2013 IEEE 14th International Conference on Mobile Data Management

Privacy challenges in third-party location services (Invited paper)

Maria Luisa Damiani
Dept. of Computer Science, University of Milan
Milan, Italy
Email: damiani@di.unimi.it

Colette Cuijpers
TILT, Tilburg University
Tilburg, NL
Email: C.M.K.C.Cuijpers@uvt.nl

***Abstract*—The concern for location privacy in mobile applications is commonly motivated by a scenario in which a mobile device communicates personal location data, i.e. the device holder location, to a third party e.g. LBS provider, in exchange for some information service. We argue that this scenario offers a partial view of the actual risks for privacy, because in reality the information flow can be more complex. For example, more and more often location is computed by a third party, the *location provider*, e.g. Google Location Service. Location providers are in the position of collecting huge amounts of location data from the users of diverse applications (e.g. Facebook and Foursquare to cite a few). This raises novel privacy concerns. In this paper, we discuss two issues related to the protection from location providers. The first focuses on the compliance of emerging location services standards with European data protection norms; the latter focuses on *hard* privacy solutions protecting from untrusted location providers.**

***Keywords*-Location privacy, geolocation services, W3C geolocation standard, data protection directive, LBS**

I. Introduction

Individual location is an enabling factor in a variety of mobile applications such as LBS (location-based services) and mobile sensing: LBS provide spatial information upon location based spatial queries (e.g. where is the closest restaurant?); mobile sensing enables the collection of geo-referenced data from sensor-equipped mobile phones (e.g. air quality). In all of these situations the user's location is communicated to some other party. As location can reveal details of one's personal life, such communication may result into a loss of user's control over personal data. In the context of information privacy [6], such loss of control translates into loss of privacy. In what follows, we refer to location privacy as the capability of controlling the way personal location is disclosed and used by a remote third party.

A large body of literature on location privacy focuses on privacy protection from untrusted application providers, e.g. a honest-but-curious LBS provider can take advantage of the location information it receives upon a request of service. To prevent such a risk, the solution commonly adopted is to map the actual location of the mobile device (client), obtained from e.g. GPS, into a different representation which somehow masks the true location. For example, location can be mapped onto a coarse location, a fake location, or be encrypted or even suppressed. Once transformed (and not suppressed) the client location is conveyed to the application provider.

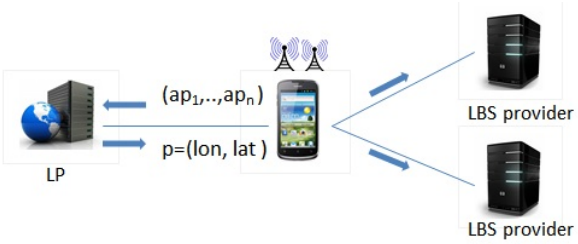

Figure 1. Architecture: the client transmits to the LP contextual information (e.g. set of Wi-Fi access points, ap_1,..ap_n) to obtain location p which is then conveyed to some LBS provider

Quite different is the situation in which the client requests the location from a third party location provider (LP), as illustrated in Figure 1. Note that location services are currently offered by major IT players, e.g. Google, Microsoft, Apple, Skyhook Wireless and used in a myriad of apps and Web applications. The protocol is as follows: the client senses the communication infrastructure in proximity and transmits to the LP e.g. the Wi-Fi access points and GSM/CDMA based stations being detected. This information is then matched onto a database of geo-referenced network components and used to estimate the location which is finally returned to the client. The advantage of this architecture is that consumer devices, e.g. smartphones, tablets, notebooks, can be located pervasively both in indoor and indoor settings, and across urban and rural areas, with an accuracy which can be of a few tens of meters.

It is evident, however, that in such scenario users' locations are disclosed not only to the application provider but also to the LP. This raises the question of how to ensure location privacy from the LP. Note that the aforementioned location privacy techniques (i.e. the LBS scenario) are not of help because the communication protocol as well as the semantics of the communication is different.

978-0-7695-4973-6/13 $26.00 © 2013 IEEE
DOI 10.1109/MDM.2013.67

In this paper we briefly describe recent research conducted on this topic along two different directions. The first direction focuses on the compliance of a geo-location standard, i.e. the W3C geo-location API for accessing location services, with European data protection norms. This problem is of practical relevance because the standard is widespread used while its impact on privacy is probably not widely clear yet. The second direction is more research oriented and explores a possible approach to minimizing the communication with potential untrusted LPs. In this case, the challenge is to provide comprehensive protection of location from both the LP and the application provider. The rest of the paper is organized as follows: Section II provides background knowledge on location systems and data protection regulation in Europe; Section III introduces the two aforementioned research directions; Section IV reports final considerations.

II. Location techniques and privacy norms

A. Location systems

The location of a mobile device, i.e. a smartphone, can be estimated using a multiplicity of techniques [5]. Crucial for privacy is the distinction between handset-based and network-based solutions: in the former case the location is computed by the mobile device itself, in the latter by a third party, e.g. telecommunication operator. The most popular handset-based positioning technique is GPS. GPS provides worldwide coverage and supports a range of location services with accuracies that range from a few meters to a few millimeters. However, when integrated in mobile phones, GPS localization presents severe limitations i.e. it is power consuming, moreover the device must be located in a position in line of sight with at least 4 satellites for the location to be estimated. This means that GPS cannot be used inside buildings, underground and in the so called urban canyons [5]. In those environments network-based localization techniques are appropriate, such as cellular (GSM/CMDA) and Wi-Fi (802.11) based systems. Taken singularly, all of these technologies are limited. Recent hybrid location systems overcome these limits integrating different technologies, e.g. cellular, Wi-Fi and IP-based positioning, to offer unprecedented opportunities in terms of location coverage and accuracy. This explains the increasing concern for business models offering location as a service and that is what we focus on.

B. Data protection in Europe.

In view of what will be discussed shortly, in particular the compliance of standards with European privacy regulation, we briefly describe the general Data Protection Directive and the so-called ePrivacy Directive, the main pillars of the EU legal framework regarding processing of personal and location data.

Data Protection Directive (95/46/EC). The general Data Protection Directive (hereafter: DPD) consists of a layered system of three levels. The first level is the general level that applies to all processing of personal data. The second level is applicable when sensitive data are being processed. The third level is applicable when personal data are being transferred to third countries. The layered system is cumulative, meaning that if sensitive data are being transferred to third countries, all three levels apply. The DPD applies to the processing of personal data which is defined as any information relating to an identified or identifiable natural person (data subject), while processing covers any operation or set of operations which is performed upon personal data. Both concepts are interpreted in a very broad manner. Besides the three levels within the DPD, the EU legal framework on data protection consists of two more levels of protection. Of these, the fourth level concerns sector specific regulations, such as Directive 2002/58/EC (ePrivacy Directive) which is relevant for location data.

The ePrivacy Directive (2002/58/EC) The provisions of Directive 2002/58/EC (as amended by 2009/136/EC) particularize and complement the DPD in the field of electronic communications services. Importantly this Directive lays down rules regarding the processing of location data, in particular it states that: “location data may only be processed when made anonymous or with prior consent and only for the duration necessary for the provision of a value added service”, being: “any service which requires the processing of traffic data or location data other than traffic data beyond what is necessary for the transmission of a communication or the billing thereof”. Prior consent and anonymisation are the only valid grounds for processing location data. In practice, the only valid ground in most cases will be prior consent as the Art. 29 WP has stipulated that “true anonymisation is increasingly hard to realize and (. . .) the combined location data might still lead to identification” [8]

III. Soft and hard location privacy solutions

Given this technological and normative framework, we consider two issues related to the use of third party location services. In the first case the data controller, i.e. the third party who receives the data, is trusted; in the second case, is untrusted. Note that when the data controller is trusted the data protection goal is to ensure that users are aware of how their personal data are used. Conversely if the data controller is untrusted, the data protection goal is to minimize the transfer of data so as to reduce the need of trust. These two situations are at the basis of the concepts of soft privacy and hard privacy discussed among the others by Danezis [3]. This distinction can be easily transposed to location privacy. Accordingly, *soft* (location) privacy solutions include privacy policies, users’ consent, audit controls and

so on. *Hard* (location) privacy solutions include privacy-enhancing techniques such as location obfuscation methods and cryptographic solutions. The two approaches discussed in what follows exemplify a soft privacy and hard privacy solution respectively.

A. Soft privacy: location standards

The W3C organization has proposed a standard API (Application Programming Interface) to request location services from a Web application [7]. Specifically the so called W3C geo-location API (simply API hereinafter) is *recommended standard* since May 2012. This API provides the abstract specification of a set of operations, which embedded in Web pages, enable to estimate the location of the users visiting the Web site. The API is coupled with HTML5 and supported by all major Web browsers. Moreover it is used by popular applications such as Foursquare and Facebook.

This standard prescribes that users must give explicit consent to the computation of their location. For example Figure 2 shows the home page of a geo-enabled Web site, called *wayn* (http://wayn.modap.org), developed as case study. This application uses the geo-location API to estimate the location of each user visiting the Web site from any device, mobile or not, and stores this information together with a few additional information in a repository on the server. Figure 2 highlights the request of user's consent. The research question we address in [1] is whether

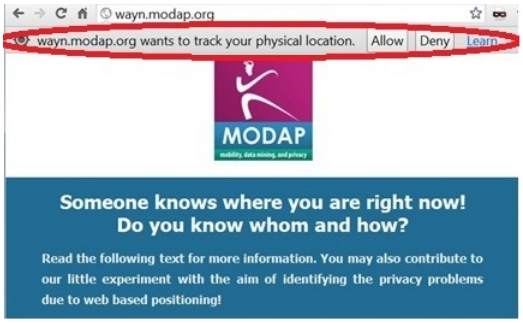

Figure 2. As the Web page is accessed the user is prompted with the request of consent (highlighted by the red ellipse)

the privacy mechanisms offered by the W3C geo-location API for the collection of personal location data (i.e. user's consent), is sufficient to comply with data protection norms in Europe. Indeed the question is of practical relevance as the use of standard location services aligned with privacy regulations would make the collection of mobility data easier. In what follows we briefly describe key features of the API and summarize major findings.

The W3C API. The API is agnostic of the positioning technology, i.e. the location can be requested without specifying how to estimate the location, e.g. using GPS, Wi-Fi based positioning or IP addressing. This simplicity of use is however paid for in term of flexibility and transparency, because the application provider is not in the position of exercising any control on the way the location is computed, in particular on whether the location is computed locally or by a third party while this is important for privacy. Indeed, what happens in practice is that the Web browser translates such an operation into a geo-location service request for the LP. Accordingly, any time the user is located the location is communicated to the LP.

Privacy analysis. Back to the initial question, i.e. whether the privacy mechanism of the W3C is aligned with data protection norms in Europe, this is a complex issue which calls for an interdisciplinary analysis. The analysis conducted in [1] assumes the applicability of the ePrivacy Directive without entering into the discussion whether or not the ePrivacy Directive is applicable. The two major issues regard the nature of user's consent and the role of data controller. According to the DPD the data controller is the person or body which alone or jointly with others determines the purposes and means of the processing of personal data. The data controller is responsible for full compliance with data protection regulations.

Ideally the data controller is the application provider. What happens in reality is that the application provider is not aware of who is going to compute the client location, because that depends on the Web browser chosen by the user. Therefore, the data controller does not have control on location processing while the qualification of the Web browser and the LP under the Data protection and the ePrivacy Directive is problematic as they do not fall under the strict authority of the application. This issue is especially problematic in view of the legal requirement of consent. Consent is only valid if specific, freely given, and informed. As the application provider is not in a position to properly provide information, as he himself is not aware of the specifics of the processing determined by user's Web browser, the legal basis for the processing falls to pieces. In other words, making it illegal. Besides this major issue, the analysis leads to a whole array of legal questions while suggesting technical enhancements of the API [1].

B. Hard privacy: protecting location from untrusted LPs

Let us turn to consider the case in which the LP is not trusted. We recall that in our scenario the LP computes the client location upon request, based on contextual information, e.g. the Wi-Fi networks in proximity. One could argue that the location could be processed locally so to avoid any privacy issues, like for example in the Intel PlaceLab [5]. Unfortunately PlaceLab-like architectures conflict with the dominant business model offering location services for free in exchange for location data. We have thus devised an approach which seeks to minimize the communication with the LP.

Minimizing the interaction. The idea behind [2] is that the amount of information that the user transmits to the LP exceeds what is really necessary to determine the users' location. In fact every time a service is requested from a given place, e.g. home, the client transmits the same or similar contextual information, e.g. Wi-Fi access points. One can thus observe that if clients would acquire the capability of recognizing autonomously the places that have been already visited, the location information would only be requested to the LP when it is strictly necessary. As a result, the communication would be minimized. To implement this idea, we devise an approach based on the metaphor of *private place*. Private place is an abstraction which conceptualizes the intuition that there are some regions of space that belong to the personal sphere. The intuition is that whenever the user is in a private space, the location should not be disclosed to the LP. In order to recognize whether the position is inside or outside a private place, using a consumer device, without interacting every time with the LP we take inspiration from previous approaches such as [4] to develop a solution which associates every place a radio fingerprint, specified in terms of Wi-Fi access points. Private places are recognized by comparing the networking infrastructure detected in a point, e.g. the Wi-Fi access points, with the set of radio fingerprints.

Privacy rules Minimizing the interaction with the LP, however, does not forestall the disclosure of the private place to the application provider. Every time, the user requests a service from, say, home, where home is a private place, the position conventionally associated with the private place at the time the place is defined, is disclosed to the application provider (conversely the service could not be requested). Therefore if the application provider is untrustworthy or collude with the LP, location privacy is again at risk. To achieve a comprehensive protection of location from both the LP and the application provider, the approach is to use privacy rules. An example of privacy rule is the following:

$$Home, [19:00, 08:00] \rightarrow cityOf(Home)$$

$Home$ is the name of a private place. The rule means that when the user is at home during the night, the location communicated to the application provider is the city in which home is located.

Architecture. The architecture of the system is illustrated in Figure 3. It consists of two main building blocks, called Place Handler and Policy Handler, respectively. The Place Handler provides the user with a set of functionalities to create private places and to automatically recognize whether the user is inside or in proximity of one of the private places previously defined. The Policy Handler enables the specification and enforcement of privacy rules. When a location is to be communicated to the application provider, the system first checks whether the mobile device is located in a private place and if it is so, enforces the privacy policy.

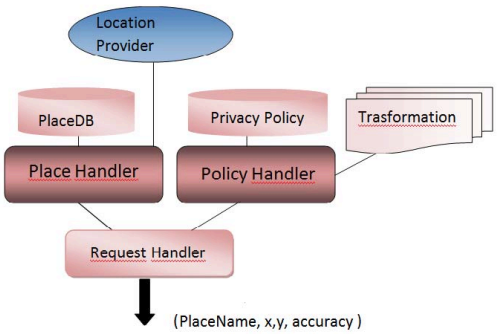

Figure 3. Privacy-enhanced location system: client architecture

IV. Final remarks

Location privacy requirements and solutions much depend on the characteristics of the applications. Nevertheless, a conceptual framework on location privacy general enough to provide guidance across different typologies of applications is still lacking. Building such a framework is a major challenge for future research on location privacy.

References

[1] M.L. Damiani and C. Cuijpers. Privacy-aware geolocation interfaces for volunteered geography: a case study. In *Proc. ACM SIGSPATIAL Workshop GeoCrowd*, 2012.

[2] M.L. Damiani and M. Galbiati. Handling user-defined private contexts for location privacy in LBS. In *Proc. ACM GIS '12*, 2012.

[3] G. Danezis. An Introduction to Privacy Technologies. http://www.cl.cam.ac.uk/teaching/1112/SecurityII/2012-security2-4-danezis.pdf,2012.

[4] J. Hightower, S. Consolvo, A. LaMarca, I. Smith, and J. Hughes. Learning and recognizing the places we go. In *Proc. of the 7th international conference on Ubiquitous Computing*, UbiComp'05, 2005.

[5] E. De Lara, A. Lamarca, and M. Satyanarayanan. *Location Systems*. Morgan & Claypool Publishers, 2008.

[6] H. Nissembaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life* Stanford University Press, 2009.

[7] W3C. Geolocation API Specification. http://dev.w3.org/geo/-api/spec-source.html, 2012.

[8] Article 29 Data Protection Working Party WP185. Opinion 13/2011 on geolocation services on smart mobile devices. http://ec.europa.eu/justice/policies/privacy-/docs/wpdocs/2011/wp185_en.pdf.