Chapter 1

An Introduction to Privacy Aspects of Information and Communication Technologies

A. Martínez-Ballesté and Agusti Solanas

Rovira i Virgili University of Tarragona, UNESCO Chair in Data Privacy
Av. Països Catalans 26, 43007 Tarragona, Catalonia, Spain
{antoni.martinez,agusti.solanas}@urv.cat

Contents

1.1. Introduction

Information and communication technologies (ICT) are the fundamental tool for the proper development of the Information Society. These technologies allow the automatic collection, storage and processing of large amounts of data. Most of these data are related to human activity: e.g. processing census data, electricity bills, payslips, medical histories, etc. The processing of these data by governments or companies has always entailed discussions about privacy. Even the Universal Declaration of Human Rights[1] states that *"no one shall be subjected to arbitrary interference with his privacy"* and that *"everyone has the right to the protection of the law against such interference or attacks"*. The privacy threat entailed to computer systems was already pointed out by Hoffman in 1969.[2]

Internet is the world-wide platform for communication, information search and other activities, namely shopping, expressing opinions, sharing pictures, etc. Since the mid 1990s, millions of people from different cultures and education levels interact with a plethora of interconnected information systems. Internet and ubiquitous computing involve an easier interaction between information systems and their users. Hence, data collection by service providers has become straightforward and, consequently, the amount of collected data has dramatically increased. This allows, for instance, the rapid creation of profiles for customer relationship management. However, data collection and analysis comes into conflict with users privacy, since their habits can be profiled by third parties and their actions tracked as well. Here follow some examples:

- *Digital Rights Management systems* require the users to identify themselves when subscribing or paying for some content. Hence, their tastes and preferences are being exposed.[3]
- *Authentication and identification of users* in a variety of services (e.g. transportation, e-mail services, online banking, etc.) allow the tracking of their activities.
- *Computerised video surveillance* entails several privacy concerns. Its wide deployment clashes with a diversity of data protection laws. Hence, several techniques must be applied to guarantee the privacy of individuals.[4]

The access to these data must be limited by governmental regulations to ensure confidentiality and guarantee the privacy of the users.[5]

In that sense, the Organisation for Economic Co-operation and Development published in 1980 a set of guidelines on the protection of personal data.[6] They mention that information should only be collected with the *consent of the individuals* and *should not be disclosed to third parties.* These guidelines have been adopted as legal initiatives in several countries.

In Europe, the Directive 95/46/EC on the protection of personal data[7] develops the aforementioned guidelines. In addition, the Directive 2002/58/EC on privacy and electronic communications[8] elaborates on data collection and processing by means of ICT. According to the directive, the public and massive access to ICT gives rise to requirements concerning the protection of the privacy of their users. It states that *"the successful cross-border development of these services is partly dependent on the confidence of users that their privacy will not be at risk"*. It is observed in this directive that publicly available electronic communication services over the Internet open new possibilities for users but also new risks for their personal data and privacy. As a result, service providers should take appropriate measures to safeguard the security of their services and take care of the privacy of their users. Since Directive 95/46/EC also remarks that invading privacy may be necessary to safeguard national security, the Directive 2006/24/EC on data retention[9] provides a framework for legislation on the collection and storage of data with the purpose of detecting and prosecuting serious crime. Japan, Russia, the United States and Canada[a] currently have no comprehensive privacy laws.[10]

However, owing to the increasing availability and ubiquity of ICT, these legal initiatives will only result effective if users are concerned with ICT privacy threats and privacy enhancing technologies are widely deployed.

In order to clarify what providers can do with users' information, *privacy policies* were introduced. They are published by service providers and users make a decision on asking for the service based on these policies. Hence, if service providers do not properly follow their privacy policies, users have the right to take legal action against providers. These policy-based schemes are widely used on the Internet by e.g. e-commerce sites that define their privacy policies in e.g. P3P (Platform for Privacy Preferences).[11]

In this chapter we present a summary of the privacy concerns related to ICT. We also point out how ICT can contribute to enhance the privacy of their users. The next topics are addressed:

- Privacy of Internet users.
- Privacy in databases.
- Privacy in ubiquitous computing.

1.2. Privacy and the Internet

Internet is used by millions of people 24 hours a day to access all kinds of information. In addition, Internet technologies allow people to create and share a variety of content. Since the spread of broadband network access technologies, most Internet users are almost permanently connected to the network. This makes them more vulnerable to several attacks and perils, namely viruses and system hacking. Current operating systems include a firewall and antivirus or, at least, give the users advice on installing security software. Fortunately, computer users are aware that only with these tools installed and updated their systems may remain safe. However, *users are not concerned about privacy risks*.

In this section we elaborate on some aspects of the privacy of Internet users.

[a]The non-European countries in the *Group of Eight*.

1.2.1. *Monitoring User Activity*

As mentioned before ICT can be used to create personal profiles. The collection of traces from Internet services (e.g. visited websites, frequency of browsing, queries at search engines, etc.) allows the monitoring of user activity. Clearly, using Internet services entails the logging of some data, usually the Internet address of the user and a timestamp. Moreover, some Internet services explicitly ask for personal data in order to proceed with the service. Users are informed about security aspects and information sharing policies of the service. Hence users are conscious of the secure management of their personal data.

Surfing the Internet gives users the impression of a *high degree of privacy*.[5] As a result, most of them do not take into account that their activity can be *monitored*.[12] Next, we describe several techniques used to monitor the activity of Internet users.

1.2.1.1. *Tracking Cookies*

Cookies are small pieces of information that web servers are allowed to put into user computers. They are retrieved from user computers to achieve several purposes, e.g. keeping user preferences, remembering the identity of the user and making web sessions feasible. Moreover, a web server is able to collect some information about user habits, since cookies allow the storage of the necessary information.

Tracking cookies[b] are not handled by the websites being visited, but by a central third party (usually an Internet marketing company). Hence, this third party will be able to track users browsing across several websites and, as a result, to build accurate profiles about the interests of these users. This information will be mainly used to deliver advertisements targeted to the interests of the user. Since the activity of users can be monitored, cookies are considered a potential privacy threat.[13,14]

With regard to this, the United States *Online Profiling Recommendations*[15] state that Internet users may be totally unaware that their activities are being monitored. Fortunately, almost all the Internet marketing companies offer the so-called *optout* options to prevent their cookies to track the Internet user.[16] Furthermore, some security software also remove tracking cookies from users systems.

[b]Or third party cookies.

1.2.1.2. *Spyware*

Spyware are small pieces of software installed in a computer without the consent of the user. Their main task is to send information to remote databases so that it can be used to determine user habits. In addition, the worst side of spyware is that they can be programmed to send chat sessions or even passwords to the remote database, undermining confidence in online activities and the effectiveness of transport layer security.[17]

They do not spread like computer viruses do: it can be downloaded and installed by users themselves, because, for instance, they have clicked on a link that promises faster Internet browsing or even protection against spyware. Spyware can come bundled with free software and, potentially, any software can include spyware. It may also be installed as *browser helpers* or toolbars. Besides tracking browsing habits, spyware uses to slow down the system and pops-up advertisements randomly. Hence, spyware is annoying and negatively impacts the use of Internet.[18]

Although a plenty of security tools are intended to clean computer systems, spyware removal is not straightforward.

1.2.1.3. *Social Engineering Attacks*

The aim of social engineering attacks is to manipulate people into disclosing private information. *Phishing* is a social engineering attack that makes use of massive sending of fraudulent e-mails on behalf of trusted entities (e.g. online banking services or Internet service providers). In these e-mails, users are requested to visit a link that points to a fake website. These websites contain a form asking for several information (including private information such as passwords) that must be filled in to avoid "*serious security problems*".

The prevention of phishing relays on the successful detection of fraudulent e-mails and fake websites. Current versions of web browsers or e-mail managers are capable of warning the user in case of detecting a site or e-mail suspicious of phishing. Besides, fake websites are not authenticated by trustworthy certificate authorities and, in addition, they often present misspellings and grammar errors. In that sense, only the *shrewdness* and *training* of users may prevent them from becoming victims.[19]

1.2.2. *Privacy in modern web services*

During several years, the web has been merely a large public-access information database allowing easy information search and retrieval. In contrast, only skilled users were able create websites. Current Web 2.0[20] services allow the creation and sharing of content in a straightforward manner. In addition, these services provide users with tools to *collaborate* and express opinions. As a result, there is a huge amount of information from individuals in the Internet: comments on a blog post, opinions about a product, pictures in a sharing service, etc. This information is published in social network services, e-commerce sites, etc. Some of these services act as *mashup services*[21] which gather and analyse the information from several other services. The popular use of these technologies raises new questions on privacy, specially in social networks and e-commerce services.

1.2.2.1. *Privacy in social networks*

Social network services allow people to be connected to an online community. Widely used social networks gather millions of users whose goal is to build large groups of *friends*. The use of these large communities has dramatically increased in the last years[c]. In these online services, users publish a profile that includes personal facts, i.e. their birth date, religious and sexual orientation, etc.[d], and can also share pictures and inform of their daily activities. This information can be accessed by virtually every member of the social network.

Social networks have developed privacy policies in which users are responsible for controlling the information they post and deciding who can access it.[23] Hence, most users are aware that they must take control over their information. However, there are some situations in which the privacy of non-members is exposed. As an example, if a user posts a picture of her with other people which are not members of the social network, the latter cannot control who views this picture. Moreover, there are some critical situations which are difficult to overcome: e.g. dishonest users creating fake profiles with harmful intentions such as monitoring their friends.[24] As a conclusion, only the consciousness of the users and informing them of a

[c]For example, in January 2008, the most popular social networks had more than 100 million users.

[d]In a recent study[22] on how a large community of students were using a social network, authors claim that a 28% of profiles posted contained a cellphone number and the majority of users disclosed their dating preferences, current relationship status or political views.

code of ethics may mitigate the privacy threats in social networks.

1.2.2.2. *Privacy in e-commerce*

In e-commerce services the retailer knows what products have been purchased by their customers. This is feasible in real-world shopping as well. Moreover, online vendors store a large amount of data about their customers and their preferences, opinions and purchases. In addition, service providers ask customers for e-mail address and other personal data. These personal data is transferred to third parties to perform a variety of functions on behalf of the service provider, e.g. delivering packages, sending advertisement e-mails, data analysis for marketing strategies, etc. Online shopping services specify the data collection and management by means of privacy policies that *users implicitly accept when they use the service*.

In online shopping, service providers make use of customer's data to inform about the popularity of their products or to recommend additional purchases to the customers. Moreover, *collaborative filtering* techniques are used to provide online shopping services with complex *recommendation systems* based on the preferences of their users. If service providers do not handle customer's data properly, users may give biased opinions to the recommendation system or even refrain from using it.[25] In that sense, several privacy-preserving collaborative filtering techniques are proposed: e.g. data perturbation,[26] distributed architectures[27] and data obfuscation.[28]

Modern web services also provide Internet users with social networks specialised in shopping, namely *social commerce* and *social shopping* networks. These networks gather thousands of users with the aim of knowing from each others shopping habits. Hence, users comment on real-world or Internet shopping and rate products, stores and even restaurants. In these cases, privacy threats are similar to those in social networks. Moreover, social commerce networks involve emerging collaborative commerce services, e.g. *group buying* (a model in which several buyers cooperate and buy at a discount price). These new services must rely on privacy enhancing technologies, so as to guarantee their proper development and rapid growing.

1.2.3. ***The Big Brothers of Internet***

We have mentioned that data is collected with several purposes: from *cybercrime* acts made by individuals to *marketing strategies* to improve customer service. In this case several marketing companies reach an agreement with some Internet services so they can gather the data from the latter. By means of data mining, marketing and Internet advertising companies are aware of different customer needs.

However, these companies may have only the picture of a specific kind of Internet users. In contrast, the few companies that produce and manage the most popular software and Internet services are able to track the activity of the majority of ICT users. These companies offer several services which can collect information from *virtually every activity done by means of ICT*. In addition to operating systems and productivity software suites, these companies provide users with efficient Internet search engines and offer access to widely used e-mail and instant messaging services. They also offer location-based applications, multimedia sharing and searching engines, calendar managers, etc. Due to the large amount of information collected by these companies, some criticisms about privacy concerns have arisen,[29,30] basically on these companies becoming *the Big Brothers of Internet*.

Regarding search engines, the queries of the users are stored and analysed to improve search results and to provide *personalised advertisements*. In addition to that, the queries that users send to a search engine can disclose their interests and even their political views or sexual dispositions.[5]

With respect to Internet communications, e-mail and instant messaging services manage the information flow between millions of users. Although most of these communications are private, they can be *scanned* to track offenders and spammers. In order to comply with data retention policies, *instant communications can be stored and e-mail messages retained* even users delete them from the servers. The information in blog services and opinion forums may also be retained and scanned with similar purposes.

Naturally, users are warned by the service providers that all the data collected and stored in their large information services will be managed according the legislation.

1.3. Privacy in Databases

Most of information systems and information technologies rely on databases. They store a variety of data and provide users with several tools for their management. In this section we describe the privacy aspects of databases and how this privacy can be preserved.

The meaning of privacy in databases is an ambiguous concept, whose meaning is context-dependent.[31] For instance, in statistical databases containing information from individuals, it refers to *respondent privacy*. In querying search engines, the concern is about preventing service providers

Table 1.1. Example data set, where *Name* is an identifier to be suppressed before publishing the data set

Name	ZIP	Weight	Height	Salary	Cancer
John Doe	43152	85	182	26000	Yes
Maggie Simpson	43152	100	165	17000	Yes
Robert Smith	43132	85	183	32000	No
Mary Martin	43152	56	165	22000	No
Brian Johnson	43152	84	205	17400	Yes
Luisa Lopez	43121	65	198	23000	No
James Dawson	43122	66	160	18500	Yes
Peter Petrelli	43144	102	195	43500	Yes
Bridget Jones	43152	108	180	21600	Yes

from profiling users (hence, it refers to *user privacy*). In this section, we elaborate on both aspects.

1.3.1. *Privacy of the Respondents*

The main goal of respondent privacy in statistical and administrative databases is to prevent the re-identification of the respondents to which the records of a database correspond. If an statistical agency releases or publishes the data as is, there could be a serious privacy loss. It could be thought that deleting the direct identifiers (*e.g.* the name of the respondents) from the data before release could be enough to guarantee respondent privacy. Unfortunately, protecting the privacy of the respondent is not so straightforward.

Data attributes can be classified as direct identifiers, quasi-identifiers, sensitive information and non-sensitive information. Table 1.1 shows an example data set where *Weight* and *Height* are quasi-identifiers and *Salary* and *Cancer* are sensitive attributes. We can build a dataset without sensitive information and another one with no direct identifiers, hence the latter can be public. However, if intruders having the dataset without sensitive information, make queries to the public data set, they will be able to disclose sensitive information (e.g., Maggie Simpson, with *ZIP*=43152, *Weight*=100 and *Height*=165, earns a salary of 17,000 and has cancer[e]). Moreover, intruders can disclose the salary of the "tallest man in a town" simply by querying the public data set.

To protect the privacy of the respondents and, in addition to suppressing the identifiers, *statistical disclosure control* (SDC) techniques must be

[e]If the intruder belongs to a health insurance company, the latter may not be interested in offering their products to Maggie Simpson.

Table 1.2. The example data set with identifiers suppressed and attributes *Weight* and *Height* micro-aggregated with $k = 3$

ZIP	Weight	Height	Salary	Cancer
43152	84.6	181.6	26000	Yes
43152	*103.3*	*163.3*	*17000*	Yes
43132	84.6	181.6	32000	No
43152	62.3	163.3	22000	No
43152	84.6	199.3	17400	Yes
43121	62.3	199.3	23000	No
43122	62.3	163.3	18500	Yes
43144	103.3	199.3	43500	Yes
43152	103.3	181.6	21600	Yes

applied to databases before releasing them. There are a variety of SDC techniques to protect databases. For instance, if quasi-identifiers achieve the k-anonymity property[32] (i.e., if a quasi-identifier is indistinguishable from other $k - 1$ identifiers), it will become harder to link sensitive attributes with identifiers. k-Anonymity can be achieved by means of several methods, e.g. micro-aggregation. When micro-aggregation is applied, the set of original records is partitioned into several groups in such a way that records in the same group are *similar* to each other and the number of records in each group is at least k. An aggregation operator (for example, the mean for numerical data) is used to compute a centroid for each group. Then, each record in a group is replaced by the group centroid. Table 1.2 shows the data set without identifiers once quasi-identifiers have been masked with $k = 3$. Unfortunately, k-anonimity may not be enough to protect the privacy of the respondents, e.g. if records sharing a combination of quasi-identifiers have the same value for the sensitive attribute. Hence, several properties complement k-anonimity to improve respondent privacy.[33]

Naturally, applying *masking methods* such as micro-aggregation results in some *information loss* of the data. Thus, SDC techniques look for the best trade off between information loss and *disclosure risk*.[34] That is, both the risk of disclosing private confidential information and the loss of data utility should be kept below reasonable thresholds.

1.3.2. *User Privacy*

User privacy is about guaranteeing the privacy of queries to dynamic databases in order to prevent user profiling and re-identification, for instance:

- Preventing the users of Internet search engines or location-based services from being profiled. As mentioned in Section 1.2.3, Internet search engines claim this profiling is to improve search results.
- Preventing the administrators of the database from getting important information from the queries sent by users. For example, if someone having a great idea queries a database to know if it has already been patented, the administrator, who is aware the query, is able to patent the idea first.

A naive approach for achieving user privacy is the downloading of the entire database, so the user could query the database locally. Naturally, this is not applicable in large databases such as Internet search engines. *Private Information Retrieval* (PIR) techniques allow retrieving information from a database while maintaining the query private from the database administrators.

The first PIR proposal[35] uses at least two copies of the same database which do not communicate with each other. A second proposal allows single-database PIR by means of public-key encryption.[36] In these proposals, the database is modeled as a vector from which the user wants to retrieve the i-th component while keeping the index i hidden from the database administrator. Hence it is assumed the user knows the physical address of the item. On the other hand, these models assume that the server cooperates in the PIR protocol. However, the users are those really interested in their own privacy, whereas the motivation for servers is dubious since PIR limits their profiling ability. As a result, these proposals may be unpractical in many real-world scenarios and, hence, new feasible protocols are being proposed.

In *collaborative PIR* schemes, users collaborate with each other in order to be cloaked in a user community[37] or seek assistance by a peer-to-peer community who submit queries on their behalf. Another strategy, targeted to Internet search engines, consists of a client software to *add fake keywords* to the user's query.[38]

1.4. Privacy in Ubiquitous Computing

Ubiquitous computing allows the integration of information systems to a variety of environments. In addition, it provides the Information Society with a wide range of new applications. In this section we discuss on privacy aspects of two significant and emerging technologies related to ubiquitous computing: *location-based services* and *radio-frequency identification*.

1.4.1. *Location-Based Services*

Location-Based Services (LBS) are gaining importance due to the advances in mobile networks and positioning technologies. They provide users with highly personalised information accessible by means of a variety of mobile devices that are able to locate themselves, e.g. by using a GPS or a fixed network infrastructure with GSM. Mobile devices are ubiquitous and services related to the user's current location proliferate. Examples of LBS are location-based tourist information, route guidance and location-based advertising, etc.

The simplest form of communication in a location-based information transaction involves a user and a provider. The former sends a simple query that includes her location. When users send their current locations, they are implicitly trusting the provider: They assume that the provider manages their location data honestly. Hence service providers can set up their own privacy policies, such as those of Section 1.1.

However, providers may not always be trusted and, hence, new techniques to achieve the same goals without assuming mutual trust are needed. We next review the most significant proposals and we classify them regarding the use of *trusted third parties* (TTPs), i.e. entities that are completely trusted and, hence, theoretically guarantee the privacy of their users.

1.4.1.1. *TTP-based schemes*

Most of the solutions to address the privacy problems in LBS are based on TTPs. They are common because they are easy to understand/implement and because they offer a good trade-off between efficiency, accuracy and privacy. Moreover, some of the ideas used in these schemes appeared in more mature fields like e-commerce. Providers are no longer aware of the real locations of their users since trust is handed over to the TTP. Thus, these solutions move users' trust from providers to those TTPs.

Pseudonymisers receive queries from users and, prior to forwarding them to providers, they replace the real identities of the users by fake ones. Note that users must completely trust pseudonymisers, because they know all the information (i.e. identity and location) of the users. The main problem of this approach is that an attacker can infer the real identity of a user by linking a user's location with e.g. a public telephone directory.[39]

Moreover, *anonymisers* perturb real locations to achieve the k-anonymity property (already discussed in Section 1.3.1), so that it becomes harder to bind locations to users. In that context, the location of a given user is k-anonymous if it is indistinguishable from the location of $k-1$ other users. Many examples of this approach and other similar ones based on cloaking can be found in the literature.[39–41]

1.4.1.2. *TTP-free proposals*

From the users' point of view, trusting an intermediate entity may be a problem in some situations. In addition, due to the fact that intermediate entities concentrate the queries of many users, communication delays can appear and the risk of failure is increased. With the aim to overcome these limitations, TTP-free schemes have been proposed. We distinguish those based on collaboration between users/providers from those based on obfuscation.

A first method based on *collaboration between users* is to generate *cloaking areas* in which the real locations of the users remain hidden.[42] The main problem of this approach is that users have to trust other users[f]. On the contrary, k users can add Gaussian noise with null average to their locations so they can freely share their perturbed locations and compute a *global centroid*,[43] which is used to represent them all[g]. Unfortunately, if users repeat this protocol several times without changing their locations, they will be disclosed due to the noise cancellation. To solve this problem, an improvement over this method[44] suggests the use of privacy homomorphisms that guarantee the privacy of the users even when they do not change their location. This method was extended to support non-centralized communications[45] by adding a randomly generated chain of users that share location information to compute a global centroid that identifies them. Finally, users can self-organise in on-the-fly structures which are used to permute their queries in their way to the providers.[37] This results in a *private location-based information retrieval scheme* in which the provider is unable to link queries with their originating users.

Instead of using privacy protocols among users a different approach based on *user-provider collaboration* is proposed.[46] In this approach, users and providers collaborate to implement a PIR protocol that guarantees the

[f]A malicious user can easily obtain and publish the location of other users
[g]By using this procedure, the k users become k-anonymous

location privacy of the user. Although the idea of using PIR techniques is promising, requiring the provider to collaborate with users seems unrealistic and can be computationally costly.

Obfuscation is a TTP-free alternative to collaboration-based methods that can be understood as the process of degrading the accuracy of the users' location with the aim to protect their privacy. In[47] an obfuscation method based on imprecision is presented. The space is modeled as a graph where vertexes are locations and edges indicate adjacency. Hence, in order to obtain an imprecise location, users send a set of vertexes to the provider instead of a single vertex. The provider cannot distinguish which of the vertexes is the real one. Some other recently proposed obfuscation methods can be found in,[48] where the real locations of users are replaced by circular areas of variable centre and radius, and in,[49] where a method called *SpaceTwist* is proposed. This method generates an anchor (i.e. a fake point) that is used to retrieve information on the k-nearest points of interest. After successive queries, users can determine their closest interest point while the provider cannot derive their real locations.

1.4.2. *Radio Frequency Identification*

Radio-Frequency Identification (RFID) devices will become ubiquitous in the near future. The spectacular market push of RFID technology is due to the interest by large retailers, important manufacturers and governments.

Simple RFID tags can be seen as a proper substitute of bar codes since they are mainly used to identify objects. Almost every object is liable to carry a *tag*[h]. Unlike bar codes, RFID devices allow objects to be identified without visual contact and help in improving and automating many processes. In addition, complex tags, with computing hardware embedded, open the door to a variety of applications, e.g. contactless authentication.

To integrate RFID tags with information systems, RFID *readers* are needed. A tag will send some information (e.g. the identifier of the product, the result of some cryptographic function, etc.) upon a request from a reader in its range. Consequently, an eavesdropper could collect several information from people carrying tags (e.g. the brand of their clothes, the amount of banknotes they have in their pockets, the use of prosthesis, etc.) Moreover, making use of several readers strategically deployed, the location of tags can be tracked and, consequently, the motion of the scanned people. Hence, *anonymity* and *untraceability* must be tackled in RFID technology

[h]Even banknotes.

if it is to be widely deployed.[50]

In that sense, several protocols for RFID take into account the privacy concerns of this technology. The design of the tags itself can consider to implement a command (i.e. *kill* or *sleep*) to permanently or temporarily disable its ability to answer requests from readers.[51] However, these proposals may not be adequate in many scenarios.

Juels suggested the use of *a collection of pseudonyms*[52] for a given tag with the idea of answering to each query with a different identifier (i.e. pseudonym). An *authorised reader* stores this collection of identifiers and, consequently, is able to match the identifier with the tag. Thus, a *non-authorised reader* will only see different identifiers. However, if the same reader polls a tag a sufficient number of times, it will be able to collect the whole list of pseudonyms of a given tag. A possible solution to this attack consists in providing the tag with the ability of detecting the reader that is making the queries and refusing to answer when these queries are too much frequent. The use of pseudonyms is used in some other proposals.[53,54]

As mentioned before, the use of RFID technology in authentication involves the deployment of tags enabled with a variety of *cryptographic hardware*. Hash functions and simple operations are sufficient for simple applications, whereas public key cryptography allows the deployment of more complex services and protocols. These cryptographic capabilities can also be used to provide the identification and authentication applications with privacy. In this case, the challenge is to design protocols for *low-cost* tags (i.e. with minimal set of cryptographic hardware).

As an example, in the *Improved Randomised Hash-locks* proposal[55] the tag uses a hash function to hide the identifier in the identification procedure. Several protocols based on simple cryptography operations can be found in the literature[56–60] and their privacy is analysed as well.[50] Some of these protocols suffer from scalability problems, which can be solved by means of the deployment of a *distributed structure of readers*.[61]

1.5. Conclusions

In this chapter, we have reviewed the most significant privacy concerns of information and communication technologies. We have focused on privacy of Internet users, privacy related to the use and deployment of databases and privacy aspects of two ubiquitous technologies: RFID and location-based services.

Regarding privacy in Internet, we have shown that service providers are *able to profile and track their users and we have pointed out the privacy threats of social networks. In addition, we have considered spyware and social engineering attacks as powerful techniques to threaten the privacy and security of users.*

The privacy in data bases has also been discussed. We have illustrated the importance of protecting statistical databases so as to protect the privacy of respondents. Moreover, we have elaborated on privacy preserving techniques for using databases and on preventing queryable information services from profiling their users.

Finally, regarding ubiquitous computing, we have reviewed the trends in protecting the location privacy in location-based services and have pointed out current research topics in privacy in RFID.

...

Disclaimer and Acknowledgements

The author is solely responsible for the views expressed in this paper, which do not necessarily reflect the position of UNESCO nor commit that organization. This work was partly supported by the Spanish Ministry of Education through projects TSI2007-65406-C03-01"E-AEGIS" and CONSOLIDER CSD2007-00004 "ARES", and by the Government of Catalonia under grant 2005 SGR 00446.

References

1. Universal Declaration of Human Rights. `http://www.unhchr.ch/udhr/lang/eng.htm`.
2. L. J. Hoffman, Computers and privacy: A survey, *ACM Computing Surveys*. **1**(2), 85–103, (1969).
3. J. E. Cohen, DRM and privacy, *Communications of the ACM*. **46**(4), 46–49 (Apr., 2003).
4. F. Coudert and J. Dumortier. Intelligent video surveillance networks: Data protection challenges. In *The Third International Conference on Availability, Reliability and Security – ARES*, pp. 975–981. IEEE Computer Society, (2008).
5. L. Kluver et al. ICT and Privacy in Europe: Experiences from technology assessment of ICT and Privacy in seven different European countries. `http://epub.oeaw.ac.at/ita/ita-projektberichte/e2-2a44.pdf`, (2006).
6. Guidelines on the Protection of Individuals with regard to Automatic Processing of Personal Data . `http://eur-lex.europa.eu/`.
7. Directive 95/46/EC on the Protection of Individuals with regard to the

Processing of Personal Data and on the Free Movement of such Data. http://eur-lex.europa.eu/.
8. Directive 2002/58/EC on Privacy and Electronic Communications. http://europa.eu.int/eur-lex/pri/en/oj/dat/2002/201/20120020731en00370047.pdf.
9. 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. http://eur-lex.europa.eu/.
10. Privacy International. The 2007 International Privacy Ranking. http://www.privacyinternational.org.
11. W3C. Platform for privacy preferences (P3P) project. Webpage (October, 2007). URL \url{http://www.w3.org/P3P/}. http://www.w3.org/P3P/.
12. M. Petković and W. Jonker, *Security, Privacy and Trust in Modern Data Management*. (Springer, 2007), 1st edition. ISBN 978-3-540-69860-9.
13. D. Kristol, Http cookies: Standards, privacy, and politics, *ACM Transactions on Internet Technology*. **1**(2), 151–198, (2001).
14. E. Sit and K. Fu, Web cookies: Not just a privacy risk, *Communications of the ACM*. **44**(9), 120, (2001).
15. United States Federal Trade Commission. Online Profiling Recommendations. http://www.ftc.gov/os/2000/07/onlineprofiling.htm (July, 2000).
16. P. Dixon. Consumer tips: How to opt-out of cookies that track you. http://www.worldprivacyforum.org/cookieoptout.html (June, 2006).
17. Q. Hu and T. Dinev, Is spyware an internet nuisance or public menace?, *Communications of the ACM*. **48**(8), 61–66, (2005).
18. R. Thompson, Why spyware poses multiple threats to security, *Communications of the ACM*. **48**(8), 41–43, (2005).
19. P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. I. Hong, and E. Nunge. Protecting people from phishing: the design and evaluation of an embedded training email system. In eds. M. B. Rosson and D. J. Gilmore, *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 905–914. ACM, (2007).
20. T. O'Reilly. What is web 2.0: Design patterns and business models for the next generation of software. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html (September, 2005).
21. Mashup services (wikipedia). http://en.wikipedia.org/wiki/Mashup_(web_application_hybrid).
22. R. Gross, A. Acquisti, and H. J. H. III. Information revelation and privacy in online social networks. In eds. V. Atluri, S. D. C. di Vimercati, and R. Dingledine, *Workshop on Privacy in the Electronic Society*, pp. 71–80. ACM, (2005).
23. Facebook privacy policy. http://www.facebook.com/policy.php.
24. L. Schiffman. False friends: How fake facebook profiles open the door to monitoring. http://www.northbynorthwestern.com/2007/10/4687/facebook-friends/.

25. S. Berkovsky, N. Borisov, Y. Eytani, T. Kuflik, and F. Ricci. Examining users' attitude towards privacy-preserving collaborative filtering. In *Workshop on Data Mining for User Modeling, at the International Conference on User Modeling*, pp. 16–22, (2007).
26. Zhang, Ford, and Makedon. A privacy-preserving collaborative filtering scheme with two-way communication. In *CECOMM: ACM Conference on Electronic Commerce*, (2006).
27. W. Ahmad and A. A. Khokhar. An architecture for privacy preserving collaborative filtering on web portals. In *IEEE Industry Applications Society Annual Meeting*, pp. 273–278. IEEE Computer Society, (2007). URL `http://doi.ieeecomputersociety.org/10.1109/IAS.2007.15`.
28. R. Parameswaran and D. M. Blough. Privacy preserving collaborative filtering using data obfuscation. In *IEEE International Conference on Granular Computing – GrC*, pp. 380–386. IEEE, (2007). URL `http://doi.ieeecomputersociety.org/10.1109/GRC.2007.129`.
29. J. Hopkins, Once-brotherly image turns big brotherly, *USA Today* (December. 2005).
30. M. Malone. Is Google Turning Into Big Brother? `http://abcnews.go.com/print?id=5727509` (September, 2008).
31. J. Domingo-Ferrer. A three-dimensional conceptual framework for database privacy. In eds. W. Jonker and M. Petkovic, *Secure Data Management*, vol. 4721, *Lecture Notes in Computer Science*, pp. 193–202. Springer, (2007).
32. P. Samarati, Protecting respondents' identities in microdata release, *IEEE Transactions on Knowledge and Data Engineering*. **13**(6), 1010–1027, (2001).
33. J. Domingo-Ferrer and V. Torra. A critique of k-anonymity and some of its enhancements. In *The Third International Conference on Availability, Reliability and Security – ARES*, pp. 990–993. IEEE Computer Society, (2008).
34. J. Domingo-Ferrer and J. M. Mateo-Sanz, Practical data-oriented microaggregation for statistical disclosure control, *IEEE Transactions on Knowledge and Data Engineering*. **14**(1), 189–201, (2002).
35. Chor, Goldreich, Kushilevitz, and Sudan. Private information retrieval. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, (1995).
36. E. Kushilevitz and R. Ostrovsky. Replication is not needed: Single database, computationally-private information retrieval (extended abstract). In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, (1997).
37. D. Rebollo-Monedero, J. Forné, L. Subirats, A. Solanas, and A. Martínez-Ballesté. A collaborative protocol for private retrieval of location-based information. In *Proceedings of the IADIS International Conference on the e-Society*, (2009).
38. Goopir. `http://crises-deim.urv.cat/unescoprivacychair/`.
39. M. Gruteser and D. Grunwald. Anonymous usage of location-based services through spatial and temporal cloaking. In *Proceesings of MobiSys 2003: The First International Conference on Mobile Systems, Applications, and Services.*, pp. 31 – 42. USENIX Association, ACM, Sigmobile, ACM (May, 2003).
40. B. Gedik and L. Liu. A customizable k-anonymity model for protecting lo-

cation privacy. In *Proceedings of the IEEE International conference on Distributed Computing Systems (ICDS'05)*, pp. 620 – 629, (2005).

41. B. Gedik and L. Liu, Protecting location privacy with personalized k-anonymity: Architecture and algorithms, *IEEE Transactions on Mobile Computing*. **7**(1), 1 – 18 (January, 2008).
42. C. Chow, M. F. Mokbel, and X. Liu. A peer-to-peer spatial cloaking algorithm for anonymous location-based services. In *GIS '06: Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems*, pp. 171–178. ACM (November, 2006).
43. J. Domingo-Ferrer. Microaggregation for database and location privacy. In eds. O. Etzion, T. Kuflik, and A. Motro, *NGITS*, vol. 4032, pp. 106–116. Springer, (2006).
44. A. Solanas and A. Martínez-Ballesté. Privacy protection in location-based services through a public-key privacy homomorphism. In eds. J. Lopez, P. Samarati, and J. L. Ferrer, *EuroPKI*, vol. 4582, *Lecture Notes in Computer Science*, pp. 362–368. Springer, (2007).
45. A. Solanas and A. Martínez-Ballesté, A TTP-free protocol for location privacy in location-based services, *Computer Communications*. **31**(6), 1181–1191 (April, 2008).
46. G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K. Tan. Private queries in location based services: Anonymizers are not necessary. In *SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pp. 121 – 132. Vancouver, BC, Canada, ACM (June, 2008).
47. M. Duckham and L. Kulit. A formal model of obfuscation and negotiation for location privacy. In *Pervasive Computing*, vol. 3468, *LNCS*, pp. 152–170. Springer Berlin / Heidelberg, (2005).
48. C. A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, and P. Samarati. Location privacy protection through obfuscation-based techniques. In eds. S. Baker and G. Ahn, *Data and Applications Security*, vol. 4602, *LNCS*, pp. 47 – 60. IFIP, (2007).
49. M. L. Yiu, C. S. Jensen, X. Huang, and H. Lu. Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services. In *IEEE 24th International Conference on Data Engineering ICDE'08*, pp. 366–375, (2008).
50. K. Ouafi and R. C.-W. Phan. Privacy of recent RFID authentication protocols. In *ISPEC*, vol. 4991, *Lecture Notes in Computer Science*, pp. 263–277. Springer, (2008).
51. A. Juels, R. L. Rivest, and M. Szydlo. The blocker tag: Selective blocking of RFID tags for consumer privacy. In *SIGSAC: 10th ACM Conference on Computer and Communications Security*. ACM, (2003).
52. A. Juels. Minimalist cryptography for low-cost RFID tags. In eds. C. Blundo and S. Cimato, *Security in Communication Networks*, vol. 3352, *Lecture Notes in Computer Science*, pp. 149–164. Springer, (2004).
53. C. Castelluccia and M. Soos. Secret Shuffing: A Novel Approach to RFID Private Identification. In *Conference on RFID Security*, pp. 169–180, Malaga,

Spain (July, 2007).

54. J. Cichon, M. Klonowski, and M. Kutylowski. Privacy protection in dynamic systems based on RFID tags. In *International Workshop on Pervasive Computing and Communication Security (PerSec)*, pp. 235–240. IEEE Computer Society, (2007).
55. A. Juels and S. A. Weis. Defining strong privacy for RFID. In *International Workshop on Pervasive Computing and Communication Security (PerSec)*, pp. 342–347. IEEE Computer Society, (2007).
56. M. Ohkubo, K. Suzuki, and S. Kinoshita. Efficient hash-chain based RFID privacy protection scheme. In *International Conference on Ubiquitous Computing - Ubicomp, Workshop Privacy: Current Status and Future Directions*, (2004).
57. G. Tsudik. YA-TRAP: Yet another trivial RFID authentication protocol. In *PerCom Workshops*, pp. 640–643. IEEE Computer Society, (2006).
58. R. D. Pietro and R. Molva. Information confinement, privacy, and security in RFID systems. In *ESORICS*, vol. 4734, *Lecture Notes in Computer Science*, pp. 187–202. Springer, (2007).
59. M. Conti, R. Di Pietro, L. V. Mancini, and A. Spognardi. FastRIPP: RFID Privacy Preserving protocol with Forward Secrecy and Fast Resynchronization. In *33th Annual Conference of the IEEE Industrial Electronics Society (IEEE IECON 07)*, pp. 52–57, Taipei, Taiwan (November, 2007).
60. S.-C. Kim, S.-S. Yeo, and S. K. Kim. MARP: Mobile Agent for RFID Privacy Protection. In eds. J. Domingo-Ferrer, J. Posegga, and D. Schreckling, *International Conference on Smart Card Research and Advanced Applications – CARDIS*, Lecture Notes in Computer Science, Tarragona, Spain (April, 2006). IFIP, Springer-Verlag.
61. A. Solanas, J. Domingo-Ferrer, A. Martínez-Ballesté, and V. Daza, A distributed architecture for scalable private RFID tag identification, *Computer Networks*. **51**(9), 2268–2279, (2007).