Towards a Privacy Risk Assessment Methodology for Location-Based Systems

Jesús Friginal, Jérémie Guiochet, Marc-Olivier Killijian

Towards a Privacy Risk Assessment Methodology for Location-Based Systems

Jesús Friginal, Jérémie Guiochet and Marc-Olivier Killijian

LAAS-CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse Cedex, France
{jesus.friginal,jeremie.guiochet,marco.killijian}@laas.fr

Mobiquitous systems are gaining more and more weight in our daily lives. They are becoming a reality from our home and work to our leisure. The use of Location-Based Services (LBS) in these systems is increasingly demanded by users. Yet, while on one hand they enable people to be more "connected", on the other hand, they may expose people to serious privacy issues. The design and deployment of Privacy-Enhancing Technologies (PETs) for LBS has been widely addressed in the last years. However, strikingly, there is still a lack of methodologies to assess the risk that using LBS may have on users' privacy (even when PETs are considered). This paper presents the first steps towards a privacy risk assessment methodology to (i) identify (ii) analyse, and (iii) evaluate the potential privacy issues affecting mobiquitous systems.

Key words: privacy, risk assessment, location-based systems

1 Introduction

The vast deployment of myriads of sensors and the rapid growth in the number of mobile devices per person is providing enormous opportunities to create a new generation of innovative Location-Based Services (LBS) addressed to improve the welfare of our society. LBS are used in a variety of contexts, such as health, entertainment or work, like discovering the nearest cash machine, parking parcel, or getting personalised weather services.

Parallel to this revolution, numerous studies reveal potential privacy breaches in the use of these services given the sensitivity of the collected information, and how it is stored and exchanged [4]. From a privacy viewpoint, the main characteristic of LBS systems is that, apart from personal identity, users' location becomes a new essential asset to protect. A recent study from MIT [1] showed that 4 spatio-temporal points (approximate places and times), were enough to unequivocally identify 95% people in a mobility database of 1.5M users. The study shows that these constraints hold even when the resolution of the dataset is low. Therefore, even coarse or blurred information provides little anonymity. However, very few users are aware of the implications that a misuse of their location information may have on their privacy, and the potential consequences on their security and safety [2]. Tackling this issue becomes critical given the increasingly variety of attacks that may impact the privacy of users [4].

By the time being, there is a range from simplistic on/off switches to sophisticated Privacy-Enhancing Technologies (PETs) using anonymisation techniques

[8]. Today, few LBS offer such PETs, e.g., Google Latitude offers an on/off switch that allows to stick one's position to a freely definable location. Another set of techniques include location obfuscation, which slightly alter the location of the users in order to hide their real location while still being able to represent their position and receive services from their LBS provider. However, such efforts remain questionable in practice while suitable techniques to guarantee acceptable levels of risk remain unavailable. Consequently, the confident use of LBS requires not only the development of PETs, but also the definition of methodologies and techniques to assess and treat the privacy risk associated to LBS solutions. There exist many and various challenges in the deployment of LBS, but the need for identifying the risk related to the processing of personal data before determining the appropriate means to reduce them, is without doubt, one of the most important in the domain of LBS.

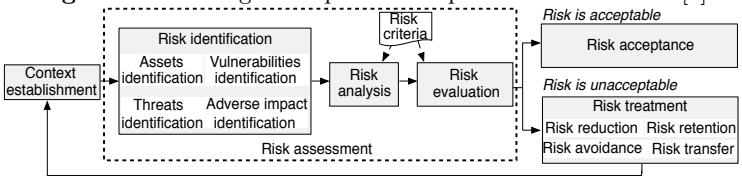
This short paper makes a step forward to explore the challenges that hinder the development of methodologies to assess the risk related to the lack of privacy of LBS. The rest of this paper is structured as follows. Section 2 shows the lack of techniques and guidelines to assess the privacy risk on LBS. Section 3 presents a privacy risk conceptual framework. Section 4 introduces our first approach towards privacy risk assess methodology for LBS. Finally, Section 5 closes the paper.

2 Privacy in risk standards

The concept of risk was first introduced in safety critical systems, but is now widely used in many domains, including information technology. Indeed, users, environment and organizations could be faced to the harm induced by the use of a new technology. The generic standard ISO/IEC-Guide73 defines the risk as the combination of the probability of an event and its consequence. This definition had to be adapted in the domain of security, where risk is defined as the "potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization" [6]. In this definition, the classic notion of *probability* of an event has been replaced by "potentiality" given the difficulty to estimate such a probability. The concept of *consequence* was also refined into "harm to the organization". The identification, analysis and evaluation of the risk, is defined in many standards and international directives as *risk assessment* within the risk management process, as Figure 1 shows.

ISO standards such as [6] or regulatory documents produced by NIST [9], deal with security risk assessment and treatment. Unfortunately, there is no privacy

Fig. 1. Risk management process adapted from ISO 27005 [6].



ISO/IEC standard dedicated to privacy risk assessment. In order to protect user data, the European Commission unveiled in 2012 a draft [3] that will supersede the Data Protection Directive 95/46/EC. It is mainly about openness and obligations of organizations managing users personal data, and it does not include methods and techniques to analyse and assess the risks. A similar approach is presented in the USA Location Privacy Protection Act of 2012 (S.1233), in order to regulate the transmission and sharing of user location data. As in the European Directives, this bill specifies entities, data and its usage, but no analysis techniques are proposed, and much less in the domain of LBS.

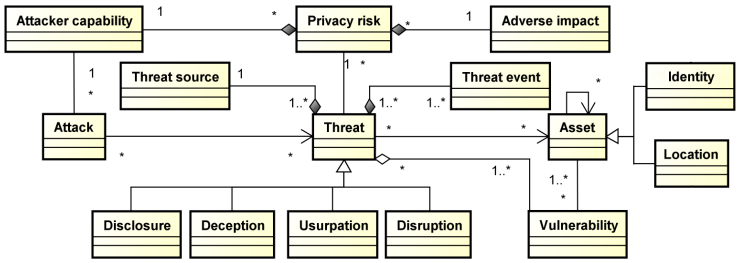
3 Privacy risk conceptual framework

This Section introduces the conceptual framework gathering the main notions relating to privacy risk. To illustrate these concepts, Figure 2 shows the general meta-model concerning the entities and relations involved in the process. Our meta-model should not be seen as a closed proposal. Instead, it is a flexible abstraction interpreting privacy risk in a generic way for LBS.

As seen in Section 2, risk is generally defined as the combination of two factors: the potential occurrence of threats, and the consequences that such threats may cause on victims, also referred to as *adverse impact*. Analogously, the definition of risk could be easily adapted to the privacy domain by refining the concept of potential occurrence as *attacker capability*, which refers to the efforts and resources employed by the *threat source* to deploy *threat events* affecting a *privacy asset*. In the case of LBS, assets are mainly intangible. Among them, *identity* and *location* are the most important. For instance, the spatio-temporal data of an individual can be used to infer his movements and habits, to learn information about his centre of interests or even to detect a change from his usual behaviour. From these basic assets it is possible to obtain derived ones. It is to note that the asset value strongly depends on its precision. Fine-grained information concerning the who-where-when tuple is much more valuable than a coarse-grained one. For example, the value of knowing that a person, identified by the full name, will be at a place with a geo-temporal error of meters and seconds is higher than knowing that a person, only identified by the surname, will be potentially at a given city within an error window of two weeks.

In the domain of privacy, the notion of *threat* particularly refers to the *disclosure* of assets. In most of the cases, conversely to traditional security, threats are

Fig. 2. General meta-model for privacy risk assessment.



based on the potential correlation of partial informations (premises), assumed to be true, to derive more valuable information. To obtain such premises, attackers may appeal to the *usurpation* of identity, the *deception* of information or the *disruption* of services. A threat presents certain precision requirements. If such requirements are compatible with the precision provided by the asset, then the threat will be potentially feasible. For example, if a threat source requires knowing the exact current position of a user, and the only information available reveals only that the user is not at home, the threat will be hardly realisable. A threat, to be realisable, needs to exploit a *vulnerability* associated to the asset. Vulnerabilities refer to inherent weaknesses in the security procedures or internal controls concerning the asset. If there is a vulnerability without a corresponding threat, or a threat without a corresponding vulnerability, there is no risk. Next section exploits this meta-model for the privacy risk assessment.

4 Towards privacy risk assessment for LBS

Thus Section presents a privacy risk assessment methodology following the framework of Figure 1. The first step, risk identification involves the identification of risk sources, events, their causes and their potential consequences. The level of risk is then estimated in the risk analysis step. For this step, we will introduce metrics as risk criteria for the privacy risk analysis.

The goal of risk identification is to collect the information concerning the meta-model in Figure 2 in a simple but structured way. First, it is necessary to identify the privacy assets. Among all the *Personal Identifiable Information*, learning the location of an individual is one of the greatest threat against his privacy. Essentially because starting with the location, many other private information can be derived [4]. Thus, once assets fixed, reasoning about the pertinent threats related to a particular asset is easier, for example inference attacks to disclose the POIs of an individual. Then, vulnerabilities will be defined as the lack of control mechanisms enabling the potential realisation of such threats on identified assets. Finally, the users of the methodology should qualitatively determine the adverse impact on the privacy of the asset. Such an intuitive technique will guide a more systematic identification of risks on the system.

The risk analysis stage aims at estimating the privacy risk. The first step towards this goal involves proposing mechanisms to quantitatively estimate the adverse impact and the attacker capability. Conversely to security, privacy presents a subjective component. Indeed, the adverse impact (AI) is a metric that depends on how a user perceives and interprets the importance of an asset [5]. The use of questionnaires can be very useful to estimate the adverse impact of a threatened asset. The estimation of the attacker capability (AC) is approached by most risk assessment standards and guidelines [6, 9] through the assignation of a numeric value. While this approach is valid to provide a coarse-grained viewpoint of the threat, no detail about how threats are exploited is provided. To cover this lack, our methodology uses the concept of *attack tree*.

The risk evaluation stage is in charge of ranking privacy risks regarding specific criteria. Applying different criteria, such as the importance of the asset for the business, the reputation of users or the regulation fulfilment, may

lead to different rankings. According to the selected criterion, risks, regardless if they were high, medium or low, will be characterised as acceptable, tolerable and non-acceptable following an As-Low-As-Reasonably-Practicable (ALARP) [7] strategy. The risks characterised as non-acceptable will be prioritised for their treatment. However, this stage is out the scope of this paper.

5 Conclusions

Privacy may be the greatest barrier to the long-term success of mobiquitous systems. However, despite many standards and approaches have been proposed to handle the problem of risk assessment, none, to the best of our knowledge has addressed the problem of managing the privacy risk for LBS. One of the major problems found in this paper concerns the identification of adequate information to carry out the risk assessment, as well as the way to process it.

Beyond this work, we are interested in studying the usefulness of our methodology to (i) guide the design PETs following a privacy-by-design approach, and (ii) compare and select (benchmark) the PETs that address the best the privacy requirements of mobiquitous systems.

Acknowledgements

This work is partially supported by the ANR French project AMORES (ANR-11-INSE-010) and the Intel Doctoral Student Honour Programme 2012.

References

[1] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3:–, 2013.

[2] J. Dobson and P. Fisher. Geoslavery. *Technology and Society Magazine, IEEE*, 22(1):47–52, 2003.

[3] European Commission. Proposal for a regulation of the european parliament and of the council on the protection of individuals., 2012.

[4] S. Gambs, M.-O. Killijian, and M. Núñez del Prado Cortez. Show me how you move and I will tell you who you are. *Trans. on Data Privacy*, 4(2):103–126, 2011.

[5] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 91–100. ACM, 2004.

[6] ISO27005. Information technology - security techniques - information security risk management. International Standard Organisation, 2008.

[7] R. E. Melchers. On the ALARP approach to risk management. *Reliability Engineering & System Safety*, 71(2):201–208, 2001.

[8] M. F. Mokbel, C.-Y. Chow, and W. G. Aref. The new casper: query processing for location services without compromising privacy. In *Proceedings of the 32nd international conference on Very large data bases*, VLDB '06, pages 763–774. VLDB Endowment, 2006.

[9] NIST800-30. Information security, guide for conducting risk assessments. U.S. Department of Commerce, NIST), 2011.