2007 IEEE International Conference on Granular Computing

Privacy in the Information and Communications Technologies

Antoni Martínez-Ballesté and Agusti Solanas
CRISES Research Group - UNESCO Chair in Data Privacy
Dept. of Computer Engineering and Maths
Rovira i Virgili University
E-43007 Tarragona, Catalonia
Email: {antoni.martinez,agusti.solanas}@urv.net

***Abstract*— Information and Communications Technologies (ICT) are fostering the appearance of a plethora of new services and applications. On the one hand, ICT will improve the efficiency of a variety of processes (*e.g.* supply chains, e-commerce, etc.) and they will make easy the development of new services. On the other hand, the massive deployment of these new technologies may lead to the invasion of users privacy.**

In this article, we analise four of the main technologies which can put the privacy of their users in jeopardy: Internet browsing, location-based services, RFID technology and statistical databases.

I. Introduction

The Information and Communications Technologies (ICT) pave the way for the storage and analysis of huge amounts of data. However when collecting and, specifically, when analysing all these data, the privacy of individuals must be taken into account. In that sense, there must exist a balance between the right of the society to information and the right of individuals to privacy. Next, we show some examples of information gathering and analysis using ICT:

- **E-commerce:** Commercial information that is collected from fidelity cards, e-commerce logs, sales software, etc. From this information customer's habits can be inferred, so these data analysis can feed marketing strategies.
- **Statistical Data:** Statistical agencies collect data from individuals in order to analyse them and publish statistical reports. These data, which are kept in large databases, are usually released to third parties (*e.g.* marketing companies, health institutions, etc.) and, thus, statistical agencies who collected the data lose control on the data once they are released. The privacy of the respondents must be guaranteed when their data is released.
- **Location Information:** Location-based services make use of the location of the users so as to provide them with personalised services. The location of the users can be used to infer their habits, thus, these data must be managed carefully.
- **Internet search:** Internet advertising companies use browsing information of Internet users to provide them with personalised advertisements. This practice could be seen as a privacy invasion if users are not properly informed.

A. Enforcing the right to privacy

The Universal Declaration of Human Rights states in Article 12 that *"no one shall be subjected to arbitrary interference with his privacy, family, [...] Everyone has the right to the protection of the law against such interference attacks"*. ICT must technically enforce the above right in the Information Society.

Moreover, the European Community Directive on Privacy and Electronic Communications [2] deals with the relationship between ICT and the privacy and security of their users. In addition to classical concepts regarding information security (*e.g.* encryption or digital signatures), the directive elaborates on some issues related to the privacy of the users, specifically those related to LBS. In fact, the location information is directly related to the privacy of the users since it could be used to determine their consumer habits. According to this directive, the LBS provider *"must inform the users or subscribers [...] whether their location data will be transmitted to a third party"*.

Unfortunately, most of the efforts regarding security in ICT are about enforcing the "corporate security". For example, e-commerce systems are perfectly protected against intrusions and data sniffing, while consumer security and specially consumer privacy tend to be disregarded.

On the other hand, since September 11th 2001, many governments make use of ICT and huge data warehouses to profile and track people for the sake of national security and war against global terrorism. Moreover, many users use the Internet as a communication channel for their harmful interests (paedophilia, terrorism, etc.). Thus, the real challenge in privacy is to be able to track "the bad people" preserving the privacy of the other.

B. Contribution and plan of this paper

In this paper we present a brief summary of the privacy aspects of ICT. The article is not merely centred in the usage of Internet and computers, since we also elaborate on current research topics such as location privacy and privacy related to statistical databases. The structure of the paper is as follows: Section II briefly summarises privacy aspects related to Internet browsing, Section III focuses on privacy in location-based services, Section IV studies the privacy threats due

0-7695-3032-X/07 $25.00 © 2007 IEEE
DOI 10.1109/GrC.2007.54

IEEE computer society

to the world-wide deployment of the RFID technology and Section V analises the privacy issues in statistical databases. Finally Section VI concludes this article and points out some future challenges.

II. Privacy of Internet users

The web is surfed by millions of people, 24 hours a day. The web is not only browsed for obtaining all kinds of information. A huge number of e-mails is being stored in servers such Hotmail or Gmail, millions of searches per hour are served by Google, Youtube hosts plenty of videos, etc. In this section we briefly summarise the privacy aspects related to Internet browsing.

A. Tracking cookies and spyware

Websites can collect large amounts of information about their users and their habits by means of cookies. A cookie is a small text file stored into our system by the web we are browsing. Cookies were designed for "innocent" purposes, such as easier returns to already visited login forms. However cookies may not be as innocent as they seem to be: Several companies (*e.g.* Google, Doubleclick, etc.) send tracking cookies to our systems to track our browsing activity. The goal of these cookies is to deliver data, *e.g.* advertisements, targeted to the interests of the user. In that sense, the information collected may be sold to other companies or merged with other databases in order to build accurate profiles of users.

With regard to this, the *Online Profiling Recommendations* by the US Federal Trade Commission state that *"consumers may be totally unaware that their activities online are being monitored"*. Fortunately, almost all the web advertising companies offer the so-called *optout* options to prevent their cookies to track the Internet user [9].

Some security tools [1] consider tracking cookies as a tool similar to spyware. Spyware is a small peace of software that is installed in a computer without the consent of the user. Its main task is to send information to a remote database. This information is used to determine the users' habits. Moreover, spyware can be programmed to send chat sessions or even passwords.

Spyware does 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 (*sic*). Spyware can come bundled with free software and, in fact, any software can potentially include spyware. Finally, spyware may be installed as browser helpers or toolbars (*e.g.* CoolWebSearch). Besides tracking browsing habits, spyware uses to slow down the system and displays pop-up advertisements randomly.

Despite there exists plenty of cleaning software, completely removing the spyware from our computer is not straightforward [26].

B. Google: the real Big Brother?

Since its foundation, Google has become more than a simple and efficient searching tool. Google is currently a universal data warehouse whose one of its statements is *"to organise the world's information and make it universally accessible and useful"*[4].

Moreover, in the last years Google has launched several tools that, behind the proposal of organising the information of its users, could be a serious threat to their privacy (*e.g.* Web History, Desktop, Picasa, etc.).

The privacy note of Gmail states that *"Google maintains and processes your Gmail account and its contents to provide the Gmail service to you and to improve our services."* [3]. However, despite the user of a Gmail account is aware of that, people sending e-mails to Gmail accounts do not know that their messages are going to be processed without their permission[1].

Of course, Google claims this information is used to help its users and to provide them with interesting and useful services, but due to the large amount of information stored by Google, some criticisms about privacy have arisen [18]. Fortunately, Google's corporate philosophy includes statements such as "Don't be evil".

III. Privacy in location-based services

New technologies and services allow the location of users to be known. If the these locations are collected several times during the day, the movements of users can be tracked. Note that, with location tracking, people habits can be inferred (for instance, it could be inferred whether users frequently visit a certain restaurant or shop). Thus, the location information of a user should be private (of course there could be some exceptions, namely criminals and old people with senile dementia). Location privacy can be defined as preventing someone from knowing the location of a person, either the current one or past locations. Unfortunately, as in the above examples, the knowledge of this location is essential to receive the service.

Mobile devices such as cell phones or laptops have become ubiquitous. They are equipped with a variety of localisation systems that make them proper for making use of the new location-based services (LBS). Some examples of these LBS are tourist information services, route planners, emergency assistance services, etc.

Most of the times, these services are provided by a trusted company (*e.g.* a telecommunications company). However, the massive use of mobile devices pave the way for the creation of ad hoc wireless networks that can be used to exchange information based on locations. When the exchange of location information is done amongst untrusted parties, the privacy of the participants could be in jeopardy.

In LBS, a user u asks the LBS provider P for some information, sending the message $\{ID_u, query, long, lat\}$. Upon this request, P seeks the desired information in its database and returns an appropriate answer to u. Note that, if u sends her exact location to an untrusted LBS provider P_u, it can misbehave because it can relate the real location '$long, lat$' with the unique identifier of u, for instance:

[1] A very different thing is to scan e-mails for spam detection purposes.

- P_u is able to know if u is in front of certain shops or places, so it can flood her with undesired advertisements.
- P_u can track u so it knows where she has been and when.
- P_u can send the identifier of u along with her location to a spammer, and the later can send undesired location-based advertisements to the user.

In order to avert these possible misbehaviour, two main solutions have been proposed:

- *Hiding the real location amongst other users*. By using this technique inspired in the well-known k-anonymity approach [13], [35], [30], P is not able to distinguish u amongst a set of k users because they share the same fake location. This indistinguishability makes difficult the tracking of the user.
- *Giving an inaccurate location to the LBS provider*. The position should be accurate enough so the information received by u is still useful. However, since the locations collected by P are not exact, they become useless to a spammer. Note that, although this seems to be a strong assumption, it is reasonable to believe that the user, who really knows her location, could make a proper use of the information given by the provider. On the contrary, the provider has access to the fake location only.

A. Proposals based on a Trusted Third Party

Most of the existing proposals related to privacy in LBS rely on a trusted third party (TTP) that mediates between users and LBS providers. The so-called *anonymizer* behaves as follows:

- It can delete personal information from the queries of u before sending them to the LBS providers. A variant of this proposal consists of the anonymizer making use of pseudonyms to hide the real identity of the users.
- It can hide the real location of u under a *cloaked region* (*i.e.* a spatial region containing k users) so that each user becomes k-anonymous (*i.e.* not distinguishable amongst $k-1$ other users). According to [22], the cloaked region must fulfil the property of k-anonymity, but must also consider a spatial cloaking. In that sense, the cloaking algorithm considers a minimum and maximum area sizes and the anonymizer uses the requests from other users (and the location data contained in them) to compute a *masked* location, taking into account the value of k and the area requirements. The masked location can be computed as the centroid of the current locations of the users in the cloaked area.

In [15], an efficient algorithm for cloaking is presented. Similar approaches are also presented in [17], [23] and [7]. A slightly different approach is presented in [16]. It consists of a decentralised architecture based on a P2P (peer-to-peer) system. Users utilise P2P techniques to form clusters according to their real location and a certain degree of k-anonymity. In addition, a set of pseudonym servers are used to hide the real identity of the user.

These proposals have an important drawback: The anonymizers store the identifiers of their users. Moreover, they can store the location information contained in the queries so the location of their users can be tracked or disclosed to third parties. Users must trust the anonymizer completely.

B. A TTP-free proposal

In [33] a TTP-free approach is presented. It consists of providing the server with a k-anonymous location computed by user u, who asks her companions (the set of $k-1$ nearby users which u wants to get "anonymized" with) for their location. Then, the centroid of the location of u and her companions is computed and sent to the server as the k-anonymous location.

In order to prevent u from knowing the real location of her companions, the companions add some noise to their real locations (x_i, y_i) prior to sending them to u:

$$(\hat{x}_i, \hat{y}_i) = (x_i, y_i) + (N_i^x, N_i^y)$$

This noise follows a Gaussian distribution with null average $\sim N(0, \sigma)$. Note that the overall noise added to the real location of the companions tends to 0 when k is big enough. Thus, the centroid of the real location of u and her companions is properly computed.

Unfortunately, this approach has a limitation. Due to the use of Gaussian noise with null average, users should not use the same technique repeatedly without changing their real location because of the cancellation of the added noise (*i.e.* the average masked location of these users tends to the real location).

To avoid this limitation, the proposal in [33] makes use of a public-key privacy homomorphism [27]. The privacy homomorphism used must be additive, *i.e.* one of the operations maps the addition of the unencrypted values. This results in a third party being able to add values but not being able to know which values are being added. In addition, the privacy homomorphism used must be public-key, *i.e.* only the owner of the secret key is able to retrieve the result of the addition. For instance, the Okamoto-Uchiyama [25] public-key cryptosystem has an additive homomorphic property.

In this case there is a public key infrastructure supplying to LBS providers public keys of a public-key probabilistic and additive privacy homomorphism. This privacy homomorphism has both encryption and decryption functions $E_{pk}(\cdot)$ and $D_{sk}(\cdot)$. The operation mapping the addition of the encrypted values is $\Gamma(\cdot)$.

Let P_u be an untrusted LBS provider whose secret and public keys in the PKI are sk_{P_u} and pk_{P_u} respectively. To summarise, either u and her companions make use of function E_{pk} to encrypt their real (in fact masked with some Gaussian noise) location. Despite being encrypted, the locations can be added by u, who is not aware of the (masked) locations of her companions.

Finally, u sends the added value $\{\bar{U}^{\Gamma}, k\}$ to the provider, who uses $D_{sk}(\cdot)$ to obtain the *real* addition. The centroid is obtained dividing by k. Note that u also sends the message to her companions, so that they can use it with further queries to the LBS provider.

IV. Privacy in RFID

Radio-Frequency IDentification (RFID) devices have an important presence in our daily life, even when we do not see them, and they 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. As a result, almost every object is liable to carry an RFID tag.

RFID devices can be seen as a proper substitute of bar codes since they are mainly used to identify objects. Unlike bar codes, RFID devices allow objects to be identified without visual contact and help in improving and automating many processes. This is possible due to the ability of RFID tags for being read fast and in parallel.

An RFID system consists of two main components:

- *RFID tags*: They are small passive devices with a variety of possible appearances from stickers to small grains embedded in official documents. A tag basically consists of a microchip and a metal coil, which acts as an antenna. In some cases, it can also contain a battery and some other microchips intended for increasing its computational power.
- *RFID readers*: They are active devices used to read the information stored in the tags. In a nutshell, readers emit a radio wave so that all tags in their range *answer* by broadcasting their embedded information (*i.e.* a set of bits). This information, generally known as Electronic Product Code (EPC), is usually the identifier of the object into which the tags are stuck.

A priori, RFID tags can be read by any unauthorised reader in their cover range and, for this reason, some security and privacy issues must be taken into account. In fact, an eavesdropper, properly equipped with a reader, could collect several information from people carrying RFID tags. Thus, an attacker would be able to get lots of information from the scanned people *e.g.* the brand of their clothes, the amount of banknotes they have in their pockets, the use of prosthesis or medicines, etc. Moreover, making use of several readers strategically deployed, it could be possible to track the location of tags and, consequently, the motion of the scanned people.

The aforementioned problems can give cause for concern because a huge deployment of the RFID technology could pave the way for a *Big Brother effect*. Hence, if no practical solutions are proposed, this huge deployment will not be likely to happen.

Next, we give a brief overview of some proposals that can be found in the literature (an interesting survey on privacy in RFID can be found in [5]):

- *Killing and Sleeping commands.* The `Kill` command is a trustworthy solution for these tags although unfortunately it permanently disables the functionality of the tags. This technique is equivalent to extract the tag from the product once it has been bought. A different solution consists in using the `Sleep` command for temporarily disabling the tags. However, these method could pose a problem for people with little expertise in technology.
- *Use of pseudonyms.* In [19] Juels suggested the use of a collection of pseudonyms 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, thus, it will be able to match the identifier with the tag. On the contrary, 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.
- *Simple cryptographic protocols.* Several cryptographic protocols related to RFID security and privacy can be found in the literature. However, some of them suffer from drawbacks that make its implementation infeasible, at least using simple and cheap RFID tags. Some proposals are the OSK/AO protocol [24], the YA-TRAP protocol [36], the Deterministic Hash-locks [29] and the Improved Randomised Hash-locks [20]. The Hash-locks approach has scalability problems that are solved in [31] by means of a scalable protocol based on the distribution of RFID readers.

V. Privacy in statistical databases

Statistical databases contain micro-data sets which are built with data from individuals. Not only statistical agencies collect micro-data: Web advertising companies such those described in Section II also collect data from individuals. These data sets are statistically analysed (averages, standard deviations, histograms, etc.) by statistical agencies who also make their data web-accessible to user queries or release their data to third parties.

Privacy related to statistical databases can be seen from three different points of view (also known as *the three dimensions of data privacy*): respondent privacy, owner privacy and user privacy [10].

In this section we elaborate on these privacy aspects of statistical databases, specially respondent privacy.

A. Respondent privacy

The main goal of respondent privacy in statistical databases is preventing 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 is, obviously, a serious privacy loss. One could think that, deleting the direct identifiers (*e.g.* the name of the respondent) from the data before publishing them could be enough. Unfortunately, protecting the privacy of the respondent is not so straightforward.

Data attributes can be classified into direct identifiers, quasi-identifiers, sensitive information and non-sensitive information. Table I shows an example data set D where "Weight"

TABLE I

EXAMPLE DATA SET. "NAME" IS AN IDENTIFIER TO BE SUPPRESSED BEFORE PUBLISHING THE DATA SET.

Name	ZIP	Weight	Height	Salary	Car
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

TABLE II

THE EXAMPLE DATA SET WITH IDENTIFIERS SUPPRESSED AND ATTRIBUTES "WEIGHT" AND "HEIGHT" MICRO-AGGREGATED WITH $k = 3$.

ZIP	Weight	Height	Salary	Car
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

and "Height" are quasi-identifiers and "Salary" is a sensitive attribute.

From the example data set we can build a data set D' with no sensitive information (that is D with the "Salary" attribute suppressed) and a data set D_p with no direct identifiers (thus, a "protected" data set). If an intruder, who owns D', makes queries to the publicly-available data set D_p he will disclose salaries simply linking the quasi-identifiers between D' and D_p. For instance, Maggie Simpson, with ZIP 43152, weight 100 and height 165, earns a salary of 17,000. Moreover, if the intruder only has access to D_p, he can sort the registers by "ZIP" and for instance, disclose the salary of the "tallest man in town".

In order to protect these databases, Statistical Disclosure Control (SDC) techniques must be applied to databases before publishing them. There is a variety of techniques for protecting micro-data, for instance micro-aggregation. In this technique, 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 so that 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 II shows the D_p data set once quasi-identifiers have been masked with $k = 3$.

Micro-aggregation has been used for several years in different countries: it started at Eurostat [21] in the early nineties, and has since then been used in Germany [28] and several other countries [14].

Naturally, applying masking methods results in some *information loss* of the data. Thus, masking methods look for the best trade off between information loss and *disclosure risk* [12]. That is, both the risk of disclosing private confidential information and the loss of data utility should be kept below reasonable thresholds.

There exist several algorithms for micro-aggregation, which must provide good cluster homogeneity at a reasonable computational cost [11], [32], [8], [13], [34].

B. Owner privacy

Owner privacy is about two or more autonomous entities being able to compute queries across their databases (which are containing *their* data) in such a way that only the results of the query are revealed.

Next we show an example: let there be four thematic parks that take part in a statistical survey. They are interested in this survey and its results (because it can be used to improve their services), but they do not want to release their data so that private company results and statistics remain secret.

The so-called Privacy-Preserving Data Mining Techniques (PPDM), based on multiparty computation, cryptographic tools and SDC techniques, are used for providing data warehouses with owner privacy.

C. 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, in Google-like databases or in queries to LBS providers). Note that user privacy is not only related to statistical databases.

We also show two examples to illustrate user privacy in databases. A first example is the profiling of users regarding their queries to Google. User privacy would consist of a user being able to query a database, preventing the database owner from profiling and tracking the user. A second example: A scientist having a great idea wants to know if it has already been patented. She queries a patent database, but its administrator, who has access to the scientist query, tries to patent the idea first.

A naive approach for user privacy is the downloading of the entire database, so the user could query the database locally. Of course, this is not applicable in large databases.

Private Information Retrieval (PIR) techniques, look for proposals for guaranteeing the user privacy in databases. Unfortunately, there are currently only theoretical or non-realistic techniques [6].

VI. CONCLUSIONS

In this survey we have presented the main topics on privacy related to the Information and Communications Technologies. The massive collection, analysis and usage of data are inherent to the great advances achieved in ICT and the Information Society. We have seen that the Internet is able to profile their users and how our privacy can be put in jeopardy due to spyware risks or Google becoming a Big Brother. Privacy is

an important concern regarding LBS and RFID technology, not just location privacy but also user privacy. Finally, the privacy in statistical databases is not achieved only by means of deleting the identifiers. The privacy of ICT users must be seriously taken into account, since the lack of privacy undermines most of all other fundamental rights. There are many parties involved: from service providers to governments. We have seen open problems and techniques to overcome them. The main research lines that should be studied in the future in this field can be summarised as: propose new protocols to properly protect the location of the users of LBS; analise the RFID protocols to make them widely applicable and scalable; improve the SDC methods in order to obtain a better balance between privacy protection and information loss; protect the user privacy in large databases (*e.g.* Internet search engines).

ACKNOWLEDGEMENT

The views expressed by the authors 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 SEG2004-04352-C04-01 "PROPRIETAS" and CONSOLIDER CSD2007-00004 "ARES", and by the Government of Catalonia under grant 2005 SGR 00446."

REFERENCES

[1] Ad-ware by lavasoft. website. http://www.lavasoftusa.com/products/ad_aware_free.php.

[2] Directive 2002/58/EC on privacy and electronic communications. http://europa.eu.int/eur-lex/pri/en/oj/dat/2002/201/20120020731en00370047.pdf.

[3] Gmail privacy policy. website. http://mail.google.com/mail/help/privacy.html.

[4] Google corporate information: Company overview. website. http://www.google.com/corporate/.

[5] J. Aragones, A. Martinez-Balleste, and A. Solanas. A brief survey on rfid privacy and security. In *World Congress on Engineering*, 2007.

[6] D. Asonov. *Querying Databases Privately: A New Approach to Private Information Retrieval*. Springer, 2004.

[7] R. Cheng, Y. Zhang, E. Bertino, and S. Prabhakar. Preserving user location privacy in mobile data management infrastructures. In *Proceedings of Privacy Enhancing Technology Workshop (PET)*, 2006.

[8] D. Defays and P. Nanopoulos. Panels of enterprises and confidentiality: the small aggregates method. In *Proc. of 92 Symposium on Design and Analysis of Longitudinal Surveys*, pages 195–204, Ottawa, 1993. Statistics Canada.

[9] P. Dixon. Consumer tips: How to opt-out of cookies that track you. website, 6 2006. http://www.lavasoftusa.com/products/ad_aware_free.php.

[10] J. Domingo-Ferrer. The three dimensions of data privacy. In *4th VLDB Workshop on Secure Data Management-SDM'2007*, Lecture Notes in Computer Science, page (to appear), 2007.

[11] J. Domingo-Ferrer, A. Martinez-Balleste, J. M. Mateo-Sanz, and F. Sebe. Efficient multivariate data-oriented microaggregation. *International Journal on Very Large Databases*, 15(4):355–369, 2006.

[12] 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.

[13] J. Domingo-Ferrer, F. Sebe, and A. Solanas. A polynomial-time approximation to optimal multivariate microaggregation. *Computer and Mathematics with Applications (in press)*, 2007.

[14] Economic Commission for Europe. Statistical data confidentiality in the transition countries: 2000/2001 winter survey. In *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*, 2001. Invited paper n.43.

[15] B. Gedik and L. Liu. A Customizable k-Anonymity Model for Protecting Location Privacy. 2004.

[16] G. Ghinita, P. Kalnis, and S. Skiadopoulos. PRIVÉ: Anonymous location-based queries in distributed mobile systems. In *Proceedings of the Sixteenth International World Wide Web Conference*, pages 371–380, May 2007. Alberta, Canada.

[17] M. Gruteser and D. Grunwald. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. *Proceedings of the First International Conference on Mobile Systems, Applications, and Services*, pages 31–42, 2003.

[18] J. Hopkins. Once-brotherly image turns Big Brotherly. *USA Today*, 12 2005.

[19] A. Juels. Minimalist cryptography for low-cost RFID tags. In *In Fourth International Conference on Security in Communication Networks-SCN 2004*, volume 3352, pages 149–164. Springer-Verlag, 2004. Lectures Notes in Computer Science.

[20] A. Juels. RFID security and privacy: A research survey. Manuscript, September 2005.

[21] A. De Michelis and A. Chantraine. Memoirs of eurostat: Fifty years serving europe. http://epp.eurostat.cec.eu.int/cache/ITY_OFFPUB/KS-49-02-183/EN/KS-49-02-183-EN.PDF.

[22] M.F. Mokbel. Towards Privacy-Aware Location-Based Database Servers. *International Workshop on Privacy Data Management, PDM, Apr*, 2006.

[23] M.F. Mokbel and Chi-Yin Chow. Challenges in preserving location privacy in peer-to-peer environments. In *Seventh International Conference on Web-Age Information Management Workshops (WAIM '06)*, 2006.

[24] 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.

[25] T. Okamoto and S. Uchiyama. A new public-key cryptosystem as secure as factoring. *Lecture notes in computer science*, pages 308–318.

[26] B. Posey. How spyware and the weapons against it are evolving. website, 11 2004. http://www.windowsecurity.com/articles/Spyware-Evolving.html.

[27] R.L. Rivest, L. Adleman, and M.L. Dertouzos. On data banks and privacy homomorphisms. *Foundations of Secure Computation*, pages 169–178, 1978.

[28] M. Rosemann. Erste ergebnisse von vergleichenden untersuchungen mit anonymisierten und nicht anonymisierten einzeldaten amb beispiel der kostenstrukturerhebung und der umsatzsteuerstatistik. In *G. Ronning and R. Gnoss (editors) Anonymisierung wirtschaftsstatistischer Einzeldaten, Wiesbaden: Statistisches Bundesamt*, pages 154–183, 2003.

[29] R. Rivest S. Weis, S. Sarma and D. Engels. Security and privacy aspects of low-cost radio frequency identification systems. In W. Stephan D. Hutter, G. Mller and M. Ullmann, editors, *International Conference on Security in Pervasive Computing - SPC 2003*, volume 2802, pages 454–469. Springer-Verlag, 2003.

[30] P. Samarati. Protecting respondents' identities in microdata release. *IEEE Transactions on Knowledge and Data Engineering*, 13(6):1010–1027, 2001.

[31] A. Solanas, J. Domingo-Ferrer, A. Martinez-Balleste, and V. Daza. A distributed architecture for scalable rfid identification. *Computer Networks*, 51:2268 – 2279, 2007.

[32] A. Solanas and A. Martinez-Balleste. V-MDAV: a multivariate microaggregation with variable group size. In *Proceedings in Computational Statistics COMPSTAT 2006*, pages 917–925. Springer's Physica Verlag, 2006.

[33] A. Solanas and A. Martinez-Balleste. Privacy protection in location-based services through a public-key privacy homomorphism. In *Fourth European PKI Workshop: theory and practice*, 2007.

[34] A. Solanas, A. Martinez-Balleste, J. M. Mateo-Sanz, and J. Domingo-Ferrer. Towards microaggregation with genetic algorithms. In *3rd IEEE Conference On Intelligent Systems*, pages 65–70, Westminster, 2006. IEEE Computer Society Press.

[35] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, 10(5):557–570, 2002.

[36] G. Tsudik. YA-TRAP: Yet another trivial RFID authentication protocol. In *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06)*, pages 640–643, 2006.