Controlling Your Personal Information Disclosure

NORJIHAN ABDUL GHANI,
Information Science Department
University of Malaya
50603 Kuala Lumpur
MALAYSIA
norjihan@um.edu.my

ZAILANI MOHAMED SIDEK
Centre for Advanced Software Engineering (CASE),
Universiti Teknologi Malaysia,
City Campus, Jalan Semarak,
54100 Kuala Lumpur,
MALAYSIA
zailani@citycampus.utm.my

Abstract: - As organizations come to rely on the collection and use of personal information in order to complete the transaction and providing good services to their users, more and more user personal information is being shared with web service providers, leading to the need to protect the privacy. Within the electronic scenario, personal information have been collected, stored, manipulate and disclosed without owner consent. This paper will discuss on the relationship between personal information and its privacy. Besides that, we extended the model introduced by Al-Fedaghi as a way to control the personal information disclosure.

Key-Words: - personal information, privacy, personal information flow model

1 Introduction

During the past decade, there has been an increasing number of personal information that is being collected, used and disclosed, and the expansion of the World Wide Web has significantly facilitated to this growth. Today, more people rely on electronic commerce in their daily tasks. People not only buying groceries, booking air tickets via online application, but any other tasks can be done by using e-commerce application. Today, the emerging trends of e-commerce have become more convenient and easy for people to do anything online.

Peri, 1998 said that personal information has become the "basic fuel" for modern business and government to carry out their services (as cited in [1]). Such parties is collecting, analyzing, storing and sharing more personal information. Unfortunately, people don't realize that once they gave their personal information, they no longer have authorities to control it. People have lost their ownership once they released their personal information. They do not have their privacy towards the personal information anymore. The more personal information has been disclosed, the less privacy they have. It caused the ability to protect that information and enforce privacy polices becomes more important.The main issue here is, people have less control over what types of information about them have been collected, used, stored and disclosed by various agencies; both private and government sectors.

The rest of this paper as follows; Section 2 will give an overview of personal information and its relationship with privacy. It's also covers the important of privacy for personal information. Section 3 explains the OECD's principles that have been adopted as guidelines in this paper and Section 4 continued the discussion on the extended the personal information flow model introduced in [2].

2 Personal Information and Privacy

Data is important in any transaction; either off-line transaction or online transaction. Unfortunately, not many people realize and understand how important and valuable their personal information. Some personal information can be classified as sensitive and need to keep as a private information. Some of personal information are sensitive but, no need to keep it private. There are four types of data involved in processing [3]:

i) *Personal data* : any data that can be used to identify a person such as name, address, telephone number.
ii) *Sensitive data* : any data that disclose information about racial or ethnic origin, religious, philosophical or other belief, political opinion, membership of

parties, as well as personal data disclosing health such as health history, race.

iii) *Identification data* : personal data that permit the direct identification of the data subject such as DNA, identity card number

iv) *Anonymous data* : any data that cannot be associated to any identified or identifiable data subject such as gender, type of disease.

From the above classification, the first three types of data can be considered as private information. Private information is personal information that requires protection due to risks that could result from its disclosure, alteration, or destruction. This personal and private information should be protected to ensure the privacy.

Personal data will go through a process to become information. There are various definitions for personal information. In 93, personal information is defined as any information that is related to the individual person. Personal information is any linguistic expression that has referent(s) of type person [4]. There are three categories of referent(s) :

- Zero personal information – having no individual referent
- Atomic personal information – having a single referent
- Compound personal information – having more than one referents

Heikkinen et al. [5] define personal information as any information that is related to the individual person.

In web-based environment, personal information is disclosed by the data owner and used by the organizations. The organization will collect, store, manipulate information to fulfill their organizations needs. From information system views, information privacy can protect individuals from misuse of data, or unauthorized access to, or modification of information could adversely affect, or be of risk to the owner of that information. Information play a fundamental role in privacy domain as they shall be collected, manipulated, stored, and disclosed according their needs. Clarke, 1999 define privacy as below :

> *the interest that individuals have in sustaining a 'personal space', free from interference by other people and organisations.*

Clarke, 1999 claims that there are several dimensions of privacy :

- *privacy of the person*, sometimes referred to as 'bodily privacy'. This is concerned with the integrity of the individual's body. Issues include compulsory immunisation, blood transfusion without consent, compulsory provision of samples of body fluids and body tissue, and compulsory sterilisation;
- *privacy of personal behaviour*. This relates to all aspects of behaviour, but especially to sensitive matters, such as sexual preferences and habits, political activities and religious practices, both in private and in public places. It includes what is sometimes referred to as 'media privacy';
- *privacy of personal communications*. Individuals claim an interest in being able to communicate among themselves, using various media, without routine monitoring of their communications by other persons or organisations. This includes what is sometimes referred to as 'interception privacy'; and
- *privacy of personal data*. Individuals claim that data about themselves should not be automatically available to other individuals and organisations, and that, even where data is possessed by another party, the individual must be able to exercise a substantial degree of control over that data and its use. This is sometimes referred to as 'data privacy' and 'information privacy'.

Personal information only should be kept by the owner itself or control the disclosure in order to ensure its privacy. But, in web-based application, this information should be disclosed in order to fulfill the transaction. Although the private information is being disclosed, normally, for the security and privacy reason, it can't be accessed by unauthorized users. For this reason, there are three main issues that need to be considered :

i) personal information shouldn't be access by unauthorized users.

ii) only required personal information will be posed

iii) personal information can't be passes for those do not need the information

3 OECD Principles

Over a few years ago, there are a number of guidelines exists to protect the PI. This guideline is important to ensure the PI privacy. In 1980, Organization of Economic Cooperation and Development or OECD, adopted and expanded eight principles as part of the "Guidelines on the protection of Privacy and Transborder Flows of Personal Data". The OECD has therefore been focusing on how these Guidelines may best be implemented in the 21st century to help ensure respect of privacy and protection of personal data on

line. There are eight principles describe by OECD's guidelines as following [7]:

1) *Collection Limitation Principle*
There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

2) Data Quality Principle
Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

3) *Purpose Specification Principle*
The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

4) *Use Limitation Principle*
Personal data should not be disclosed, made available or otherwise used for purposes other than those specified except: *a)* with the consent of the data subject; or *b)* by the authority of law.

5) *Security Safeguards Principle*
Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data.

6) *Openness Principle*
There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

7) *Individual Participation Principle*
An individual should have the right:
   a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him;
   b) to have communicated to him, data relating to him
      - within a reasonable time;
      - at a charge, if any, that is not excessive;
      - in a reasonable manner; and
      - in a form that is readily intelligible to him;
   c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and
   d) to challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed or amended.

8) *Accountability Principle*
A data controller should be accountable for complying with measures which give effect to the principles stated above.

Next section will explain on how this guideline has been incorporated to personal information flow model. This guideline has been used as guides of extended the work done by Al-Fedaghi, 2005.

4 Personal Information Flow Model

Previous section discussed on OECD's principles that have been adopted to protect the PI privacy. In [4], Hippocratic Database (HDB) is a database concept that adopted these eight principles to come out with ten principles to protect the PI privacy in a database system. Personal information flow model (PIFM) has been introduced by Al-Fedaghi in [2, 8, 9] consists of four main modules or phases; creating, collecting, processing and disclosing the PI.

Besides this four phases, we decide that it' important to control the PI before disclose it. Figure 1 shows an extended version of PIFM introduced by Al-Fedaghi. Al-Fedaghi, in his paper introduced four phases in PIFM. This model reflects the personal information pattern that guides and restricts relationship among objects (e.g., proprietors, pocessors, miners) and phases [2]. The purpose is to show the relationship of recognizing, understand and manipulate personal information. This model complements other descriptions such as the data protection EU directive as an explicit representation of personal information flow in realty. EU directive lumps together all processing of personal data to mean collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction [2]

Dorsey, 2000 introduced different types of categories applied to information mentioned in personal information : retrieving information, evaluating/assessing information, organizing information, analyzing information, presenting information, securing information, and collaborating around information (as cited in Al-Fedaghi, 2). In the context of personal information privacy, this category can be applied to several phases such as creating, collecting, processing, controlling and disclosing the personal information.

But, in a way to protect the personal information, there is a need to control the personal information disclosure. Because of that, we add one phase between processing and disclosing personal information. An extended model was carried out to adopt the principles introduced by OECD "Guidelines on the protection of Privacy and Transborder Flows of Personal Data" discussed above. This model stated that any personal information should be disclosed only to authorize users, with a specific purpose and for a limited time. Because of this reason, we add another phase named "*Controlling the personal information*" before "*Disclosing the personal information*" phase.

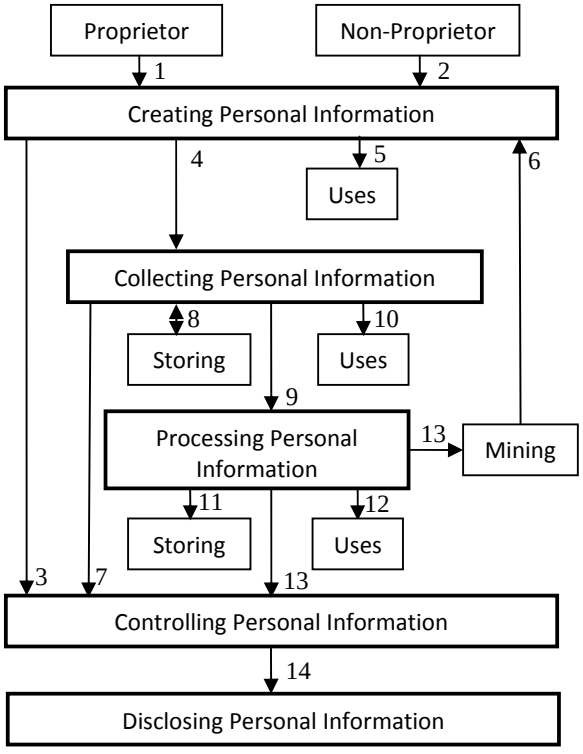

Figure 1 : Personal Information Flow Model

As in Figure 1, there are five main phases in PIFM. These five phases explain how the personal information is created, collected, processed, controlled and disclosed.

Creating personal information

Creating personal information is the first phase on the PIFM. Personal information can be created by two parties; proprietor and non-proprietor (e.g. medical diagnostic procedures performed by physicians) or by deduced by someone (e,g. data mining that generates new information from existing information)[2]. Figure 1 show that personal information can be created at point labeled 1, 2 and 6. Any atomic personal information of an individual is proprietary personal information of its proprietor. Once the personal information have been created, it can be either used (point 5) or collected (point 4) or go to controlling phase before disclose it (point 3). Uses means that the personal information is used in decision making process. Point 3 stated that the personal information should be controlled before disclosed it. It means that the personal information is only will be disclosed if it passes the fourth phase.

Collecting Personal Information

After the personal information is created, it can be collected at point 4. Personal information is collected from various sources and for various purposes of collected. The collected personal information can be either keep as records for future used (point 8), used it (point 10), process the personal information (point 9) or proceed to controlling phase (point 7).

Processing Personal Information

The processing phase of personal information involves acting like storing (point 11), using (point 12) and mining (point 12) the personal information. Personal information is processed based on the purpose it being collected. Besides this three personal information also can be controlled (point 13)

Controlling Personal Information

Previous model introduced by Al-Fedaghi is modeled without "controlling personal information phase". In this paper, we extended the work done by him by adding this phase. In this new era of internet, it's important to control the personal information before disclose it. This phase will check the personal information before it goes to last phase; disclosing the personal information. figure 1 shows all the personal information are controlled at point 3, 7 and 13 before decided either the personal information can be disclosed or not.

Disclosing Personal Information

Disclosing personal information meaning that the personal information is going to be released to insiders or outsiders. Personal information is only being disclosed if it is authorized to do so.

4 Conclusion

In this paper, we extended the work done by Al-Fedaghi on personal information flow model. This model was designed to control the personal information disclosure. Personal information should be disclosed only to authorized users with specific purposes for a limited time

References:

[1] Al-Fedaghi, S. Personal Information eWallet, *2006 IEEE International Conference on Systems, Man, and Cybernetics*, October 8-11, Taipei, Taiwan. (2006a).

[2] Al-Fedaghi, S. Aspects of Personal Information Theory, *Proceedings of the 2006 IEEE Workshop on Information Assurance*, United States Military Academy, West Point, NY. (2006b).

[3] P. Guarda, N. Zannone, Towards the development of privacy-aware systems. Inform. Softw. Technol. (2008).

[4] Al-Fedaghi, S. How to Calculate the Information Privacy. *The Third Annual Conference on privacy, Security and Trust*, St, Andrews, New Brunswick, Canada. (2005).

[5] Heikinen, K., Juha E., Pekka J., and Jari, P. Personalized View of personal information. *WSEAS Transactions on Information Science and Applications*, vol. 2, No. 4, 2004.

[6] We Clarke, R. 1999. Introduction to Dataveillance and Information Privacy, and Definitions and Terms.[Online] Available : http://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html#Priv.

[7] OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. OECD Publications, Paris. Available online at : http://www.uhoh.org/oecd-privacy-personal-data.PDF.

[8] Al-Fedaghi, S. Personal Management of Private Information. Innovations in Information Technology, 2006. Pp 1-5. (2006).

[9] Sabah Al-Fedaghi, "Personal Information Flow Model for P3P", W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, Ispra Italy, October 17-18, 2006.

[10] Perri, 6, The Future of privacy. Volume 1:Private Life and Public Policy, Demos, London, 1998

[11] Dorsey, P/ (2000). What is PKM? http://www.millikin.edu/webmaster/seminar/pkm.html.

[12] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. Hippocratic databases. In *The 28th International Conference on Very Large Databases (VLDB)*, 2002.

[13] EU Directive 95/46/EC – The Data Protection Directive, http://www.dataprotection.ie/viewdoc.asp?m=&fn=/documents/legal/6aii-2.htm#5.