' < freq$ such that $d'_M < d_M$ exposing the user at the borders, as depicted in Fig. 4. In particular, at any border, the attacker can choose a frequency $freq'$ such that the area $d'_M \cdot d < d^2$, which implies that $\mathbb{P}(y_i = x_i) = \frac{1}{d'_M \cdot d} > \frac{1}{d^2} = C(d)$. □

Theorem 2. *If $\mathcal{A}$ chooses $freq$ and $\mathcal{U}$ moves at $v \leq v_M$ then there is no obfuscation function satisfying displacement indistinguishability.*

Proof. Let M be the set of all possible coordinates and $\mathcal{G}$ the set of obfuscated coordinates, such that $M = \bigcup_{G \in \mathcal{G}}$ and $\forall G \in \mathcal{G}\ G \cap M \neq M$. Also, let f be an honest obfuscation function in the sense that $\forall x \in f(x, d, t)$. It may happen that $\forall i, j\ G_i \cap G_j = \emptyset$, that is, all regions are disjoint. Indeed, this is a general situation of Case 2 and the proof is direct when the user enters a new region. Additionally, it may happen that the obfuscation function returns overlapping regions, which is equivalent to say that $\exists i, j\ G_i \cap G_j \neq \emptyset$. In this case, the user will be at some point in time in a position $x \in G_i \cap G_j$ and might behave in two different ways:

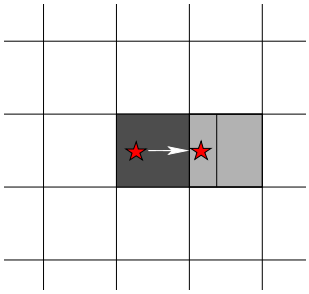


Fig. 4: The adversary can estimate the new position with better precision than intended by the user

- The user always returns either G_i or G_j. Since f is honest and $\forall G \in \mathcal{G}\ G \cap M \neq M$, the user will report G_j when he leaves the overlapping region. As a direct consequence of Case 2, the adversary is able to reduce the uncertainty region.
- The user returns $G \xleftarrow{D} \{G_i, Gj\}$ based on some probability distribution D. Given a sufficient number of samples, which can be obtained by reducing the requesting frequency $freq$, the attacker is capable of determining the distribution and thus reduce the uncertainty region.

□

Lemma 3. *Consider the case where the attacker $\mathcal{A}$ and the user $\mathcal{U}$ have the following properties:*

- *$\mathcal{U}$ chooses a replying frequency $freq = \varepsilon > 0$*
- *$\mathcal{U}$ travels at speed $v \leq v_M$*

then there exists a function f such that displacement-ind and past-displacement-ind hold.

Proof. This case presents several situations where the user might trade-off between speed and location privacy. In any case, since the speed of the user is bounded by v_M, the area covered during his movement is also bounded by $A \leq \pi \cdot (v_M \cdot freq)^2$, as depicted by the circle in Fig. 5. Therefore, the user might adapt either its frequency $freq > 0$ or speed $v < v_M$ such that $A \cap (G_i \cup G_{i+1}) \geq 1/C(d)$, where Gi is the i-th reported region and G_{i+1} is the subsequent reported region. That is, he never leaves the gray region adjacent to the initial square in the above mentioned figure. Therefore, the adversary can not infer in which point of the reported obfuscation region the user actually is after time $freq$. The user can repeat this for all subsequent location reports. The trade-off between privacy and speed is evident since the user could either:

- By default go as further as possible within the allowed range
- Slow down and reach a random point within the wished reported obfuscation region

If the user chooses the first option then past-disp-ind does not hold after two steps or more if we assume that the adversary is aware of this strategy since he knows the user will be in a position near a border in the obfuscation region. In the second case, both displacement-ind and past-disp-ind hold because of the randomness of the choice at the reported time $freq$.

□

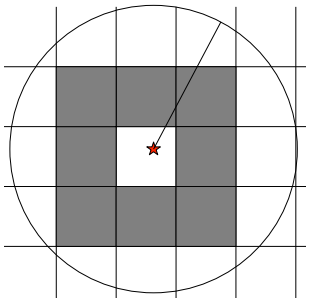


Fig. 5: The user can control the frequency to preserve displacement-indistinguishability

It is interesting to notice that an even better trade-off between privacy and speed could be achieved if the user is **not** honest about his current position at the time of the location request. It will be matter of future work to formalize this possibility and to compare it with our current definitions.

5 Related Work

Several techniques have been proposed to protect the privacy of the users when accessing location-based services. These might be categorized into policy-based, anonymity-based and obfuscation-based techniques. The first group of solutions consider regulations, policies and access control mechanisms to protect the location information of the user [2, 9, 10]. There are two main limitations to this approach. First, the definition of accurate policies is a difficult task for inexperienced users, while system-supplied policies are usually simply accepted by the users without considering future implications. Second, whenever the location information leaves the domain of the user, any entity accessing that information might potentially make improper use of it without the user even noticing. Additionally, though the service provider makes no fraudulent use of individuals' location information, a break into the system might still compromise the privacy of the user. Therefore, a better privacy protection requires the use of computational mechanisms capable of obscuring location information to a certain degree before being finally delivered.

Anonymity-based solutions aim to hide the real identity of the user when sharing his location information with other entities. This type of solutions are suitable when the real identity of the user is irrelevant for the provision of the service, for example, if the user wants to know which is the closest restaurant to his current location. Usually, these solutions make use of periodically changing pseudonyms, which aims to prevent user tracking. They rely on mix zones [3] or highly-populated areas [11], where the pseudonyms of the users are modified and thus the identities of the users are somehow mixed. Another set of solutions are based on the concept of k-anonymity [8] and alter the user location in such a way that there are at least $k-1$ other individuals with similar features in the same geographical region to prevent re-identification. Proposed techniques differ in the way regions are computed. Basically, two main groups of solutions are considered, those that rely on a trusted authority [5], and those in which users collaborate to anonymize their location [4].

On the contrary, obfuscation-based techniques are devised for protecting location information instead of identity information. These solutions are practical when the user is willing to share its current location but is reluctant to provide precise location information. For example, a user might want to share that he is visiting a certain city but not where he is exactly located. Therefore, obfuscation-based solutions are intended to reduce the quality of the location mea-

surements being shared by the user. This degradation can be achieved in different ways, for example, by adding noise or generalizing the actual measurements, as proposed in [7]. However, the techniques more closely related to our research are those based on the provision of an obfuscation region (e.g. a disc of 10 km), which contains the actual location of the user. Three obfuscation operators (radius enlargement, radius reduction, and center shifting) as well as the composition of these operators are analyzed in [1]. Also, a recent IETF draft [12] presents an obfuscation algorithm that reports a disc of some radius that is shifted from the actual location a factor that is calculated based on a secret key held by the user. Moreover, previously mentioned schemes based on the concept of k-anonymity can be also regarded as a form of obfuscation because they report a cloaking region containing the actual location of the user but they present different objectives.

6 Conclusions

In this work we have formalized a number of notions of geographic location privacy that allow us to prove results on an important class of obfuscation functions: honest functions with indistinguishability regions. We have seen how this notions are satisfied under different assumptions about the adversary capacity and knowledge. This is a promising first step towards a better understanding of obfuscation functions under scenarios with weaker assumptions.

In future work we plan to study more realistic scenarios, where the user moves for example on roads or following certain routes, and where the destinations are not all equally likely to be visited within a region. Moreover, we will investigate what are the implications of using dishonest obfuscation functions in the provision of practical location-based services as well as their impact on the privacy protection level.

Acknowledgements

This work has been partially funded by the European Commission through the FP7 project NESSoS (FP7 256890) and the FP7–ICT project WebSand (FP7 256964). Additionally the paper has been supported by the MoDelSec Project of the DFG Priority Programme 1496 "Reliably Secure Software Systems – RS^3". The last author is supported by the Spanish Ministry of Education through the National F.P.U. Program.

References

[1] C. A. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati. An Obfuscation-Based Approach for Protecting Location Privacy. *IEEE Trans. Dependable Secur. Comput.*, 8:13–27, January 2011.

[2] R. Barnes, M. Thomson, J. Winterbottom, and H. Tschofenig. Location Configuration Extensions for Policy Management. http://www.ietf.org/id/draft-ietf-geopriv-policy-uri-01.txt, June 2011.

[3] A. R. Beresford and F. Stajano. Location Privacy in Pervasive Computing. *IEEE Pervasive Computing*, 02(1):46–55, 2003.

[4] G. Ghinita, P. Kalnis, and S. Skiadopoulos. PRIVE: Anonymous Location-Based Queries in Distributed Mobile Systems. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 371–380, New York, NY, USA, 2007. ACM.

[5] 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, pages 31–42, New York, NY, USA, 2003. ACM.

[6] R. Kim. Location-based services worth $10B by 2016. http://gigaom.com/2011/06/09/location-based-services-worth-10b-by-2016/, June 2011.

[7] J. Krumm. Inference Attacks on Location Tracks. In *Proceedings of the 5th international conference on Pervasive computing*, PERVASIVE'07, pages 127–143, Berlin, Heidelberg, 2007. Springer-Verlag.

[8] P. Samarati and L. Sweeney. Protecting Privacy when Disclosing Information: k-Anonymity and its Enforcement through Generalization and Suppression. In *Proceedings of the IEEE Symposium on Research in Security and Privacy (S&P)*, Oakland, CA, May 1998.

[9] H. Schulzrinne, H. Tschofenig, J. Morris, J. Cuellar, J. Polk, and J. Rosenberg. Common Policy: A Document Format for Expressing Privacy Preferences. http://tools.ietf.org/html/rfc4745, Feb 2007.

[10] Senator Edwards, J. Location Privacy Protection Act of 2001. http://www.techlawjournal.com/cong107/privacy/location/s1164is.asp, July 2001.

[11] J.-H. Song, V. W. Wong, and V. C. Leung. Wireless Location Privacy Protection in Vehicular Ad-Hoc Networks. *Mob. Netw. Appl.*, 15:160–171, February 2010.

[12] M. Thomson. Obscuring Location. http://tools.ietf.org/html/draft-thomson-geopriv-location-obscuring-03, June 2011.

[13] C. R. Vicente, D. Freni, C. Bettini, and C. S. Jensen. Location-Related Privacy in Geo-Social Networks. *IEEE Internet Computing*, 15:20–27, 2011.

[14] A. F. Westin. *Privacy and Freedom*. New York Atheneum, 1 edition, 1967.