

Location Privacy – An Overview

Michael Decker
Institute AIFB, University of Karlsruhe (TH)
76 128 Karlsruhe, Germany
decker@aifb.uni-karlsruhe.de

Abstract

Location-based services (LBS) are mobile services which evaluate knowledge about a user's location. In most cases this is done to provide the user with personalized information, e.g. to list point of interests in his nearer surrounding. But users have concerns when their whereabouts are tracked and their location information is disclosed to external service providers. So it is no wonder that a lot of publications deal with the specific privacy issues of LBS usage. The aim of this article is to give an overview about the most relevant works in this field: we first discuss different attack scenarios where location information is misused. Afterwards we present a classification of technical approaches to prevent such misuse of location information. We also summarize some results from pertinent empirical studies.

1. Introduction

Mobile services are services provided over wireless data communication (e.g. GPRS, UMTS) that are used with handheld computers like PDAs, cellular phones or smartphones. Location based services (LBS) are a special class of mobile services: these services evaluate knowledge about the position of a mobile device to deliver the service. A suggesting example would be a service that lists points-of-interest (POIs) like restaurants, ATMs, petrol stations or shops in the nearer surrounding of the current position of the calling user. But our definition of LBS also includes services that take the position of a mobile device distinct from the calling user into account; an example for such a service would be a "track-my-kid" or "friend-finder" service to find out about another person's current whereabouts. There are several reasons that explain the great popularity of LBS in the research community:

- An approximate estimation of a mobile user's position can be retrieved as side effect of the mobility manage-

ment in cellular networks like GSM or UMTS [21]: in such networks the operator has to maintain a database with information about the base station currently used by the user's mobile device. This is necessary to route mobile terminated calls and messages to the right cell and also to provide a handover when the user is moving between cells. Cells may have a radius of up to several kilometres in rural or costal areas (up to 35 km for ordinary cells in the case of GSM) but for many LBS the precision of location information of this "Cell-ID" is sufficient (e.g. location aware weather forecast or advertising messages).

- There are many demonstrative application scenarios for LBS, for example those already mentioned.
- Navigation services in general and especially those for motorists are one kind of LBS that already enjoys great commercial success [7] and could act as trailblazer for other kinds of LBS. For example after reaching the destination the navigation service could inform the user about points-of-interest in his surrounding, e.g. to find a hotel or restaurant.
- The user's location will be subject to many changes in typical application scenarios because mobile devices are used when being "on the move" so the user will appreciate if he hasn't to enter his current location all the time.
- In the USA the Federal Communications Commission (FCC) passed the so called "Enhanced 911" (E-911) mandate that obligates Mobile Network Operators (MNO) to provide the originating position of a call to emergency services (e.g. ambulance, police) [21]. In the first phase MNOs were required to be able to report the location of the serving base station to the emergency service; in the second phase it was demanded to provide location information with a precision of 50 to 100 meters in 67 % of cases. Especially the requirements of the second phase turned out to be challenging and lead to a great innovation push. The possibility to

retrieve a mobile device's position also inspired a lot of ideas for new services.

- Due to size and weight constraints a mobile device has only limited means for data input, e.g. it has some buttons but no full keyboard or pointer device. So by evaluating the user's current location the service can save the user from the cumbersome act of entering his current location. In some scenarios the user might even not be aware of his position, e.g. when visiting new places or if he got lost.
- The display of a mobile device is also of inferior quality (small size, limited colour depth and resolution, poor contrast). So the user will appreciate it if the service limits the displayed information to the relevant one according to his current location.

The last two points deal with assisting the user when interacting with a mobile device by using location information. This approach can be generalized to use other kind of information (e.g. time, personal profile data, etc.) and is then termed "context awareness" [12]. Location awareness is by far not the only source of context information but the most prominent.

The already mentioned Cell-ID-method is just one method for obtaining a mobile device's position (for an overview see [21]): Within cellular networks the precision of locating can be further enhanced by considering the radio signal's runtime-differences to several base stations (lateration). Satellite position systems like the "Global Positioning System" (GPS) or the projected Galileo enable position determination with a much higher precision all over the world, but the costs for launching and maintaining satellite navigation systems are enormous and such systems cannot be used for indoor locating. That is why special indoor positioning systems like "Active Bat", "Active Badge" or WiFi-based systems like "RADAR" were developed. For this article it is just of relevance if the location can be determined by the mobile device itself (self locating) or by the network (remote locating).

But beside the fascination for the new services which are enabled by those technologies there are considerable concerns with regard to privacy problems: people are afraid of being "tracked" like in Orwell's dystopia when using mobile technologies. To underpin this statement we present the results of several studies in chapter five. One of these studies was conducted by ourselves: 44 % of the $N = 604$ respondents stated that they would change the way they use their cellular phone if it were possible to track them.

Because of these concerns it is no wonder that there is a great body of research discussing several misuse scenarios of location information and proposing technological solutions to prevent such misuse. However [5] criticizes that

"[...] current contributions are mostly based on example scenarios, and the assumptions made in each scenario about the knowledge available to the attacker are not always clear [...]". Our article thus wants to give an overview about various research results concerning location privacy. The reader is also referred to Liu's tutorial lecture about this topic which is available as a set of slides [23]; this tutorial is focused more on a detailed description of pertinent algorithms while our article gives a broader overview.

The remainder of the article at hand is organized as follows: In the second chapter we discuss several scenarios where location data is misused. Chapter three covers technical approaches to prevent the misuse of location data and shows which techniques can be applied for which scenarios. A short overview concerning some legal regulations concerning location data is given in chapter four. Chapter five introduces some results from pertinent empirical studies. In the last chapter we summarize and give an outlook to further research directions.

2. Specific Misuse Scenarios of Location Data

2.1. Assumptions

In the following article we make the assumption that LBS are provided by independent third parties, i.e. the LBS provider (LBSP) cannot determine the end user's position by himself and therefore there are two basic approaches to obtain location information about a user: In the first one a specialized location service is queried. This location service has to be authorized by the end user to provide his location information to the respective LBS provider. The second approach is that the location information is included in the request by the mobile device itself (e.g. using an attached GPS-receiver) or by a mediator that complements the request with location information before forwarding it to the LBSP. Having independent LBS providers is also a reasonable assumption from an economic perspective [11].

We further assume that the communication between mobile device, service provider and an optional mediator is encrypted. This means that the "attacker" is the LBSP or a LBSP whose backend systems are compromised. It could be also the case that a LBSP is required by law to provide collected location information to authorities for criminal prosecution (see chapter four). As an "attack" we regard the case that a LBSP gained more knowledge about a user than the user intended to let the LBSP know. This knowledge may be subject to uncertainty but the attack could reduce the uncertainty. There are approaches thinkable how the LBSP could obtain location information without the user's consent, e.g. (i) triangulation of the radio waves emitted by the mobile device, (ii) hiding a tracking transmitter in someone's clothes or car or (iii) employing a camera system to

identify all the users at a given location. However we do not consider scenarios that rely merely on such methods but rather misuse scenarios (attacks) where the location information was deliberately passed to the LBSP. The triangulation of radio waves can be circumvented by the so called spread spectrum technique [13]: the basic idea is to modulate the data signal based on a shared secret only known to the mobile device and the network operator so the signal is not discriminable from background noise without knowing the shared secret.

2.2. Direct and Indirect Location Privacy Problem

Knowing that user Alice visited location L may represent a privacy impairment because L may reveal Alice’s political attitude (if L is the headquarter of a certain political party), personal interests (if L is a place where certain things are traded or one can perform leisure activities), employer (if L is the premise of a company), her circle of friends and acquaintances (if L is a house where one of her friends lives) or health problems (if L is a specialized hospital). Knowing about Alice’s location one could even be able to intercept her to perform a hold-up.

We further can distinguish if the time when Alice visited L is known or not. Knowing the time increases the risk of privacy infringement because L may only be a suspicious place to be within certain time spans: L may be a room where only on Mondays the assemblies of a political party are held or L was the scene of a crime at a certain point in time. It’s also far less suspicious being inside a shopping mall at 2 p.m. than at 2 p.a. (when the mall is closed).

Knowing that Alice visited L several times may also be of interest for an observer: visiting a hospital once a year is far less meaningful than visiting the hospital every three days. Visiting the house where criminal Bob has found shelter just once Alice still can pretend that she just had to ask for the way but this pretext gets far less believable if Alice visited the house several times.

We see from these examples that being able to map location information to a certain person might be worrying. Since for many types of LBS it is possible that the LBSP works without knowing about the user’s identity (e.g. POI- or navigation services) it is an obvious idea to apply pseudonymization (e.g. [10]): in this approach a mediator replaces the user’s identity in the request by a pseudonym before forwarding it to the respective LBSP. After processing the request the LBSP delivers his response along with the pseudonym to the mediator which can then dispatch the response to the respective end user. The pseudonym can be a randomly chosen but unique bitstring of sufficient length. Even if a user doesn’t perceive his location information as sensitive he might use a pseudonym if there are

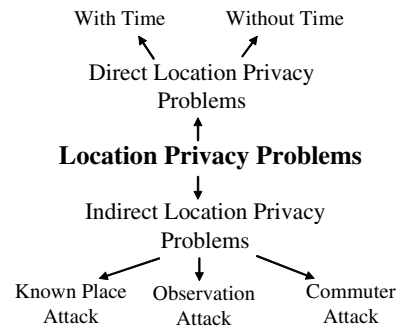


Figure 1. Classification of Location Privacy Problems

other kinds of personal information in his request he wants to conceal, e.g. profile information. Using a mediator and a pseudonym is also advantageous for push services, i.e. services where messages are sent to the mobile device without being directly requested (e.g. SMS-messages or push services), because this way unsolicited messages can be easily prevented: if a service provider doesn’t know a valid pseudonym he simply cannot push a message to a mobile device. But even if pseudonymization is applied there are privacy problems. These problems will be discussed in the next section.

The conclusion from this section’s considerations is that there are two classes of location privacy problems (see also the classification in figure 1): those with non-pseudonymized users (direct location-privacy problem) and those with pseudonymized users (indirect location-privacy problem).

2.3. Attacks on Pseudonyms - The Indirect Location Privacy Problem

There are scenarios thinkable where it is possible to infer about a pseudonymized user’s identity by evaluating the location information. We call this the ”indirect location privacy problem”. In these cases the location information in conjunction with other external information sources can be employed as ”quasi identifier” to resolve the identity behind a pseudonym.

In the case of the so called ”restricted space identification” [15] this external information is knowledge about places where certain persons typically stay, e.g. a residential house (for outdoor scenarios) or office room (for indoor scenarios). If a LBS-request comes from place L and L is the front garden of a certain house the attacker just can look up in a public directory (e.g. telephone book, maps) to find out who is living in that house. We therefore call this kind

of attack also "front garden" or "writing desk" attack. In [20] a study is described where several users could be identified by evaluation of their GPS tracking data. The external information they used were maps freely available in the internet. A similar study by [17] led to the identification of home addresses in 85 % of cases, but due to the design of the experiment the authors couldn't verify the addresses found.

The temporal-spatial sequence of places visited by a user can also constitute a quasi identifier, e.g. a commuter's daily way to get to work [6]. If a pseudonym is revealed by evaluation of a user's regular path we call this "commuter attack". Another attack on pseudonyms is the observation identification attack [15]. Here the attacker can observe a place or room (e.g. camera system) and recognize users. If Alice visits place L which is under such an observation and a LBSP receives a request from a pseudonymized user P coming from L than an attacker can infer that Alice is P . Meanwhile there are camera systems with face recognition capabilities¹. Such systems are still in their infancy but systems for automatic vehicle identification based on number plate recognition are widely employed at garages or for tolling² and therefore represent a great danger for the case of using a LBS (e.g. navigation service) while travelling in a car. But even if such an observer is not capable to identify Alice and is only able to recognise her between different visits of L this might represent a danger: If Alice started a new session with the same LBS and uses a new pseudonym for this the attacker is capable of linking the two pseudonyms.

3. Technical Approaches to Prevent Misuse of Location Data

3.1. Overview

In this chapter several technical approaches for supporting location privacy will be discussed. In figure 2 a classification scheme for this purpose is depicted. On the uppermost level we distinguish the following approaches: "policies", "modification of request", "dummy requests" and "provider change". The modification approach can be further divided into the two main branches "pseudonymization" and "deliberate impairment of locating". Pseudonymization was already mentioned in section 2.2 to introduce the indirect location privacy problem and will be discussed in more detail below. The impairment approaches again can be divided into "blind out", "reduction of precision" and "path crossing" as special cases.

¹e.g. www.crossmatch.com/facial_recognition.html

²e.g. www.licenseplaterecognition.com

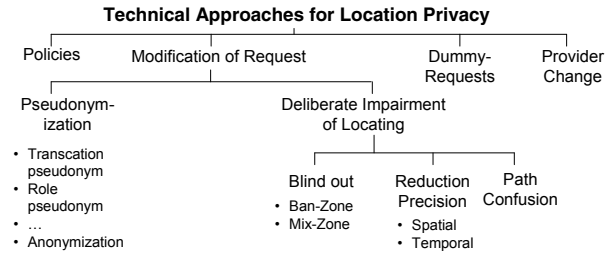


Figure 2. Classification of Different Technical Approaches for Location Privacy

3.2. Policy Approaches

Using policy approaches one tries to write down under which conditions which LBSPs are allowed to obtain which quality of location information from a certain user. One example rule of a policy could state that a certain LBSP is allowed to query a user's location on weekdays with a resolution of 100m but only if the user is in a certain part of the country. Without using policies it would be necessary to prompt a user for explicit allowance each time a LBSP wants to query his location information which would greatly impair the user experience, especially for LBS that require continuous location information. However, policy approaches alone are not sufficient since there is the need for further technical arrangements to guarantee the enforcement of the policies, e.g. a mediator that obeys a policy in his decision if a LBSP's request for a certain user's location information can be performed or not. Policies can be seen as a way to steer all the other technical approaches depicted in figure 2. In the remainder of this subsection we name some approaches to implement policy language for LBS:

Myles et al. [25] propose a policy approach where the decision if a LBSP will get access to a user's location information can depend on context parameters like time, the location itself (e.g. someone only wants to be located when being at his workplace), type of request, collocated people and the organization the LBSP belongs to. They employ a specification language based on APPEL, which is a part of the W3C's P3P (Platform for Privacy Preferences Project). APPEL is a XML-language to specify a user's privacy policies when browsing in the web, e.g. a user could declare that he doesn't want to access websites that collect the eMail-address for marketing purposes.

In Langheinrich's privacy awareness systems "pawS" [22] P3P and APPEL are also used to describe policies and requirements for privacy issues in ubicomp scenarios. In rooms where such services are available a privacy beacon announces what data will be collected for each service. The user's mobile device delegates this information to a privacy

proxy in the internet. This proxy is configured using APPEL and will contact the services and compare their policies with the preferences of the mobile user. A service might have even several policies which imply different quality levels depending on how much personal information the user is willing to disclose.

Besides the W3C there is the IETF as an association that works on non-binding standards for the internet. "Geopriv" is a working group of IETF committed to the field of privacy problems that arise from the use of LBS³. In their publication "RFC 4745" they propose "a document format for expressing privacy preferences" based on XML. A policy consists of several rules. Each of these rules again consists of a three parts: conditions, actions and transformations. If someone requests location information about the policy holder first the conditions are evaluated. Conditions describe entities that might query location information (e.g. a certain user or a LBSP) and possible places where the policy holder resides. If such a condition matches the associated action part says whether the requested location information can be sent to the receiver; the transformation part of a rule describes optional transformations of the requested location information before sending it to the requestor. Such a transformation might be the reduction of the location information's granularity.

3.3. Pseudonymization

How pseudonymization works was already introduced in section 3.2. There are different kinds of pseudonyms [26]: A pseudonym that is used for a certain purpose (e.g. usage of LBS) is called "role pseudonym". If a role pseudonym is used for just one transaction or session made up of several communication steps (e.g. turn-by-turn navigation service) we have a transaction pseudonym. A relation pseudonym is used for several transactions but just for the communication with a distinct partner; it could be used for a LBS where the LBSP has to remember a user's recent sessions so the user hasn't to reconfigure the LBS for each session. The usage of a person's pseudonym (e.g. e-mail-addresses and telephone number) is not bound to a certain purpose. Here one can distinguish if the identity behind the pseudonym is publicly known or not; these kinds of pseudonyms are called public or respective non-public person pseudonyms. Public person pseudonyms are not of interest for the purpose of data protection; they are more a mean to simplify communication.

The strongest form of pseudonymity is anonymity that is a pseudonym used only for one communication step (e.g. one "response-request" step) and thus acts as token to technically relate a SP's response to the respective request. But pseudonymization is not a LBS-specific technique. Using an e-mail-address can be a form of pseudonymization if the

recipient cannot find out about the identity behind the address.

One question with pseudonymization is what organization operates the mediator service that provides the pseudonymization. The end user should deem this organization as trustworthy. Since as already mentioned the MNO knows the approximate position of the user due to mobility management he could play the role as mediator. If the mediator is operated by an external party this arises the question of the business model because if money flows from the LBSPs to the mediator the thereby resulting dependence might impair the integrity of the mediator. It is therefore preferable to have the mediator organization reimbursed by the users directly or to have a non-profit organization as mediator (e.g. government agency or a non-profit league).

3.4. Deliberate Impairment of Locating

The category "deliberate impairment of locating" subsumes "blinding out", "reduction of precision" and "path crossing". Each approach will be discussed below in this section whereas for "reduction of precision" we also cover the special case of "k-anonymity".

Blinding Out. To "blind out" means just to block submission of location information to the LBSP. This approach could be seen as creation of artificial "dead zones" as there were an alleged area with no radio reception. For the direct location privacy problem such "ban zones" should cover "dangerous places". In the case of the indirect location privacy problem a ban zone prevents the front garden and the observation attack if it overlaps the front garden or the observed place. But especially for the latter case it might be not so easy to know about all observed places. A large enough ban zone can also prevent the commuter attack.

The naive approach when defining the bounding of such a zone would be to draw a circle around location L but in this case the attacker could guess the circle's centre. Therefore one reasonable approach for the definition of ban zones would be the one proposed by [20]: first an "invisible circle" with L as centre point is drawn and afterwards a random point in that invisible circle is chosen as the actual centre point of the ban zone. This method requires that enough other users have a "known place" in the ban zone.

"Mix zones" [4] are a special kind of ban zones for pseudonymized services which are shared by several pseudonymized users. As in ordinary ban zones the submission of location information to the LBSP is suppressed while the respective user resides within a mix zone but after leaving the zone his requests will be forwarded using a new pseudonym. The areas that are not covered by mix zones are called "application zone". The rationale behind mix zones is that given enough users entering and leaving a mix zone

³<http://tools.ietf.org/wg/geopriv>

a LBSP cannot follow a user’s trail through a mix zone. For example two users with the pseudonyms “*abc*” and “*def*” enter a mix zone at almost the same point in time and leave as users “*ghi*” and “*jkl*” the mix zone; the LBSP cannot be sure if “*ghi*” or “*jkl*” is the pseudonym for the user that was “*abc*” before entering the mix zone.

Reduction of Precision. “Reduction of precision” is another suggesting approach to enforce location privacy. There are many services which can be provided with relative imprecise location information, e.g. a service that sends information like location-aware news, weather forecasts, events or special offers (advertisements). For such services it is sufficient to have location information with a precision of several kilometers. If the location information is retrieved with a higher precision than necessary (e.g. a GPS-Receiver can provide a precision of up to several meters; CellID in GSM-networks of up to several 100 meters in city areas with many base stations) the precision can be reduced by adding “random noise” [20] or just by truncation or rounding the coordinate’s decimal places. The truncation or rounding method has one drawback when used for LBS which requires continuous location information: for example when all of the x-coordinate’s decimal places are truncated and the resulting value “jumps” from “2” to “3” the LBSP can infer that the user just crossed the line between “2.9” and “3.0” for the truncation case respective between “2.4” and “2.5” for the rounding case.

There are also scenarios where it is reasonable to reduce not or not only the spatial precision of the location information but also the temporal precision, which means spatial and temporal reduction of precision are orthogonal approaches [15] (see also figure 3). There are even examples for services where the highest possible spatial precision of location information is preferable (e.g. retrieved by a GPS receiver) but the delivery of the location information to the LBSP can be deliberately delayed for several hours or even months.

One example of such a service would be the “road hazard detection” service mentioned in [15]: a mobile device built into a motor vehicle records “near accident” situations, e.g. emergency stops along with the position where the incident took place. These records are then transmitted once a month or quarter to an agency (LBSP) that analyzes the collected data to detect places in the road network with accumulations of near-accident situations. If such places are detected the road network can be improved accordingly (e.g. placement of warning signs or traffic lights, changing the course of the road). Other examples of such services are location aware mobile blogging (supplementing a blog entry with the current position), virtual graffiti/memos (depositing virtual messages at current position and being able to read messages deposited by other users in the nearer surround-

		Spatial precision	
		High	Low
Temporal precision	High	Turn-by-Turn-Online Navigation, POI-Finder, Tourist-Guide	Weather Warner, Time-Critical Advertisement
	Low	Mobile Blogging, Virtual Grafitti/Memo, Road Hazard Detection, Mobile Data Gathering	Location Aware News, Weather Forecast

Figure 3. Examples for LBS to show that temporal and spatial precision are orthogonal

ing) or data gathering (e.g. surveying passers-by, reading off electric meters). There is also the opposite case of LBS that can be offered with relatively imprecise location information but where a real-time delivery of location information to the LBSP is required, e.g. a service that warns a user of an approaching thunderstorm or informs him about time-critical local offers (e.g. special limited offer by a local department store with expected high demand).

But when reducing the precision of location information one should keep in mind that even rather imprecise location information could pose risks for the user. This is the case when from the location information the attacker can conclude that the user didn’t stay at a place he was obliged to stay. For example even if the location information is so imprecise that it only says in which country the users stay this still could lead to problems if it was anticipated that the user is in another country on official journey.

k-Anonymity. Another way of spatial precision reduction is to provide a rectangular area (so called “cloak-box”) to the LBSP as location information [15]. The dimension of this cloak box is calculated to ensure that k (potential) users (with $k > 1$) where staying within the area described at the time of the respective service request. This provides k -anonymity which means that a user is indistinguishable from the $k - 1$ other users so a higher k -value should provide a better degree of anonymity.

The first algorithm for this problem with regard to location privacy was described by Gruteser & Grunwald [15] and is based on quadtrees, a data structure originally designed to index spatial data in database systems. Following this approach the whole area is divided into a partition of four quadrants with equal size. The quadrant with the requesting user is then again divided into four sub-quadrants. This is repeated as long as the quadrant with the respective user contains at least $k - 1$ other users. If this restriction is violated the algorithm returns to the previous quadrant

which is returned as result. Based on real-world road networks and traffic count statistics the algorithm was evaluated in a simulation: for areas with a high density of vehicles the median precision provided by the cloak boxes was 30 to 65 meters; for other areas the resolution decreases up to 250 meters. Altogether the authors report that the median resolution of the cloak boxes was 125 meters.

A more sophisticated algorithm for the k -anonymity problem is the CliqueCloak-algorithm developed by Gedik & Liu [14]. The algorithm first builds a graph with all pending requests, where two requests depicted as vertices are connected by an edge if they could potentially be grouped into the same cloak box without violation of the precision constraints of the respective requests. Putting as many requests as possible into one cloak box is then the problem of finding so called cliques, i.e. subgraphs where each node is connected with all other nodes. The algorithm not only allows customizable precision constraints but also allows setting an individual k -value for each user. Gedik & Liu evaluated their algorithm in a simulation similar to the one for the quadtree-algorithm. They generated user requests with randomly chosen values for the spatial and temporal precision demands (mean value 100 m respective 30 seconds) as well as for the k -value (between 2 and 5). For 90 % of the requests the algorithm was able to find an appropriate cloak box. If no such cloak box can be found the request has to be discarded.

One essential question is that for an appropriate value for k . Gruteser & Grunwald mention $k = 5$ as "a fair level of anonymity". When a LBS user is afraid of getting in the sight of a criminal prosecution based on his location information a value of $k = 2$ should be sufficient in a constitutional state because of the legal right of "in dubio pro reo" (presumption of innocence): when there was at least another potential user at the scene of crime at the time of matter the judge cannot convict him alone on his location information.

Beside k -anonymity there is also the proposal for other parameters to measure the quality of a cloak-box, e.g. l -diversity: the parameter l describes the number of locations (e.g. buildings, landmarks) that are within a cloak-box [23]. The idea behind this concept is that a cloak-box might not provide sufficient anonymity if all of the k users are gathering at the same place (e.g. meeting in a room or public event). Several algorithms to compute cloak-boxes that fulfil certain k - and l -values are part of the "PrivacyGrid"-framework [1]: The name comes from the grid that is used to cover the whole reference space. The "top-down"-variant of the algorithm starts with the grid-cell that contains the calling user. If this cell doesn't contain $k - 1$ other users or l locations it is consecutively enlarged by adding adjacent cells.

Improving the Results of Cloaked Queries When spatial obfuscation is applied this brings along a reduced quality of service, of course. For LBS whose purpose it is to return a list of nearest POIs this problem can be tackled in the following way [24]: the LBSP returns all POIs that would constitute a correct answer for a user at any position within the cloak-box. Afterwards a mediator or a mobile client application with awareness of the actual user position can filter out the inappropriate ones. This "subsequent rectification" method leads to processing overhead since many of the results will be discarded but allows to provide the quality of service while protecting a user's location information.

Path Confusion. One special case for impairing location information is the "path confusion" approach described in [16]: the approach aims at service scenarios that require the service provider to be continuously aware about a mobile user's position. The basic idea is to perturb the location information only for a short period because long term perturbation would impair the quality of the service noticeably. When two users' paths meet the algorithm can decide to perturb the location information of these users so the observer of the location information gets the impression that the user crossed their paths.

Discussion. Reduction of spatial and/or temporal location precision are applicable for the direct location privacy problem, e.g. if the spatial precision is reduced the LBSP cannot find out if Alice visited the meeting room of that political party and if the temporal precision is reduced the LBSP cannot find out if Alice's visit of place L was at the time when at L a crime was committed. The reduction approaches have also the potential to prevent all described indirect location privacy problems. The drawback however is that reduced precision of locating might lead to reduced quality of service.

3.5. Further approaches

Dummy Requests. There is also the idea of generating dummy request using simulated users along with realistic movement patterns, e.g. the system developed by Kido et al. [19]. They describe an implementation where the dummy-requests are computed on a mobile device, but a stationary third-party could also do so. When there are dummy requests all kinds of attacks on pseudonyms are hampered because the attacker has first to distinguish between real users and dummy users. Even for non-pseudonymized services dummy-requests can be reasonable because then the end user can deny of having visited a certain place where something suspicious or even unlawful took place ("it wasn't me, it was a dummy!").

But if a LBS isn't free of charge there is the problem

who will pay for the dummy users' requests; maybe the real users are willing to pay the additional charge but if there are k dummies per users it would increase the service charge by the factor of $k + 1$. If the LBSP can provide the service without variable costs per sessions or request (which should be the common case for a digital service) he should be willing to offer the service for a reduced rate of $1/(k + 1)$ because he loses no potential revenue by offering the service for free to dummy users without spending power. This consideration might not work if the LBSP needs content from other providers ("white label content", e.g. maps, weather or traffic information) to enrich his service and has to pay for each request.

Provider Change. A final approach to mention is the deliberate change of the respective LBSP for one service from time to time. An example would be to use the navigation service by LBSP *A* for one hour and than switching to the navigation service by LBSP *B*. This approach requires that for a given type of service there are several providers that can deliver a comparable service and thus the approach cannot be applied for very special service only provided by one LBSP. As far as we know this approach wasn't analyzed in detail yet.

Commuter attacks can be prevented by this method. For all other attacks of the direct as well as the indirect location privacy problem this approach reduces the misuse potential because one LBSP sees only a fraction of a user's path.

3.6. The Roles of Mediators

Most technical approaches for the enforcement of location privacy found in literature make use of a mediator party. A mediator is necessary for all forms of pseudonymization, because in this case the mobile device cannot communicate directly with the LBSP or otherwise the LBS would at least figure out the user's IP-address. The Mix-Zone-approach therefore also requires a mediator because mix zones only work for pseudonymized users.

The usual approach to implement k -anonymity is based on some kind of mediator because it is necessary to have a central entity that is aware of the locations of several users so their requests can be grouped together to reach the desired k -values. This is also the case for the "path confusion" technique. In [8] an approach is described where the cloak-box is calculated by direct collaboration of mobile devices in a peer-to-peer-fashion.

The remaining approaches could be implemented without a mediator entity: a mobile device with self-locating capabilities could alter its location information (e.g. adding random noise, truncation of decimals) before submitting this information to a LBSP. Temporal reduction of precision is also trivial to implement because the mobile device

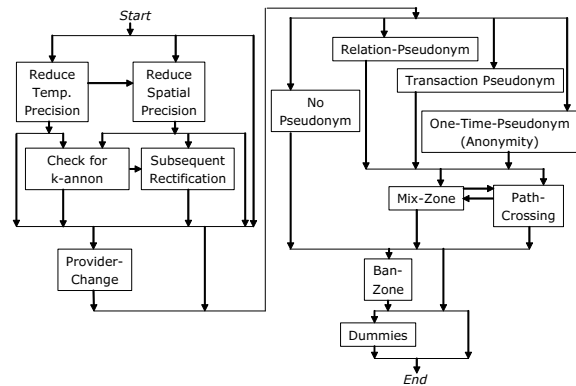


Figure 4. Decision Path Diagram

has just to wait for some time before submitting the location information. "Blinding out" can also be performed on the mobile device alone as well as the generation of the dummy requests; indeed the above mentioned system developed by Kido et al. [19] follows this approach. They even describe a special encoding to reduce the size of a request containing several dummy locations. The "provider change"-method can be also implemented locally on the mobile device if the device has a list of several service providers for the same service.

3.7. Combination of Approaches

The decision which methods for location privacy should be applied has to be made on a case by case decision for each LBS. This can be exemplified by the decision path diagram depicted in figure 4: Each path from "start" to "end" represents a possible combination of techniques for location privacy. First we have to decide if spatial and/or temporal precision reduction can be applied; it is also possible to combine both. If we opt to reduce the spatial precision it is maybe possible to apply the "subsequent rectification" approach. K -anonymity can be applied for both spatial and temporal precision reduction. We then have to decide if the "provider change" method can be applied. Afterwards it is decided if some kind of pseudonymity can be applied or not. For pseudonymized services we have the possibility to apply the "mix zone" and/or the "path crossing" approach. "Ban zones" and "dummies" can be applied whether a pseudonymization technique is applied or not.

4. Legal Regulations

One might consider legal regulations as instrument for the enforcement of location privacy. However, there are several drawbacks: Laws are only of use if the offender is

caught indeed and put in front of a judge. In some cases the "attacker" might even be a country's executive force, e.g. in the case of criminal prosecution. Legislation will differ from country to country. We therefore only mention two directives of the EU:

The European Unions directive "2006/24/EC" on "[...] the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks [...]" demands from the EU's member states to implement national legislation to oblige MNOs to store "[...] data identifying the geographic location of cells by reference to their location labels (Cell ID) [...]" for at least six month. This regulation is obviously motivated to support criminal prosecution. It should be mentioned that parties acting as mediators (e.g. to pseudonymize user requests) are obliged to log and store communication data so it can be resolved which pseudonym was assigned to a certain user. This rule can be circumvented by operating mediation services outside the EU or by chaining several mediators.

But other EU directives are concerned with protecting the privacy of LBS users [27]. Article 9 of the directive "2002/58/EC" makes the statement that location information other than "traffic data" (e.g. information location required for mobility management) is only processed when pseudonymized or with the prior consent of the respective user. To give his consent the user has to be informed what kind of location data and for which purpose will be processed. At any time the user can withdraw his consent.

If an LBS is deployed within an enterprise (e.g. to support mobile workers) in some countries (e.g. Germany) this requires the agreement of the labor council because LBS are a measurement that could be used to supervise the behaviour of employees.

5. Empirical Studies

5.1. Own Study

In October 2007 we had the opportunity to ask the users of an eLearning system at our university a few questions related to the usage of mobile technologies. One of these questions asked how they would change the usage of their cellular phone if it were possible to locate them with a precision of about 100 Meters. This is a "trick questions" because the MNO's mobility management requires determining the base station in use so this is already possible. It is therefore remarkable that only 56 % of the total of $N = 604$ users that answered this question stated they would use their cellular phone as they already do. The others stated they would change the way how they use their phone: 33 % would turn it off more often, 7 % would turn it on only for

making telephone calls and 4 % would even abolish their cellular phone.

Another question asked for how many hours per day on average they had their cellular phone switched on (stand-by state). This question was answered by $N = 606$ users. The average of the values stated is 18,6 hours whereby 56 % said they had turned on their cellular phone for 24 hours. These long "stand by"-times imply a high risk for supervising someone by tracking his cellular phone.

5.2. Further Studies

Danezis et al. [9] conducted a study with $N = 74$ participants to find out about the perceived monetary value of location privacy. They pretended to be able to determine user's position for 28 days all around the clock with the Cell-ID approach and asked students for which amount of money they would participate in such a study. The students could give a (sealed) bid. Since it was assured that the n lowest bids would be accepted there was a strong motivation to mention the actual amount of money deemed as a fair compensation. The median of the amount of money determined in this study was 10 GBP. Students which did travel more tended to have higher prices to "sell" their location privacy.

In an often cited study Barkhuus & Dey [3] asked $N = 16$ users about how they perceived risk after using LBS for five days. One of the results was that LBS capable of performing "remote locating" were rated as more dangerous than LBS with "self locating". A result of another study conducted by Barkhuus was that users perceived the risk of using LBS higher before using such services than during using such services [2].

Junglas & Spitzmüller [18] employed a structural equation model ("PLS") to analyze how personal traits influence the intention to use LBS. In their study $N = 470$ students participated. The factor with the highest influence according to their results is "perceived usefulness"; the factory "perceived privacy", "perceived risk" and "trust" influence the intention to use such a service at approximately the same degree but weaker than "usefulness".

6. Summary and Outlook

In the article we first explained why Location Based Services (LBS) are an important class of mobile services. We then discussed several misuse scenarios of location information when using such services. To prevent misuse of location information there are several technical approaches which were classified and discussed afterwards. We also covered the legal situation with regard to location privacy problems and presented empirical work that shows that users have specific concerns when using LBS.

Current research in the field of the location privacy works on problems that occur in location-aware community service [27], e.g. services like friend-finder. The salient question with community services is to decide whether another user is allowed to obtain location information. Since location privacy is a complex subject it is a challenging task to provide tools for end users which enable them to configure mobile technologies and location privacy approaches in a way to suit their needs.

References

- [1] B. Bamba and L. Liu. PrivacyGrid: Supporting Anonymous Location Queries in Mobile Environments. Technical Report GIT-CERCS-07-17, Georgia Institute of Technology, 2007.
- [2] L. Barkhuus. Privacy in Location-Based Services, Concern vs. Coolness. In *Proceedings of the Workshop on Location System Privacy and Control at Mobile HCI 2004*, 2004.
- [3] L. Barkhuus and A. K. Dey. Location-based Services for Mobile Telephony: A Study of Users' Privacy Concerns. In *Interact 2003, 9th IFIP International Conference on Human-Computer Interaction*, Zürich, Switzerland, 2003.
- [4] A. R. Beresford and F. Stajano. Mix Zones: User Privacy in Location-aware Services. In *Proceedings of the IEEE Workshop on Pervasive Computing and Communication Security (PerSec)*, pages 127–131. IEEE, 2004.
- [5] C. Bettini, S. Mascetti, X. S. Wang, and S. Jajodia. Anonymity in Location-Based Services: Towards a General Framework. In C. Becker, C. S. Jensen, and J. Su, editors, *MDM*, pages 69–76. IEEE, 2007.
- [6] C. Bettini, X. S. Wang, and S. Jajodia. Protecting Privacy Against Location-Based Personal Identification. In *Proceedings of the Conference on Secure Data Management 2005*, pages 185–199, Trondheim, Norway, 2005.
- [7] Canals. Mobile GPS Navigation Market Doubles Year-on-Year. Research Report 2006/081, Reading, U.K., 2006.
- [8] C.-Y. Chow, M. F. Mokbel, and X. Liu. A Peer-to-Peer Spatial Cloaking Algorithm for Anonymous Location-based Services. In *Proceedings of ACM-GIS '06*, pages 171–178, Arlington, Virginia, USA, 2006.
- [9] G. Danezis, S. Lewis, and R. Anderson. How Much is Location Privacy Worth? In *Proceedings of the Fourth Workshop on the Economics of Information Security*. Harvard University, 2005.
- [10] M. Decker, R. Bulander, G. Schiefer, and B. Kölmel. A system for mobile and wireless advertising. In *Mobile Information Systems II*, pages 287–301, Leeds, U.K., 2005.
- [11] M. Decker, K. Issel, J. Mrozik, and G. Schiefer. The Role of Small and Medium-Sized Enterprises in Repeating the Success of the Internet in the Wireless World. In *Proceedings of eChallenges 2007*, pages 1442–1449, Den Haag, Netherlands, 2007.
- [12] A. K. Dey. Understanding and Using Context. *Personal and Ubiquitous Computing Journal*, 5(1):4–7, 2001.
- [13] H. Federrath, A. Jerichow, D. Kesdogan, and A. Pfitzmann. Security in public Mobile Communication Networks. In *Proceedings of the IFIP TC 6 International Workshop on Personal Wireless Communications*, pages 105–116, Prague, 1995. IFIP.
- [14] B. Gedik and L. Liu. A Customizable k-Anonymity Model for Protecting Location Privacy. In *Proceedings of the 25th International Conference on Distributed Computing Systems (IEEE ICDCS 2005)*, pages 620–629, Columbus, Ohio, USA, June 2005. IEEE.
- [15] M. Gruteser and D. Grunwald. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In *Proceedings of the First International Conference on Mobile Systems, Applications and Services*, pages 31–42. ACM, San Francisco, CA, USA 2003.
- [16] B. Hoh and M. Gruteser. Protecting Location Privacy Through Path Confusion. In *SECURECOMM '05: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks*, pages 194–205, Washington, DC, USA, 2005. IEEE Computer Society.
- [17] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Enhancing Security and Privacy in Traffic-Monitoring Systems. *IEEE Pervasive Computing*, 5(4):38–46, 2006.
- [18] I. A. Junglas and C. Spitzmüller. A Research Model for Studying Privacy Concerns Pertaining to Location-Based Services. In *Proceedings of the 38th Hawaii International Conference on System Science*, page 180b, 2005.
- [19] H. Kido, Y. Yanagisawa, and T. Satoh. An Anonymous Communication Technique Using Dummies for Location-based Service. In *Proceedings of the IEEE International Conference on Pervasive Services 2005 (ICPS2005)*, pages 88–97, Santorini, Greece, 2005. IEEE.
- [20] J. Krumm. Inference Attacks on Location Tracks. In *Proceedings of the Fifth International Conference on Pervasive Computing (Pervasive 2007)*, Toronto, Canada, May 2007.
- [21] A. Küpper. *Location-based Services – Fundamentals and Operation*. John Wiley & Sons, Chichester, U.K., 2005.
- [22] M. Langheinrich. A Privacy Awareness System for Ubiquitous Computing Environments. In G. Borriello and L. E. Holmquist, editors, *4th International Conference on Ubiquitous Computing (UbiComp 2002)*, number 2498 in LNCS, pages 237–245. Springer-Verlag, Sept. 2002.
- [23] L. Liu. From Data Privacy to Location Privacy: Models and Algorithms. In *VLDB*, pages 1429–1430, 2007. Slides available at www.vldb2007.org.
- [24] M. F. Mokbel, C.-Y. Chow, and W. G. Aref. The New Casper: Query Processing for Location Services without Compromising Privacy. In *VLDB '06: Proceedings of the 32nd International Conference on Very Large Data Bases*, pages 763–774. VLDB Endowment, 2006.
- [25] G. Myles, A. Friday, and N. Davies. Preserving Privacy in Environments with Location-Based Applications. *IEEE Pervasive Computing*, 2(1):56–64, 2003.
- [26] A. Pfitzmann and M. Köhntopp. Anonymity, Unobservability and Pseudonymity: A Proposal for Terminology. In *International Workshop on Design Issues in Anonymity and Unobservability*, Berkeley, CA, USA, 2000. Springer, Heidelberg.
- [27] E. I. Tatli. Extending P3P/Appel for Friend Finder. In *Proceeding of the International Workshop on Privacy-Aware Location-based Mobile Services (PALMS)*, pages 243–247, 2007.