INGENIERÍA Y DESARROLLO

ARTÍCULO DE REVISIÓN / REVIEW ARTICLE

A Survey on Privacy in Location-Based Services

Revisión en privacidad en servicios basados en localización

Mayra Zurbarán*
Liliana González**
Pedro Wightman Rojas***
Universidad del Norte (Colombia)
M. Labrador****
University of South Florida

* Receiveda Bachelor degree in Computer Science from Universidad del Norte. She is currently Junior Researcher and a student of the Computer Science Doctorate Program at Universidad del Norte. *mzurbaran@uninorte.edu.co*

**Received a Bachelor degree in Computer Science from Universidad del Norte. She is currently Junior Researcher and a student of the Computer Science Master Program at Universidad del Norte. *lghernandez@uninorte.edu.co*

***Received a Ph.D. degree in Computer Science from the University of South Florida. He is currently a Senior Researcher, professor and Director of the Department of Computer Science at Universidad del Norte. *pwightman@uninorte.edu.co*

****Received the M.S. in Telecomunications and the Ph.D degree in Information Science with concentration in Telecomunications from University of Pittsburg. Since 2001 has worked with the University of South Florida where he is currently an Associate Professor in the department of Computer Science and Graduate Program Director. *mlabrador@usf.edu*

Correspondencia: Pedro M. Wightman. Universidad del Norte, km 5, vía a Puerto Colombia. Of. 4L6. Tel. 3095095 – 3010.

Volumen 32, n.º 2
Julio-diciembre, 2014
ISSN: 0122-3461 (*impreso*)
2145-9371 (*on line*)

Abstract

Location services have become popular over the last years due to the global adoption of smartphones and the worldwide availability of the Global Positioning System (GPS) and other positioning methods. Location-based services (LBSs) offer relevant information to users based on their location. Some common applications of LBSs are traffic or public transportation information, search of points of interest (restaurants, stores, etc.), navigation, among others. Despite all the desirable features that these services provide, most of them do not provide adequate protection of the geographical location of the users, putting them at risk if their information falls in wrong hands. This paper presents a compendium of techniques to protect the location privacy of the users, and introduces an approach to compare and evaluate the presented mechanisms and their viability to be used in different kinds of LBSs.

Keywords: location obfuscation, location privacy, location tracking, private information retrieval, points of interest search.

Fecha de recepción: 13 de febrero de 2014
Fecha de aceptación: 30 de mayo de 2014

Resumen

Los servicios de localización se han popularizado en los últimos años debido a la adopción global de teléfonos inteligentes y la disponibilidad a nivel mundial del Sistema de Posicionamiento Global (GPS) y otros métodos de posicionamiento. Los servicios de localización (LBSs) ofrecen información relevante para los usuarios en función de su ubicación. Algunas aplicaciones comunes de LBSs son el tráfico o información de transporte público, la búsqueda de puntos de interés (restaurantes, tiendas, etc.), la navegación, entre otros. A pesar de todas las características deseables que estos servicios prestan, la mayoría de ellos no ofrece una protección adecuada de la ubicación geográfica de los usuarios, lo que los pone en riesgo si la información llega a manos equivocadas. En este trabajo se presenta un compendio de técnicas para proteger la privacidad de localización de los usuarios, y una matriz de valoración para evaluar los mecanismos presentados y su viabilidad para ser utilizados en diferentes tipos de LBSs.

Palabras clave: búsqueda de puntos de interés, intercambio privado de información, ofuscación de localización, privacidad de localización, rastreo de localización.

1. INTRODUCTION

During the last few years, the market associated with mobile technology has grown at an impressive rate, becoming attractive to all the actors involved: manufacturers, operators, governments and research centers, due to the massive adoption of this technology, the improved computing power of new devices, commercialization opportunities not only for traditional voice services but for more sophisticated applications that allow interaction at a higher level for people and with information on the Internet. This evolution of the mobile communication has turned cell phones into the essential way people communicate daily, and users are constantly demanding for new applications that suit their needs.

Apart from the improved computing and communication capabilities of these phones, one of the main advantages that they present compared to other devices like laptops, is that they usually have integrated the possibility to calculate their geographical position, either via Global Positioning System (GPS), or other technologies like WiFi-based or Cell tower-based location; even though laptops have the alternative of adding a USB GPS device or by using WiFi routing information as well, the implicit integration of these technologies into smartphones, their portability, connectivity and their personal nature; increased the exploitation potential through the latest mobile applications in order to offer a personalized service improved by context.

These applications are called Location Based Information Systems (LBIS). LBIS are defined as "Applications that provide users with information based on their geographical position, which could be obtained from the mobile device they are accessing the service, or using a manually defined location"[26]. The origin of LBS was the E911 (Enhanced 911) in the United States in 1996, it required the mobile operators to locate the callers of the emergency line with prescribed accuracy [3]. In order for LBSs to provide the requested information properly, sensitive data about the subject's location is required. While this private information is sent unprotected from the mobile device, it is in danger of being intercepted and misused by un trusted third parties and even by the LBIS itself. Location privacy violation attacks can include targeted spamming, stalking, physical assaults, fraud, robbing, kidnapping, etc. Attacks are not limited to the use of the current location of the individual, there are also prediction attacks that can infer where a person is going to be based on the intercepted information from GPSs. Location information can give attackers the opportunity of physically harming victims, also there are places where the mere fact of revealing a subject's permanence there could give away too much information, such as workplace, home, hospitals, rehab centers, jails, church, political centers, etc.

It is known that governments may obtain telephone records and location information of persons involved in judicial acts. However if this information is accessed without restrains, it may be used improperly. With this in mind, aiming to protect citizens privacy, governments became interested and raised concerns about the adequate protocols to handle these communications[16]; in EU, the European Union Directive on Privacy and Electronic Communications [17] specifically defines location information, user consent requirements, and corporate disposal requirements. This provides EU citizens an explicitly stated, protected right to the privacy of their location information. In Colombia there is the law Ley 1581 [30] that forces entities that treat personal data to notify their users to what extent their data will be stored or manipulated. In the US there is a project that specifically aims to provide location privacy to citizens when subscribing to an LBIS, the Location Privacy Act was presented in 2011 and was passed by the Senate in late 2012 [6].

To provide location privacy while making use of LBIS, the use of a protection mechanism is necessary. Many methods have been introduced in the literature, however very few have been implemented in commercial applications. Location privacy has been defined by [11] as: "A special type of information privacy which concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others…"

There are three aspects to location information: identity, location and time. If an adversary is able to link between them, location privacy is broken. Historical location data is also important since it allows establishing behavior patterns and possibly identifying user's home, work and usually frequented places.

Privacy itself is a complex subject, one of the first definitions to this term came earlier on the 19th century by Louis Brandeis and is actually quite simple: "The right to be let alone"[46]. Naturally, this meaning has evolved along with humanity and technology, but the essence still remains. It is a day-by-day challenge to maintain the balance within the developed technology and available privacy measures. In [34] propose a conceptual framework for everyday privacy in ubiquitous computing.

There are diverse approaches to satisfy LBIS Privacy requirements; some of them are designed to protect user´s identity while issuing queries, others focus on protecting specifically the user's location and some offer protocols to obfuscate queries as well, in [51] is presented a framework to provide query privacy specifically. Access-control is also a way of protecting users from undesired context requests of the applications at specific events, however it is not always admissible to dispense with the service as described in [40].

An ideal approach for providing location privacy would provide statistics on the LBIS users' behavior and at the same time protect each individual identity and location information (figure 1). Statistics are important in order to ensure improvements on the service allowing techniques of Am I (Ambient Intelligence) and bring better user experience to the application.

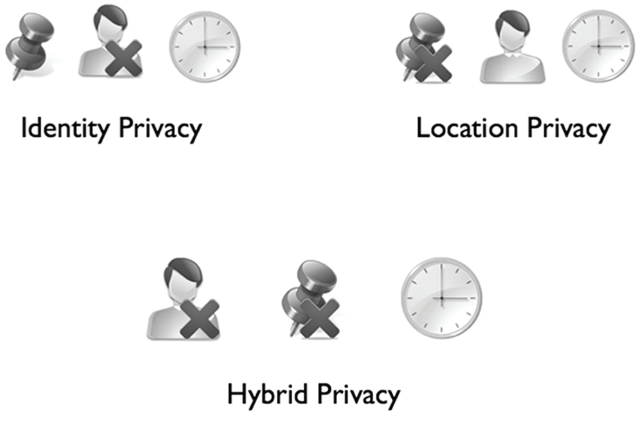


Figure 1. Types of privacy in LBIS

On figures 1 and 2 it is shown how one component can damage the quality of information gathered on an LBIS and give perspective of what would be available if undesired communication of this happens, for a no protection scheme, an attacker would be able to construct the user's full path (figure 2a). In another scenario, suppose an attacker gains access to identity and time but has no knowledge of what places the user visited (figure 2c), she could infer very little since there is no context information about that user in the system, this corresponds to Location Privacy. Also, given the case where the identity is unknown (Identity Privacy) but location and time are specified (figure 2b), the obtained location information could help infer who the person is through matching on a directory and serve as a prediction base for a user's weekly routine. A different alternative shown on Figure 2d is when an attacker has knowledge of the identity of the subject and the places that were visited, but the timestamps of that information is blurred or not available. In this case the attacker can construct a behavior pattern for that user and analyze the information provided to place those events on a feasible yet not exact time interval.

There are many definitions of identity, the one we will refer to in this paper is known as Idem Identity or the Diachronic Meaning of Identity which is better expressed by this quote of Beller and Leerssen: "Identity becomes to mean being identifiable, and is closely linked to the idea of 'permanence through time'" [44].

This paper will provide insight for Location Privacy Protection Mechanisms (LPPM) available, identify their usability and provide a measure to compare between methods. In Section II will be explained the type of LBISs commonly offered in the market and later in Section III will be explained the LPPMs available for such services.

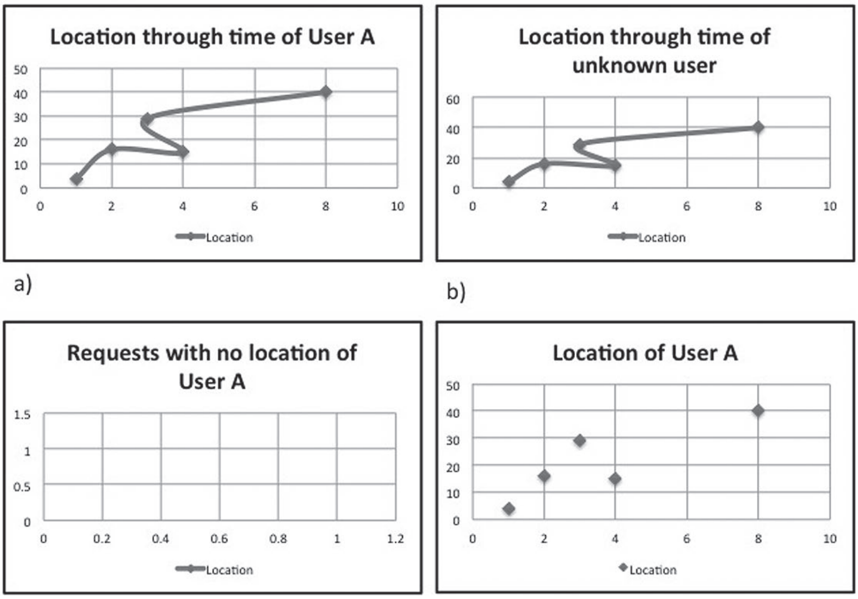


Figure 2. Charts representing data available in different privacy scenarios

2. LOCATION SERVICES

In this section, different location services will be reviewed and categorized based on their intended purpose. LBIS are constituted by a client-server architecture as defined in [25]. Within Location Services are found actors that perform specific roles that make possible the use of the service (figure 3), these roles may be as follows:

- User: Is the LBIS subscribed user that makes a request from a mobile device capable of obtaining the user's location.
- Server: Is the LBIS server that processes the query and provides relevant location information requested by the user, such as Points of Interest (PoI) or navigation services.
- Communication Network: Refers to a communication network such as the Internet, General Packet Radio Service (GPRS) or an ad-hoc network and any other means that make possible the communication between the user and the LBS server.

- Proxy: Is a service that provides security at network level to protect clients' location and identity through IP lookups, these services could be distributed such as The Onion Router (TOR) [22] or centralized like Virtual Private Networks (VPNs), the later is not recommended for protecting specifically identity privacy for the arguments presented in [5].

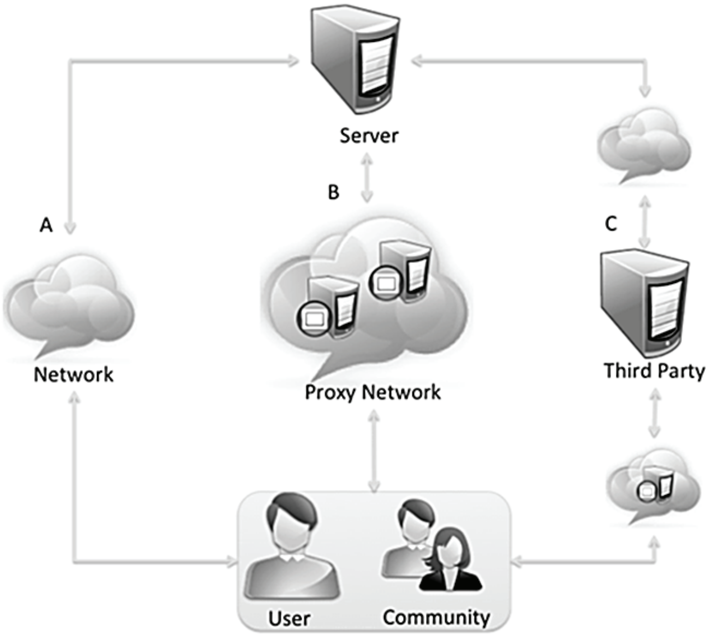


Figure 3. Location services proposed architecture

- Community: Denotes all the users of the LBIS, the community may intervene in the functionality of the service, as is the case of applications used to monitor traffic. Community members could participate in methods for providing location privacy, however not every LPPM requires a community to work.

- Third Parties: Are external relations that intervene to provide location privacy in conjunction with the LPPM. Third parties relations act as proxy-like servers at application level that centralizes the architecture, in [44] it is defined as: "A subjective, dynamic, context-dependent, non-transitive, non-reflexive, non-monotone, and non-additive relation between a trustor and a trustee".

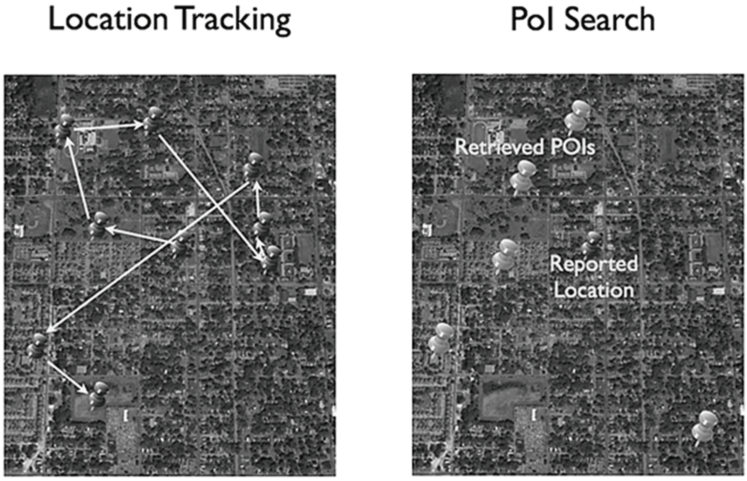


Figure 4. Types of location services

The most commonly used positioning component for providing user's mobile devices with location information required by LBIS is the Global Positioning System (GPS), with installed receivers in most smartphones available nowadays. Other alternatives are the Global System for Mobile Communications (GSM) and the use of WiFi signals for estimating position on mobile devices [1].

On figure 3 the components of a communication schema for a LBIS are shown. Note that all components are not essential for basic functionality. In option A only a network is required, for option B the users communicate through a proxy network and option C shows a third party mediating communication between users/community and the server.

After a thorough study, two main categories were identified based on their behavior: POI Search and Location Tracking as shown on figure 4.

Location Tracking

Location Tracking are services that act silentlylistening to users' location continuously, these services run in the background and allow users to track a device's position and offer functionality based on that trace. On Location Tracking LBIS we have identified the following actors:

- **Monitoring User** is a LBIS user that requires to track locations of a tracked device, many monitoring users could track one device.

- **Tracked device** is the device that is constantly reporting its location to the LBIS to be observed by approved monitoring users.

- **Server**is what provides a platform for communicating between users and tracked devices, also stores logs and historical traces and any additional functions that the LBIS offers.

Location Tracking services are helpful for tracking goods (*i.e.* applications used to find cell phones or laptop computers), traffic monitoring (*i.e.* Waze), friend finder applications, navigation, geomarketing and geofencing; this is, determining when a tracked device trespasses an area delimited by the monitoring user, however, for this it is not necessary that the tracked device communicates its exact location, *i.e.* applications that are used for supervising persons sentenced to house arrest; being the convict a tracked subject and the police being the monitoring user that only needs to receive alerts when the convict trespasses delimited areas. Friend Finder applications are a particular type geofencing. In this case monitoring users are also tracked subjects of their friends or buddies and vice versa, for this kind of applications it is not required for the monitoring users to receive the exact location of the tracked subjects at all times, rather receive alerts when the tracked subjects are within proximity.

Geomarketing applications let users know of available up to date information of their interest and relevant to their location without the user manual request, rather act in background once the user subscribe to the service and are useful to get notification on offers of stores in the moment users are passing by.

PoI Search

Services categorized as PoI Search are designed to allow users query a LBIS that provides nearby places information based on the user's interest and location. These services focus on processing requests and work in a reactively manner, they do not work on background requesting constant location updates from the user, except the location where the user is located when she makes the request to the system, *i.e.*: Requesting the cheapest gas station around a user's current location.

In PoI Search application we have identified the following actors:

- **Requesting User** is a LBIS user through a mobile device with positioning capabilities that makes requests based on her current location or subscribes to receive push notification of nearby places of her interest.

- **Server**; which acts as a PoI Database providing relevant information to the subscribed users about their requests. The service should provide users with places of their interest in the vicinity, available offers or relevant information of such places.

- **Places of Interest** corresponds to the specified kind of location information that the user is willing to receive, not necessarily the type of place where the user is located, *i.e.*: A user requests vegetarian restaurants around her workplace.

Types of LPPMs

LPPMs are designed to provide location privacy to LBIS users, depending on the type of LBIS, there are some requirements it needs to cover. When applied to location tracking services, LPPMs ought to keep location information about the tracked devices undecipherable to anyone different from the allowed monitoring users, even to the LBIS itself. Some location tracking services require a high level of accuracy on the tracked subject's location relying mostly on how good is the approximation of the positioning system. The challenge for LPPMs used for this kind of applications is to maintain the level of accuracy that the positioning component in the device can provide and at the same time assure that the information does not get disclosed to anyone other than the allowed monitoring users. On the other hand, mechanisms intended for PoI search services may alter the subject's position in order to provide location privacy or require special implementation on the server side to be privately queried and process requests.

LPPMs' available techniques include: Cryptography-based, Private Information Retrieval (PIR), Progressive Retrieval, Noise-based techniques, Spatial Cloaking, K-Anonymity, Pseudonyms and Dummy Queries. Some mechanisms may include more than one technique to achieve location privacy. On Figure 5 is exposed a taxonomy based on types of location services of these methods under their respective application.

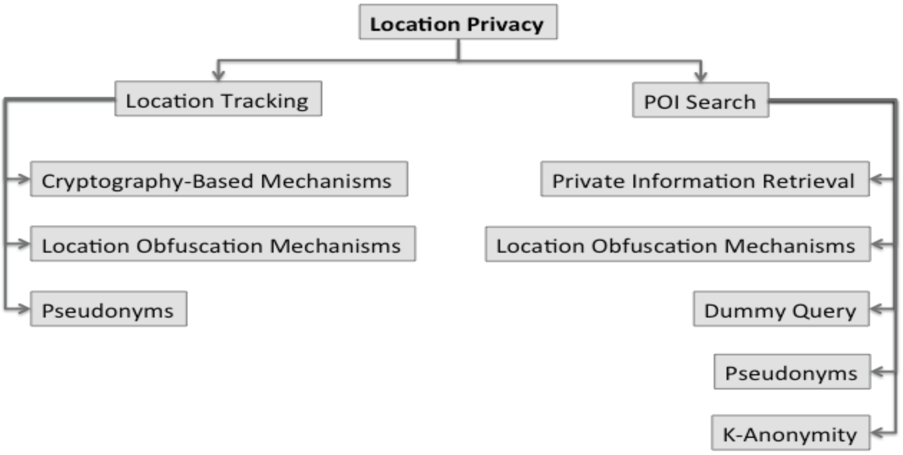

Location Privacy
Location Tracking
POI Search
Cryptography-Based Mechanisms
Location Obfuscation Mechanisms
Pseudonyms
Private Information Retrieval
Location Obfuscation Mechanisms
Dummy Query
Pseudonyms
K-Anonymity

Figure 5. Taxonomy of different mechanisms

Cryptography-Based Mechanisms

LPPMs used for tracking may be cryptography-based, these methods offer secure communication and preserve location information accuracy. In[39] a model for a location privacy aware friend finder application is proposed with two alternative protocols. In the privacy requirements specified, it states that each user should be capable of controlling the location information to be disclosed to others; the service provider should have as little information as possible and the user's friends should know the proximity but not her exact position, also, any eavesdropper in the network should not be able to filter any location information about the users. To accomplish this, it is defined a Minimal Uncertainty Region (MUR): "the user accepts that the adversary knows she is located in a MUR R, but no information should be disclosed about her position within R". In order to capture these uncertainty regions, they use spatial granularity, which is understood in many LPPMs as "a subdivision of the spatial domain into a discrete number of non-overlapping regions, called granules".

The two protocols presented are C-Hide & Seek and C-Hide & Hash, both adopt symmetric encryption techniques where each user poses a unique key that is shared with their friends and vice versa. The key exchange is performed through a secure communication before executing the protocols. In this scheme each user has to report their location to the service provider, this is done by discretizing time in update intervals. The success of these protocols lies in the fact that for every update of a user, a different key is used. This is possible due to the generation of a key stream based on the initially exchanged key of the users, each buddy will be able to generate the key corresponding to the current update interval of their approved buddies and therefore decrypt the identification of the granule the user is in. The main different between C-Hide & Seek and C-Hide & Hash protocols, is that the first one lets the buddies know the granule where the user is located, while the second requires more computational cost but manages to provide full privacy without disclosing the granule, save the case when the user is in proximity. The C-Hide & Seek protocol is designed to be used with any symmetric encryption technique and the C-Hide & Hash with a hash function.

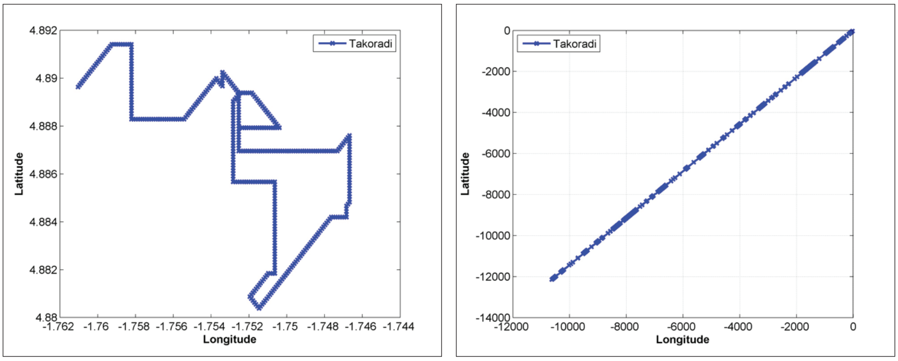


Figure 6. Matlock observable data vs. Original user trace taken from [36]

A holomorphic encryption based on matrix obfuscation technique is presented in [36], the mechanism uses a matrix $M_{(1,3)}$ containing the latitude, longitude and the time when the coordinates where obtained. According to [52]: "Holomorphic encryption is a special kind of encryption that allows operating on cipher texts without decrypting them; in fact, without even knowing the decryption key".

The method requires a second matrix $N_{(3,3)}$ to perform the obfuscation operations. The resulting matrix Q from the operations between M and N contains the location information encrypted and undecipherable by an LBIS. In order to decrypt the information, the inverse of the N matrix, N^{1} will be used as shared key between the motoring users and tracked device to access the matrix containing the obfuscated location information M. The method allows the position reported, to be recovered unaltered while providing location information. As seen on Figure 6, the information obtained by the LBIS or anyone different from the allowed monitoring users is not interpretable in any geographical sense.

Private Information Retrieval

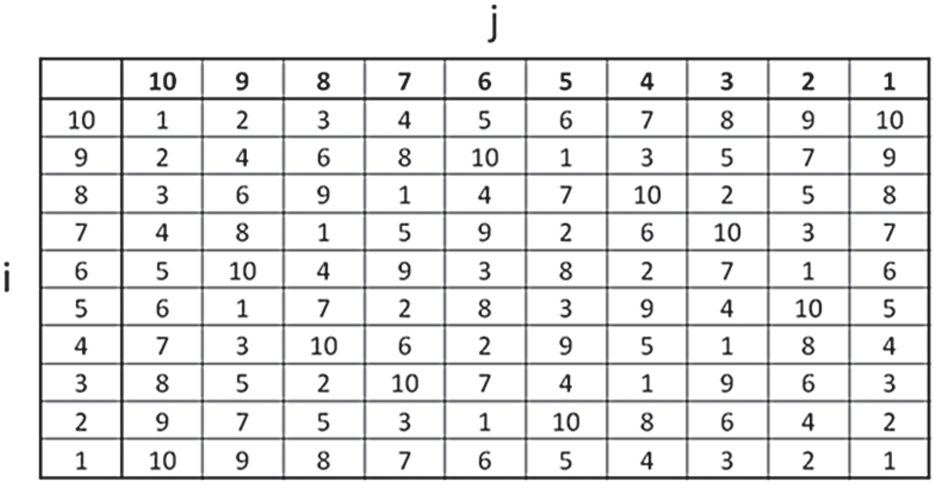
Private Information Retrieval is a widely used approach for providing Location Privacy on nearest neighbor searches, formally, it was first defined in [13] as "… schemes that enable a user to access k replicated copies of a database (k≥2) and privately retrieve information stored in the database. This means that each individual database gets no information on the identity of the item retrieved by the user", but this approach was not initially intended to be used on a single database, it required replication on at least two databases with communication restriction between them, since it aimed to provide information theoretic privacy, which demands an adversary with no knowledge of the information requested and assumes unlimited computational resources for the attacks. It was not until 1997 in [28] and[10] when a computational PIR (cPIR) technique with a single database was presented, this scheme assumes an attacker limited to probabilistic polynomial-time computations. PIR techniques are challenged to provide solutions with reasonable computation and communication costs. PIR implementations usually require special processing in the server side of the LBIS.

In [29], authors distinguish between Cryptography-based and Hardware-based PIR, where the first one utilizes cryptographic transformation to the query and/or database structure, while the second requires special hardware architecture with a Secure Coprocessor (SC) that act as a securely protected space where the retrieval of information takes places in a way that the LBIS cannot decipher. The proposed technique in [29]; SPIRAL, is hardware-based and uses random permutation of the database items with a mapping that is only stored in the SC, the SC also caches the items retrieved to a user to ensure that each item in the database is queried at most once and avoid inferences from attackers or the LBIS itself, when the cache in the SC becomes full a reshuffling of the entire database is performed, they propose to generate offline reshuffled databases to avoid increasing computational costs, a downside of this method is that it does not support k-nearest neighbor search, rather retrieves the it item requested by users. In[37] is presented another hardware-based PIR mechanism that aims to provide location privacy with a MUR of the entire spatial domain and supports k nearest neighbor search.

In [31] is presented a Cryptography-based cPIR protocol designed to be used with any PIR technique according to the authors, this method adds spatial cloaking to reduce the database domain to be searched. The protocol consists on discretizing the space in the form of a space granularity based on a Hilbert curve of the concentration of PoIs in an area, in a way that each resulting granule will contain the same amount of PoIs, the number of PoIs is set in the beginning of the processing and cannot be changed later without altering the database. The biggest area granule resulting from the calculation is set as the size of a cloaking region, a user can chose a bigger region consisting of more than one cell, which is later consulted with the chosen PIR mechanism to retrieve only the PoIs in the user Hilbert Cell.

Another technique that uses cryptography-based cPIR is presented in[47], the Mapping-Based Private Information Retrieval (MaPIR) method introduces redundant identification on a spatial granularity represented by a grid of squared zones (figure 7). The granules' IDs can be calculated with basic arithmetic operation on both, the mobile device and server, IDs are a transformation of location coordinates and go from 1 to 10 so the ID alone does not reveal any location information, for detailed explanation of the calculations see [47].

Redundancy consists on each ID matching 10 different granules on the spatial grid in order to provide a greater MUR when and ID is reported by the client device and later when retrieving the PoIs. PoIs are stored in the database along with their computed IDs, the PoIs matching an ID requested by a user are then all retrieved to that user. The specific type of place of interest of the user is not considered private and is used to filter undesired PoIs to lower communication costs and allowing the LBIS to get statistics on their user's interests.



	10	9	8	7	6	5	4	3	2	1
10	1	2	3	4	5	6	7	8	9	10
9	2	4	6	8	10	1	3	5	7	9
8	3	6	9	1	4	7	10	2	5	8
7	4	8	1	5	9	2	6	10	3	7
6	5	10	4	9	3	8	2	7	1	6
5	6	1	7	2	8	3	9	4	10	5
4	7	3	10	6	2	9	5	1	8	4
3	8	5	2	10	7	4	1	9	6	3
2	9	7	5	3	1	10	8	6	4	2
1	10	9	8	7	6	5	4	3	2	1

Figure 7. MaPIR redundancy function

Reference [21] presents PIR methods with reasonable computation and communication costs, AproxNN uses a discretized spatial domain with Hilbert Curve ordering to represent PoIs in a one-dimensional space, the PoIs are queried by the user and retrieved using a binary search. The user location is also transformed by the same Hilbert Curve function and PoIs are retrieved based on the assumption that granules that are close in the two-dimensional granularity are also close in the Hibert Curve ordering. For the processing it uses a B+-tree that contains the PoIs in ascending order with leaves not greater or equal to its root. It is also introduced ExactNN, a method that maps PoIs using a Voronoi tessellation in a way that each Voronoi cell contains exactly one PoI, also it superposes a regular granularity squared grid that is privately queried by the user and retrieves the PoIs contained in the Voronoi cells that intersect the grid cell. For grid cells that are fully contained within a Voronoi cell, it generates fake PoIs to match the number of bytes retrieved by cells with maximum intersected Voronoi cells.

Noise-based or Location Obfuscation Mechanisms

Noise-based LPPM transform the user's location in a way that the original location is permanently lost, however the resulting obfuscated location is still close enough to be used by a LBIS and provide acceptable performance. Since the induced noise cannot be too large, because the LBIS relevance would be affected, the overall users' route could be inferred, in[24] were performed tests to infer home addresses from users location tracking logs. The results found that for imprecision obfuscation with a simple Gaussian noise technique there needs to be a standard deviation of 2 kilometers of added noise, in order to reduce to near zero the amount of correct inferences in the attacks.

An advantage of noise techniques is those don't require a special implementation on LBIS, therefore can be implemented on the device without interfering with the LBIS.

In [8] is defined location obfuscation as "the means of deliberately degrading the quality of information about an individual's location in order to protect that individual's location privacy". There are three terms identified for imperfection on spatial information: "Inaccuracy, imprecision, and vagueness. Inaccuracy concerns a lack of correspondence between information and reality; imprecision concerns a lack of specificity in information; vagueness concerns the existence of boundary cases in information..."

Some LPPMs that uses noise-based techniques are presented in [2], [18], [33], in [2] the authors introduce a measure that takes into account the precision that a sensing technology may provide while calculating a subject's location and they call it relevance, through this measure, they allow the user to specify their privacy preferences, *e.g*: "100 meters" would specify that the user cannot be located with an accuracy not better that 100 meters, to satisfy this requirement, the authors introduce obfuscation operators that consist on enlargement of the reported area, generation of a reduced obfuscated area that lowers the possibility of the real position to be located within and shifting the center of the reported area.

The method presented in [18] is based on a vagueness technique which deforms the quality of the information by assigning another value with an algorithm which is O(n^2) in complexity and time is O(n). The method requires a discrete representation of the world consisting of a non-empty set of unique area called regions. Each region represents a location. There is also a set of relations, which represent how near the subject is from each region. The obfuscation process consists on changing the relation to a more vague one, i.e. the user specifies she is near x location. The obfuscation for that location would be changing the relation "near" to "around".

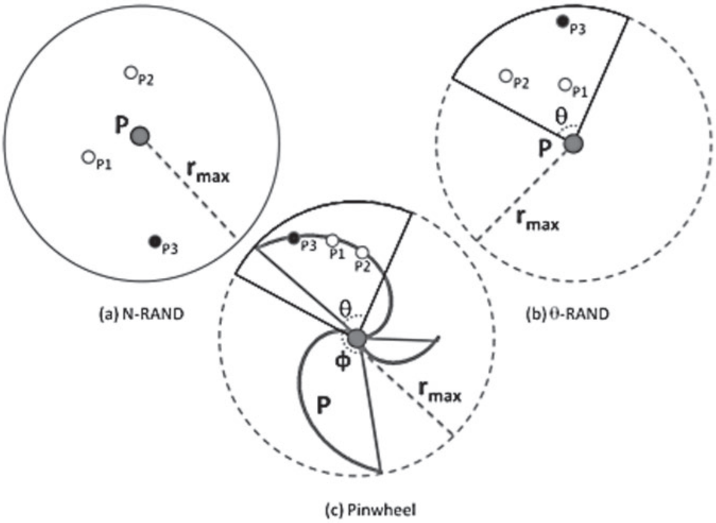


Figure 8. Examples of the generation of a random point in N-RAND, θ-RAND with θ=40o and Pinwheel with φ=140°

Authors in [33] present multiple point-based obfuscation techniques; these add noise to the original location to generate an imprecise obfuscated location. The results proved the N-Rand to be the most efficient, consisting on the generation of N random points with a radius centered on the original location to finally choose the farthest from that location. Later, the authors proposed the methods θ-Rand [49] and Pinwheel [48], θ-Rand is a variation of N-Rand, it differs in the domain used to generated random points, θ-Rand defines a sector of the circle by delimiting it to a radius not greater than the original circle radius and a defined θ angle. The generation of the random points within the domain is the same as in N-rand and the farthest randomly generated point from the center, is the resulting obfuscated location. The results are compared with its predecessor using the Exponential Moving Average algorithm (EMA), where the θ-Rand showed an improved result filtering less noise. In the latest work, the authors proposed a mechanism inspired in Pinwheel shapes, it also defines a domain in the same way as in θ-Rand and generates points contained within, Pinwheel differs from the previous methods in the generation of points, which in this case will not be at a random radius from the center, rather for each angle is defined a value for the radius by a function that follows the trajectory of a pinwheel. Figure 8 shows the dominium for the generation of points of the different mechanisms.

Dummy Query

The use of dummy query is a popular approach to provide location privacy in PoI search services, it consists on sending N fake requests along with the real one in order to disguise the user's true location, this technique poses downsides as it requires the server to process N queries additional to the one relevant to the user, this incurs in servers overhead and communication costs, however there are some techniques developed based on dummy queries that manage to decrease such costs.

In 2005 authors developed the first technique inspired in dummy queries[23], this first approach proposes the generation of n-1 fake locations to be sent to an LBIS to provide location privacy. The scheme assumes users that are constantly reporting its location to an LBIS and therefore would not be located too distant from the immediately previous reported location. The dummy locations are generated in a way that form feasible traces of a regular user; for a first query of a user, there are generated n-1 random fake location and sent to the LBIS with the real one, for the following requests, the method bases the generation of dummies in the ones previously reported in order to build n possible traces for that user.

In [43]SpotME is introduced, this method works with large scale amount of users to count people in certain areas in order to provide information related to traffic, crowd analysis, etc. SpotME requires the geographical space to be divided in locations, these defined locations are the ones users chose from to say whether or not they are present, each with 50 % of probability and the users are not forced to answer truthfully, this data is later manipulated by an algorithm that estimates the real proportions of all data received by the LBIS, it shows an accurate result of people concentration at certain locations, without identifying between users. The method is also able to indicate if people are entering or exiting a location, which results useful to estimate population flow. The vulnerability of SpotME resides in the ability of an attacker to collect more than one map of the location sent by a user and compare and intersect the maps to get one possible real location.

In[38] are proposed two techniques to generate dummy requests to LBISs, both techniques send a single message to the service to lower communication costs and require a light transformation on the server side for processing the requests, which are conformed by n positions and a type of interest, which they call query predicate that applies to all dummy requests. The first technique generates dummies forming a grid with a dummy in each vertex, while the second one generates the dummies based on a virtual circle that contains the user's location. The server processes all the locations with the same predicate and retrieves all the information that is later filtered by the client.

Pseudonyms

Pseudonyms are an alternative to provide identity privacy in location based applications, however the use of pseudonyms alone is not sufficient to provide location privacy in a LPPM since a pseudonym that stays the same over time will eventually lead to the identification of a user as stated in the definition of Idem Identity [44].

For a pseudonym-based implementation of a LPPM was introduced the notion of mixed zones in [4], the proposal is intended for applications that cannot be accessed anonymously, but do not require the user's true identity either, rather an internal pseudonym managed by the service. The mechanism requires a third trusted party middleware to provide users with pseudonyms and to guarantee that the true users' identity is not revealed to the LBIS. Identity anonymization is provided by changing pseudonyms over time in designated mixed zones, which are zones where users do not have any applications subscribed (in the case of proactive LBIS) and therefore are able to change of pseudonym. Downsides include the case when no users are available at a mixed zone or this one may be too large so the LBIS may identify users.

A technique presented by Eckhoff et al. in [15], proposes the exchange of identities or pseudonyms between users of the service, and also keeps these pseudonyms changing every period of time (slot) in order not to give an attacker the possibility to link two or more requests with the same handle.

K-anonymity and Spatial Cloaking

K-anonymity was one of the first approaches presented to achieve location privacy; it consists on making a user indistinguishable among other K-1 users, as in [20], [14], [27].

In some implementations, the user may be able to specify the *k* parameter. In many of these techniques areused cloaked regions to provide such anonymity, where *k* users are similar enough within an area, as to deceive attackers from identifying the real issuer of an LBS request, thus, regions with higher density of users, result in smaller cloaked areas. A hindering factor for cloaked*k*-Anonymity is that often incur in the *outlier problem* as described in [42].

On the LBIS side these methods may incur a computational overhead, since processing costs of a query with a region as input, instead of a discrete point requires special spatial calculations.

Cloaking was introduced in[19],where is proposed a centralized architecture for obfuscating anonymous location information, the algorithm consists on providing K-anonymity through spatial cloaking, the centralized server is a trusted third party that has to know the locations of the LBS users at all times and uses them to assure that at least K users are contained within the reported area. Another cloaking method implemented is the temporal cloaking, which instead on enlarging the reported area, delays the request until at least K users have visited the user reported location.

Authors in [7]present a mechanism that aims to provide cloaking privacy in Peer-to-Peer networks, using members of the community to exchange location information between them in order to calculate a cloaked region to be reported to the LBIS, thus eliminating the need of a trusted third party. The main downsides of this technique are the communication costs, and in some cases there may not be enough users to form a cloaked region, also the fact that any malicious user in the community may obtain members' locations. In [41] the term *information leakage* was introduced; referring to the amount of revealed location information in spatial cloaking to provide a better performance.

Mokbel et al. present Casper in [45] and [9]; a mechanism that uses a quadtree to represent the cloaked regions, where the root node is the whole domain and the leaf nodes represent a quadrant of its parents.

In [12] is presented an *imprecise location-based range query (ILRQ)*,which serves for issuing users to know if the subject of interest is within a specified range from them, similar to *geofencing*. It is imprecise since it process cloaked regions, which result in probable answers to the query but not a definitive one.

Progressive Retrieval

Methods based on Progressive Retrieval (PR) perform many requests for a single user interaction. This approach aims to reveal as least location information as possible to obtain the desired service performance.

A PR implementation can be seeing on Space Twist [50] and Anon Twist [32], the last one being an improvement to [50], both use an anchor which is a fake location that is contained within an area of radius P from the original location. Centered on the original location is defined a demand space stating how close should a PoI be from the original location to be accepted as relevant by the user, the algorithms make use of a supply space, which in first place is just the anchor, but increases the region radius to the latest nearest PoI retrieved on each iteration. The algorithm finishes when the demand space is fully contained with the supply space, guaranteeing that the nearest POI for the real location is available to the client without releasing the real location to the server. Anon Twist proposes an improvement by introducing density maps, which brings k-anonymity to the algorithm. In[35] is presented a technique that improves Anon Twist by guaranteeing absence privacy as well, allowing the user to specify a puppet location where she does not want an attacker to infer she is not present, ie. Avoid disclosing when she is not home. In order to do so, the maximum distance between a puppet location and the user's real location must be half the initial radius, this is *p/*2. With a puppet location already specified, the algorithm keeps requesting POIs until the puppet location *p* is contained within the candidate area, thus making this technique less efficient.

3. EVALUATION

Given that there are such a variety of mechanisms to protect location privacy, it is not feasible to provide a quantitative evaluation of all these solutions under a common scenario and to declare a single mechanism as the best for all situations.

This paper proposes a qualitative comparison of the techniques based on a simple comparison table that summarizes some of the main aspects identified among all techniques, and provides a guide on the desired characteristics for some of the most common types of LBIS applications.

On table I a comparison between the LPPMs referenced in this survey is presented. In order to provide at least an initial tool for comparing LPPMs mechanisms, some key factors are proposed:

- Allows PoI search
- Can be used in location tracking services -where geofencing is included-. Mechanisms applicable to location tracking services may not always provide accurate information, as is the case of noise-based LPPMs, for these, the acronym LA (Low Accuracy) will be used in the column Allows Tracking.

Table 1. Comparison Table

LPPM	Type of Technique	Allows PoI-Search	Allows Tracking	Requires Third Party/ Hardware	Reports Location Info to LBIS	MUR	Special Implementation in the LBIS
[39]	Cryptography	N	Y	N	N	ESD	Y
[36]		N	Y	N	N	ESD	Y
[29]	PIR	Y	N	SC	N	ESD	Y
[31]		Y	N	N	Region	CI/ CO	Y
[47]		Y	N	N	Region	CI	Y
[21]		Y	N	N	Region	CI	Y
[2]	Noise-Based	Y	LA	N	Region	NG	Y
[18]		Y	LA	N	Y	NG	Y
[33]		Y	LA	N	Y	NG	N
[24]		Y	LA	N	Y	NG	N
[48], [49]		Y	LA	N	Y	NG	N
[23]	Dummy Queries	Y	N	N	Y	CI	N
[43]		N	Anonymous	N	Y	CI	Y
[38]		Y	N	N	Y	CI	Y
[4]	Pseudonym	Y	Anonymous	Y	Y	N	N
[15]		Y	Anonymous	N	Y	N	N
[19]	K-Anonymity	Y	N	Y	Region	NG	Y
[7]		Y	N	N	Region	NG	Y
[50]	PR	Y	N	N	Y	NG	N
[32],[35]		Y	N	Density Map	Y	NG	N

- Requires a third trusted party or trusted hardware component
- Reports any location information to the LBIS
- Magnitude of the MUR; which could be of the Entire Spatial Domain (ESD), Country (CO), City (CI), Neighborhood (NG) or None (N) for methods that report the exact location
- Requires special implementation in the LBIS side.

An ideal implementation, based on surveyed techniques, should be general enough to be used on both PoI search and tracking; however, specialization of the techniques is not a very negative issue if the advantages against general techniques are large enough. In addition, in order to reduce the footprint of the solution, operational costs and availability, the solution should not need a trusted third party in order to work properly because the successful usage of the solution always would depend on the availability of this component; however, having a strong third party can give a certain level of security to the system that a purely distributed one may not reach, like in terms of identity verification. Also, in terms of special hardware, it would increase the costs of an actual implementation, but on some cases may become necessary, like when the actual general purpose hardware is vulnerable in nature and the application is used in critical environments, like in military scenarios where hardwired cryptography-based communication may be necessary to avoid eavesdropping from the enemy.

One important factor to take into account is if the application needs to report the location information; the ideal would be for the exact location not to be shared but only when necessary. For critical applications, cryptographic techniques and those that report regions instead of locations would be desirables, compared to techniques in which the location is shared, even if slightly altered. On the other hand, for non-critical applications in which it may not be that important to reveal the location, the cost of the cryptographic techniques can be high in terms of the access to the real information and the kind of services that you can provide over the encoded data or the slightly altered, compared to a small footprint technique like the point-based obfuscation that can offer a certain level of protection at a low cost, while preserving the geographical validity of the data and its availability for immediate usage.

The scale of the protection is also important, but depends directly on the nature of LBIS. Some applications require a maximum MUR like the case of a value truck, we should need to alter as much as possible the information in order to provide no useful information to attackers, while being able to retrace the real path. While others not as critical may allow a smaller MUR.

Finally, given that many applications already exist, requiring special implementations on the server side may require a large investment to include the mechanism. The LPPM solution should require minimal or no investment on the server side, and probably just changes in the client application level.

CONCLUSIONS

In the literature there have been numerous LPPMs to ensure protection of the users while using location services, but they have not been implemented commercially due to many reasons; the lack of interest of the location services companies and the fact that there are too many available mechanisms without a vision of real applicability on existing location services, with different characteristics and without a measurable indicator to compare between. In this survey we intend compiling representative mechanisms for each identified technique and provide an approach for evaluation that can be used with any LPPM to put in perspective the functionality it offers against other mechanisms and provide insight for both;the LBIS and the user in order to establish a start point when analyzing the suitability of a LPPM implementation for a service, and in the future give a stronger basis to support such implementation in commercial services that protect user's right to location privacy. Nowadays such privacy has been limited to the presentation of agreements of terms and conditions, leading to being tracked to whatever purposes the company desires, or simply discard the agreement and avoid using LBIS and be exempt of the benefits that they offer.

LBIS are often not isolated products, but a piece of greater information systems, which are indispensable for many individuals and business corporations who cannot risk stopping its use. Table 1 provides a good landscape of the nature of many existing LPPMs, which should become a guide for developers in order to select the mechanism that offers the best characteristics to their needs.

Acknowledgement

This work was supported by the Young Researcher program from Colciencias 2012-2013 and Universidad del Norte, Barranquilla (Colombia), under the grant number UN-OJ-2012-16910.

REFERENCES

[1] C. L. Bowen III and T. L. Martin, "A survey of location privacy and an approach for solitary users", in *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, 2007, p. 163c-163c.

[2] C. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati, "An Obfuscation-Based Approach for Protecting Location Privacy", *IEEE Transactions on Dependable and Secure Computing*, vol. 8, n° 1, pp. 13-27, Jan. 2011.

[3] P. Bellavista, A. Kupper, and S. Helal, "Location-Based Services: Back to the Future," *IEEE Pervasive Computing*, vol. 7, n° 2, pp. 85-89, Apr. 2008.

[4] A. R. Beresford and F. Stajano, "Mix zones: User privacy in location-aware services", in *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, 2004, pp. 127-131.

[5] J. Appelbaum, M. Ray, K. Koscher, and I. Finder, "vpwns: Virtual pwned networks", in *2nd Workshop on Free and Open Communications on the Internet*, 2012.

[6] *California Location Privacy Act, 2012.*

[7] Y. Che, Q. Yang, and X. Hong, "A dual-active spatial cloaking algorithm for location privacy preserving in mobile peer-to-peer networks", in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, 2012, pp. 209-2102.

[8] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *Pervasive Computing*. Springer, 2005, pp. 152-170.

[9] C.-Y. Chow, M. F. Mokbel, and W. G. Aref, "Casper*: Query processing for location services without compromising privacy", *ACM Transactions on Database Systems*, vol. 34, n° 4, pp. 1- 48, Dec. 2009.

[10] B. Chor and N. Gilboa, "Computationally private information retrieval", 1997, pp. 304-313.

[11] M. Duckham and L. Kulik, "Location privacy and location-aware computing", *Dynamic & mobile GIS: investigating change in space and time*, vol. 3, pp. 35-51, 2006.

[12] R. Cheng, Y. Zhang, E. Bertino, and S. Prabhakar, "Preserving User Location Privacy in Mobile Data Management Infrastructures", in *Privacy Enhancing Technologies*, G. Danezis and P. Golle, Eds. Berlin-Heidelberg: Springer, 2006, pp. 393-412.

[13] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval", *Journal of the ACM*, vol. 45, n° 6, pp. 965-981, Nov. 1998.

[14] Z. Gong, G.-Z. Sun, and X. Xie, “Protecting Privacy in Location-Based Services Using K-Anonymity without Cloaked Region”, 2010, pp. 366-371.

[15] D. Eckhoff, R. German, C. Sommer, F. Dressler, and T. Gansen, “SlotSwap: strong and affordable location privacy in intelligent transportation systems”, *Communicatios Magazine, IEEE*, vol. 49, n° 11, pp. 126,133, Nov. 2011.

[16] European GNSS Agency. Opportunities Abound in Growing Location-Based Services Market. Available: http://www.gsa.europa.eu/go/news/opportunities-abound-in-growinglocation-based-services-market, 2010.

[17] European Union Directive on Privacy and Electronic Communications, 2002.

[18] J. Haadi Jafarian, “A Vagueness-based Obfuscation Technique for Protecting Location Privacy”, in *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, 2010, pp. 865-872.

[19] 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*, 2003, pp. 31-42.

[20] B. Gedik and L. Liu, “Location Privacy in Mobile Systems: A Personalized Anonymization Model”, in *25th IEEE International Conference on Distributed Computing Systems, 2005. ICDCS 2005. Proceedings*, 2005, pp. 620-629.

[21] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan, “Private queries in location based services: anonymizers are not necessary”, in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, 2008, pp. 121-132.

[22] H. Kido, Y. Yanagisawa, and T. Satoh, “An anonymous communication technique using dummies for location-based services”, in *Pervasive Services, 2005. ICPS’05. Proceedings International Conference on*, 2005, pp. 88-97.

[23] J. Krumm, “Inference attacks on location tracks”, in *Pervasive Computing*. Springer, 2007, pp. 127-143.

[24] C. S. Jensen, H. Lu, and M. L. Yiu, “Location privacy techniques in client-server architectures”, in *Privacy in location-based applications*. Springer, 2009, pp. 31-58.

[25] M. A. Labrador, *Location-based information systems: developing real-time tracking applications*. Boca Raton: CRC Press, 2011.

[26] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, “Preventing Location-Based Identity Inference in Anonymous Spatial Queries,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, n° 12, pp. 1719-1733, Dec. 2007.

[27] E. Kushilevitz and R. Ostrovsky, “Replication is not needed: Single database, computationally-private information retrieval”, in *Foundations of Computer Science, 1997. Proceedings 38th Annual Symposium on*, 1997, pp. 364-373.

[28] A. Khoshgozaran, H. Shirani-Mehr, and C. Shahabi, "SPIRAL: A Scalable Private Information Retrieval Approach to Location Privacy", 2008, pp. 55-62.

[29] Ley 1581, 2012. Colombia.

[30] F. Olumofin, P. K. Tysowski, I. Goldberg, and U. Hengartner, "Achieving efficient query privacy for location based services", in *Privacy Enhancing Technologies*, 2010, pp. 93-110.

[31] S. Wang and X. S. Wang, "Anon Twist: Nearest Neighbor Querying with Both Location Privacy and K-anonymity for Mobile Users", 2009, pp. 443-448.

[32] P. Wightman, W. Coronell, D. Jabba, M. Jimeno, and M. Labrador, "Evaluation of Location Obfuscation techniques for privacy in location based information systems", in *Communications (LATINCOM), 2011 IEEE Latin-American Conference on*, 2011, pp. 1-6.

[33] S. Lederer, A. K. Dey, and J. Mankoff, "Everyday privacy in ubiquitous computing environments", in *Ubicomp 2002 Workshop on Socially-informed Design of Privacy-enhancing Solutions in Ubiquitous Computing*, 2002.

[34] D. Riboni, L. Pareschi, and C. Bettini, "Integrating Identity, Location, and Absence Privacy in Context-Aware Retrieval of Points of Interest", 2011, pp. 135-140.

[35] P. M. Wightman, M. A. Jimeno, D. Jabba, and M. Labrador, "Matlock: A location obfuscation technique for accuracy-restricted applications", in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, 2012, pp. 1829-1834.

[36] S. Papadopoulos, S. Bakiras, and D. Papadias, "Nearest neighbor search with strong location privacy", *Proceedings of the VLDB Endowment*, vol. 3, n° 1–2, pp. 619-629, 2010.

[37] H. Lu, C. S. Jensen, and M. L. Yiu, "Pad: Privacy-area aware, dummy-based location privacy in mobile services", in *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, 2008, pp. 16-23.

[38] S. Mascetti, D. Freni, C. Bettini, X. S. Wang, and S. Jajodia, "Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies", *The VLDB journal*, vol. 20, n° 4, pp. 541-566, 2011.

[39] D. Riboni, L. Pareschi, and C. Bettini, "Privacy in georeferenced context-aware services: A survey", in *Privacy in Location-Based Applications*. Springer, 2009, pp. 151-172.

[40] K. W. Tan, Y. Lin, and K. Mouratidis, "Spatial cloaking revisited: Distinguishing information leakage from anonymity", in *Advances in Spatial and Temporal Databases*. Springer, 2009, pp. 117-134.

[41] S. Mascetti, C. Bettini, D. Freni, and X. S. Wang, "Spatial generalisation algorithms for LBS privacy preservation", *Journal of Location Based Services*, vol. 1, n° 3, pp. 179-207, 2007.

[42] D. Quercia, I. Leontiadis, L. McNamara, C. Mascolo, and J. Crowcroft, "Spotme if you can: Randomized responses for location obfuscation on mobile phones", in *Distributed Computing Systems (ICDCS), 2011 31st International Conference on*, 2011, pp. 363-372.

[43] K. Rannenberg, D. Royer, and A. Deuker, *The future of identity in the information society: challenges and opportunities*. Berlin-London: Springer, 2009.

[44] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The new casper: A privacy-aware location-based database server", in *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, 2007, pp. 1499-1500.

[45] S. D. Warren and L. D. Brandeis, "The right to privacy", *Harvard law review*, vol. 4, n° 5, pp. 193-220, 1890.

[46] P. Wightman, M. Zurbarán, M. Rodríguez, and M. Labrador, "MaPIR: Mapping-Based Private Information Retrieval for Location Privacy in LBISs", in *8th IEEE Workshop on Network Security (WNS) on*, 2013.

[47] P. Wightman, M. Zurbarán, and A. Santander, "High Variability Geographical Obfuscation for Location Privacy", in *47th International Carnahan Conference on Security Technology(ICCST) on*, 2013.

[48] P. Wightman, M. Zurbarán, E. Zurek, A. Salazar, D. Jabba, and M. Jimeno, "θ-Rand: Random Noise-based Location Obfuscation Based on Circle Sectors", in *IEEE International Symposium on Industrial Electronics and Applications (ISIEA) on*, 2013.

[40] M. L. Yiu, C. S. Jensen, X. Huang, and H. Lu, "Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services", in *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on*, 2008, pp. 366-375.

[50] X. Chen and J. Pang, "Protecting query privacy in location-based services," *GeoInformatica*, vol. 18, n°1, pp. 95-133, Oct. 2013.

[51] D. Micciancio, "A First Glimpse of Cryptography's Holy Grail," *Commun. ACM*, vol. 53, n° 3, pp. 96-96, March. 2010.