Pervasive and Mobile Computing 17 (2015) 159–174

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Privacy protection in pervasive systems: State of the art and technical challenges

Claudio Bettini, Daniele Riboni *

Università degli Studi di Milano, D.I., via Comelico 39, I-20135 Milano, Italy

ARTICLE INFO

Article history:
Available online 8 October 2014

Keywords:
Data privacy
Anonymity
Obfuscation
Pervasive applications

ABSTRACT

Pervasive and mobile computing applications are dramatically increasing the amount of personal data released to service providers as well as to third parties. Data includes geographical and indoor positions of individuals, their movement patterns as well as sensor-acquired data that may reveal individuals' physical conditions, habits, and, in general, information that may lead to undesired consequences like unsolicited advertisement or more serious ones like discrimination and stalking.

In this survey paper, at first we consider representative classes of pervasive applications, and identify the requirements they impose in terms of privacy and trade-off with service quality. Then, we review the most prominent privacy preservation approaches, we discuss and summarize them in terms of the requirements.

Finally, we take a more holistic view of the privacy problem by discussing other aspects that turn out to be crucial for the widespread adoption of privacy enhancing technologies. We discuss technical challenges like the need for tools augmenting the awareness of individuals and to capture their privacy preferences, as well as legal and economic challenges. Indeed, on one side privacy solutions must comply to ethical and legal requirements, and not prevent profitable business models, while on the other side it is unlikely that privacy preserving solutions will become practical and effective without new regulations.

1. Introduction

In our everyday life, independently from our attitude towards technology, we inevitably release personal information either by filling in a job application, by obtaining a shop fidelity card, by passing through a toll-road, by using a credit card or just by walking in shopping malls and public streets where tens of cameras are deployed. In principle, this information may be improperly used in a number of ways including unsolicited advertisement, discrimination, identity theft, or even stalking. This is not a new situation, but the advance in technology has dramatically increased not only the amount of personal information that is acquired, but in particular the ability to store, process, and share this information. Cloud computing and social networks in particular have given a boost in the collection, sharing and processing of personal information that inevitably lead to several privacy breach accidents. For this reason, privacy, generally defined as *the right to be let alone*, is becoming more and more a shared concern. Nonetheless, a recent review on mobile apps for healthcare and well-being has shown that a relevant part of both free and paid apps does not expose any privacy policy; many of them send data to undisclosed third parties; and the vast majority of them transmits potentially sensitive data in plain text [1].

A natural question for our readers is what is the role of mobile and pervasive computing in this situation? Indeed, independently from the success of cloud computing, social networking, and the big data analysis that they enable, we are

* Corresponding author. Tel.: +39 0250316350.
E-mail addresses: claudio.bettini@unimi.it (C. Bettini), daniele.riboni@unimi.it, riboni.daniele@gmail.com (D. Riboni).

http://dx.doi.org/10.1016/j.pmcj.2014.09.010
1574-1192/

witnessing a shift of paradigm in computer science. In 2011, smartphones and tablets outsold workstations and portable computers by 1.5×. In 2013, this ratio reached 4×. A major share of the smartphone apps has rich functionalities exploiting location, time, and other context parameters. An increasing number of objects, from consumer electronics to home appliances, from clothes to accessories are acquiring sensing, computing and communication capabilities; Wearable sensors monitoring fitness activities, sleep disorders, motion patterns as well as user's physical parameters are becoming trendy. We start seeing our home and office infrastructure sensing and communicating energy consumption patterns, presence patterns and in general information about what its inhabitants are doing. Analogously, the city we live in is using more and more sensing technology to become aware of what its inhabitants are doing with the noble goal of optimizing its services and improve safety.

It is quite intuitive that this shift of paradigm is indeed having a deep impact on how we share personal data and on how we will be sharing it in a few years. This impact has been perceived not only by technology experts, but by sociologists, economists and last but not least by law regulators. A growing number of people are concerned about the negative consequences that may arise from the large-scale monitoring of individuals' life in terms of human rights and societal values [2]. The 2014 White House report on "Big Data and Privacy"[1] also highlights the challenges for data protection that the collection of personal data through pervasive technologies implies. A major data protection reform has been recently proposed for adoption in the EU,[2] and analogous initiatives are being discussed all over the world. It is indeed a major question if technological solutions alone can address the privacy problem; it is our responsibility as researchers in this area to identify and explain the possible privacy threats that may be hidden in pervasive applications. Not only this could help in designing a modern regulation system that protects users without preventing new business opportunities, but it would also help software developers to design applications that better inform the user about possibly hidden consequences in terms of personal data release, and possibly mitigate the risk of privacy violation. This paper is intended to give a contribution along these lines.

The rest of the paper is structured as follows. In Section 2 we identify different categories of pervasive and mobile applications, the privacy threats that they may pose, and the requisites that should be considered when designing a privacy protection strategy. Section 3 critically analyses the state of the art for the privacy enhancing technologies applicable to pervasive applications with a specific reference to the identified requirements. The gap between the state of the art and the actual requirements is discussed in Section 4 describing open issues and research challenges, including economic, legal, and usability aspects. Section 5 concludes the paper.

2. Applications, privacy threats, and requirements

2.1. Identifying privacy threats

In this survey we consider a privacy violation to occur when the association between an individual's identity and some personal information is acquired, retained and/or processed by a third party without the consent by the individual. We refer to the third party as *the adversary*.

Despite new pervasive and mobile applications being proposed everyday as the result of the rapid evolution of sensing and mobile technologies, in this section we will make an effort to identify what we consider main application categories, and then analyze them in terms of the risks for privacy violation that they may involve. Understanding the possible presence of privacy threats includes understanding which parts of the information being released are considered sensitive to the user and which parts may be used to identify or re-identify (when joined with other information) the user. Since the actual presence of a privacy threat depends also on the entities that gain access to personal information, we will also mention the *possible adversaries* to be considered for each category. Indeed, no privacy protection technique can be properly validated if a clear adversary model is not provided. The model must specify at least: (a) which part of the personal information being transferred and/or processed the adversary has access to (e.g., complete/partial, occasional/historical, etc.), (b) which external or background knowledge the adversary has access to, (c) if different adversaries can collude.

Table 1recaps the main application categories, sensitive data, and adversaries, which are discussed below. Intruders, as well as third parties that may legally access the data, are considered adversaries for all categories of applications; hence, they are omitted from the table. Note that what was listed as sensitive data, may not be sensitive by itself, but may lead to a privacy violation when joined with external information. Moreover, sensitive data can lead to a privacy violation only when the adversary can link it to an identity. The table does not specify another kind of data released by these apps and that is referred in the literature as *quasi-identifier*. This is data that may be used by an adversary to re-identify the user (e.g., location itself may be used to restrict the candidate individuals being at that location in a given time instant). These aspects will be discussed in Section 3.

2.2. Location based services (LBS), mobile advertisement (MA), and Geo-social network applications (GeoSNs)

Characteristics: This is currently the category that includes most of mobile (and pervasive) applications since users and businesses have greatly appreciated the personalization of services based on user location. Examples of these apps are

[1] http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.

[2] http://ec.europa.eu/justice/data-protection/.