Privacy Reference Monitor – A Computer Model for Law Compliant Privacy Protection

Fei Xu
College of Computer Science and Technology
Beijing University of Technology
Beijing 100124, China
xfeixfei@gmail.com

K.P. Chow
Department of Computer Science
University of Hong Kong
chow@cs.hku.hk

Jingsha He
School of Software Engineering
Beijing University of Technology
Beijing 100124, China
jhe@bjut.edu.cn

Xu Wu
College of Computer Science and Technology
Beijing University of Technology
Beijing 100124, China
xuwu@emails.bjut.edu.cn

***Abstract*—The Internet and computers did not invent or even cause privacy issues. The issues existed long before the creation of computers and Internet. The existence of The Internet, computers and large data storage make it possible to collect, process and transmit large volumes of data, including personal data. In this paper, we shall study the privacy from following two different views, namely legal framework and computer security model, and attempt to identify the difference between them. Because of the difference, we further argue that the current computer security model is not sufficient to support the privacy requirements in the legal framework. We propose a computer model "privacy reference monitor" to handle those unsupported requirements. The design of the privacy reference monitor is privacy policy neutral with a small number of functions. With minimal functionalities, we believe that it is possible to implement a verifiable privacy reference monitor.**

Keywords-Privacy Protection; Computer Model;

I. INTRODUCTION

The Internet and computers did not invent or even cause privacy issues. The issues existed long before the creation of computers and Internet. The existence of Internet, computer and large data storage make it possible to collect, process and transmit large volume of data, including personal data. Today, the study of privacy usually involves different dimensions: laws, ethics and information technology.

As network applications and services become more ubiquitous, many different ways of collecting and accessing user private information have emerged, making people's lives more convenient [1]. At the same time, however, users are often forced to provide their private or sensitive information to service providers. Due to many technical and administrative reasons, user's private information is often poorly managed by service providers and sometimes abused, resulting in serious privacy violations. To give end users privacy they can totally control for the dynamic, pervasive computing environments becomes a major research challenge.

Legal framework and computer security model are frequently used to provide privacy protection in various studies. Since privacy and security are essentially different, we notice one important difference between privacy protection and other security goals is that user's privacy preferences varies and are more complex than security needs.

Previous research on user's privacy preferences showed that privacy preferences vary not only across requesters but also across activities, situations and context [2]. Users may have special expectations for privacy when in particular places or engaged in specific activities. For instance, people would like to share location information with their boss when they are in their office during office hour, but would not like to after work. In a bar or at home, 8:00 a.m. or 10:00 p.m. makes a lot difference in user's privacy access control decision making. So the traditional computer security model is not sufficient in privacy protection and thus further research is needed to address this problem.

The data privacy problem has been studied extensively in the United States and Europe. In the United States, law is introduced in response to a specific incident or fear, rather than to address a condition that has not yet happened. As a result, there are numerous laws related to privacy, such as video privacy protection act in 1988 and telephone consumer privacy protection act in 1991. Depending on different types of personal data, there are different laws, such as medical records law and banking records law [3]. Because of the reactive approach, it is possible that some unforeseeable scenarios exist that against data privacy protection and further legislations will be needed in the future.

On the other hand, the Europe takes a different approach. The European approach sets the framework first (top-down approach) and then builds the rules. This top-down approach led to the European Union Data Directives [4], which made reference to the United States Health, Education and Wealth health information and privacy in 1973. The Directive established the fundamental privacy principles, as codified by the Organization for Economic Co-operation and Development (OECD), which specified the guidelines on the protection of privacy and transborder flows of personal data.

1521-9097/09 $26.00 © 2009 IEEE
DOI 10.1109/ICPADS.2009.63

II. LEGAL FRAMEWORK FOR DATA PRIVACY PROTECTIONG

The Organization for Economic Corporation and Development (OECD) is an independent international organization with voluntary membership. The OECD is actively involved in the area of science and technology and fosters the development and promulgation of standards, polices and regulations. The OECD has foresight to see the need for security and privacy regulations and have established the Guidelines on the Protection of Privacy and Trans-border Flows of Personal Data in 1980 [5].

In the Guidelines, the data controller is defined to be the party who is competent to decide about the contents and use of personal data regardless of whether or not such data are collected, stored, processed or disseminated by that party or by an agent on its behalf. Data controller decides how personal data is collected, processed, stored, released and used, and personal data is any information relating to an identified or identifiable individual (data subject).

When the data controller decides to collect personal data for whatever purpose, he must state the purpose (PP) and the usage of the data explicitly. In the guidelines, eight principles are presented:

- Collection limitation principle: only minimum amount of personal data necessary to perform the purpose (PP) be collected
- Data quality principle: the data controller have to ensure that the data is accurate, complete and current
- Purpose specification principle: every individual must be informed why the data is being collected, how it will be used, to whom it will be disseminated, and how long it will be retained (the purpose PP and the retention period T)
- Use limitation principle: personal data cannot be disclosed or used for any other purposes than those stated at the time of collection
- Security safeguard principle: data controllers should provide adequate safeguards to protect against unauthorized access, alteration, use, release and destruction of the data
- Openness principle: data subjects should receive regular communication from the data controllers holding their personal data which include the data, the procedures used to collect, process, store and release the data, the right to view the data, and the contact information of the data controller
- Individual participation principle: data subjects have the right to obtain their personal data
- Accountability principle: the ultimate responsibility for protecting the personal data rests with the data controller

Directive 95/46/EC, known as the Data Protection Directive (the Directive), was issued in October 1995 by European Parliament and Council [6]. The purpose of the Directive is to protect individuals' personal data and the processing and free movement of this data. The Directive establishes several security and privacy rules which are based the eight principles in the OECD guidelines.

In addition to the above 8 principles, the Directive also says under specified conditions, personal data may be released to two classes of outsiders: third parties and recipient. Third parties are individuals or organizations with whom the controller has established a contractual relationship to perform some aspects of processing personal data. The Directive requires that all privacy provisions and safeguards be invoked in contracts with third parties, and third parties be held accountable for compliance. Recipients are individuals or organizations that are legally entitle to receive processed personal data.

Each country that adopts the Data Protection Directive has its own laws implemented according to the above principles. In Hong Kong, the parties involved are the data subject (data owner) and the data user (the data controller). Based on the above principles, the Hong Kong Privacy Data (Personal) Protection Ordinance defines the following data protection principles in the ordinance [7]:

- Principle 1 - Purpose and manner of collection: This defines the information a data user must give to a data subject when collecting personal data from that subject.
- Principle 2 - Accuracy and duration of retention: Personal data should be accurate, up-to-date and kept no longer than necessary.
- Principle 3 - Use of personal data: Unless the data subject gives consent otherwise personal data should be used for the purposes for which they were collected or a directly related purpose.
- Principle 4 - Security of personal data: Appropriate security measures should be applied to protect the personal data.
- Principle 5 - Information to be generally available: The personal data that hold by the data users should be open and the purposes for which personal data are used should be available to data subjects.
- Principle 6 - Access to personal data: Data subjects should have the rights of access to and correction of their personal data.

III. COMPUTER SECURITY MODEL

Traditional computer security focused on two key aspects: access control and cryptography. Access control determines who (user or process) can access what (data or resource). Cryptography emphasizes on how to hide information from others.

Access control model includes three key components: subject, object and access control policy. The set of entities that can be accessed and consequently need to be protected in a system are called objects. The set of entities that can issue requests to access objects are called subjects. Subjects are active entities while objects are passive ones. Access control policy specifies which subject can access which object. Although very simple, access control model provides a very good abstraction for expressing the access right that any subject can have on any object in a system in which access control is used for providing security protection.

Well-known privacy protection methods such as P3P, EPAL and XACML are primarily access control policy writing languages. P3P is a standard for privacy practice definition that allows users to assess whether privacy practices that a server provides comply with stated privacy requirements [8]. P3P also provides the means for privacy policy specification and exchange. EPAL [9] is a formal language for writing enterprise privacy policies to govern data handling practices while XACML [10] is an access control policy language that can be used for describing policies and access control decisions. The most obvious weakness of these privacy access control policy languages is that they only provide a means for making promises but fail to provide the necessary mechanisms to ensure that these promises can be put into real practices.

Other privacy related access control models are often the extension of the traditional access control models such as mandatory access control (MAC), discretionary access control (DAC) and role-based access control (RBAC) that have been designed mainly for security [11]. P-RBAC model is an extension of the well-known RBAC model and is designed mainly for expressing highly complex privacy-related policies and hence provides more consideration to privacy factors [12]. However, highly complex policy systems are more prone to inconsistencies and the P-RBAC model cannot solve the inconsistency problem effectively. Purpose-based privacy access control model proposed by Sabah regards purpose as a central concept based on which access decisions can be made [13]. The model uses a complex set of purpose hierarchies and user's role hierarchies to manage the mapping between users and purposes. However, merely considering purpose as the only main factor in making access decisions renders the model less appropriate as a general model for privacy protection.

IV. The Gap between the Legal Model and the Computer Security Model

Traditional computer system security concentrates on three key issues: confidentiality, integrity and availability [11]. Computer security professionals argued that data privacy has three key aspects, namely sensitive data, affected parties, and controlled disclosure. These aspects closely resemble the three elements of access control, which are subject, object and access rights [14]. On the other hand, with respect to our discussion in the last two sections, there is no one to one mapping between the principles in the data privacy legal framework and the computer security access control model. Of course, one can always implement a software that supports the data privacy legal framework using the access control mechanisms with significant amount of development effort. One example is the middleware approach proposed by Lioudakis et al. [15]. In this section, we shall analyze the difference between the data privacy legal framework and the access control model.

It is usually believed that data privacy protection can be handled within the information security framework, either through the policy definition of access control model or the confidentiality property of the cryptographic algorithms. On the other hand, referring to the data privacy legal framework defined by the European Data Protection Directives, it is found that out of the eight principles defined, the security safeguard principle explicitly defined the security requirement, which requires a combination of technical (IT security) and organizational (physical, personnel and operational security) security mechanisms. All other principles have not been addressed within the IT security framework explicitly. It is therefore quite easy to have the situation that maximum IT security measures are enforced while there are violations of data privacy protection principles.

By further analyzing the above principles, we can classify them into the following groups according to the technologies being used to handle them:

1. Purpose, binding and retention
2. User rights and data accuracy
3. Security and accountability

We shall examine each of the above group with respect to the legal framework and the computer security model.

A. Purpose and Binding

The purpose and binding group includes the collection limitation principle, the purpose specification principle and the use limitation principle. These three principles emphasize the collected personal data should be used only within the specified purpose and nothing more. The data controller has to specify the specific usage of the personal data to be collected and its corresponding retention period, and should not be used for any other purpose without prior consent from the data subjects. In order to prevent any future unauthorized usage of data, data controller is required to specify the data retention period. All personal data have to be erased, destroyed, and rendered anonymous at the end of the specified period of use.

This is the part that receives least attention from the computer security domain as it is generally believe that these requirements depend very much on the specific application and therefore cannot be represented using any general framework. On the other hand, from the data privacy protection point of view, they are important elements as they are related to the proper usage of the personal data. We therefore propose the privacy protection monitor to support the principles in this group.

B. User Rights and Data Accuracy

The user rights and data accuracy group includes the openness principle, the individual participation principle and the data quality principle. These first two principles require the data controller to have regular communication with data subjects of which personal data are hold, and the data subject has the right to obtain and modify their personal data. The last principle emphasizes the data controller has to ensure that the data is accurate, complete and current with suitable auditing functions. All these three principles require close collaborations between the data controllers and the data subjects.

These second group of requirements requires proper communication channels between the data controllers and the data subjects should be established. The data controllers should inform the data subjects regularly about the data being held by them, and request the data subjects to update the personal data should there be any changes to their data. In order to satisfy the requirements in this group, proper user interaction technologies should be employed, such as electronic mails for notification, web search/input/update forms to allow a data subject to access his/her personal data on the Internet in a safe manner. Without proper user participation, it is impossible to enforce the requirements in this group.

C. *Security and Accountability*

The security and accountability group includes the security safeguard principle and the accountability principle. The requirements of these two principles are similar to the traditional IT security requirements, such as confidentiality, integrity, availability and accountability. Within the IT security framework, it is common to see efforts have spent in these key areas. Examples are using strong encryption algorithms to encrypt person data, proper authentication techniques for access control auditing, extensive logging for personal data access.

D. *The Gap*

With respect to the above three groups of the data privacy principles, it is obvious that IT security only supports requirements in the security and accountability group. The other two groups contain requirements that are neither essential nor not directly related to IT security functions. It is therefore quite common to see a gap between the data privacy protection requirements and the IT security support.

With the above discussion, it is clear that the computer security model fully satisfies the security and accountability group of the data privacy protection requirements. With proper design user interaction model, the data privacy protection requirements of the user rights and data accuracy group can be achieved, subject to correct implementation and thorough testing. On the other hand, the data privacy protection requirements of the purpose and binding group cannot be achieved easily within the current computer security framework and the user interface technology.

To implement a law compliant privacy protection system, we cannot solely rely on proper implementation of IT security techniques. We must emphasize that a good implementation of IT security is a necessity for a law compliant privacy protection system but not sufficient. Moreover, the introduction of the data dissemination functions in the Data Protection Directive complicates requirements in the purpose and binding group. According to the Directive, such functions are necessary in order for proper business flow among different business corporations which may belong to different countries.

In the next section, we shall introduce the privacy reference monitor to handle the purpose and binding requirements, and also the data dissemination requirements

V. Privacy Reference Monitor

One of the commonly used implementation of access control is reference monitor [16]. The reference monitor is an abstract system that mediates all access requests and functions correctly, and is tamperproof. Any access request must go through the reference monitor which grants or denies the access. The monitor must work correctly and the correctness must be verifiable. Tamperproof means the modifications of its functions by any unauthorized person or process must not be possible. Sample implementation is the NSA's Security Enhanced Linux (SELinux). Another key concept in reference monitor is policy neutral, i.e. it can implement any access control policy.

The advantages of reference monitor allow an application system to be decoupled from the access control checking, while the access control checking is divided into the checking implementation and the access control policy. Since the only algorithm implementation in the reference monitor is the access control checking, which make it possible to go through a rigorous verification process. The details of the security policy are encoded in the security policy.

Most research in data privacy protection that attempts to support the legal requirements usually work on a complicated environment [17, 18], e.g. using middleware [15] or multi-layered model [19]. With such complicated models, it is quite difficult to perform a thorough verification.

Our research focuses on defining the primitive checking functions that are necessary to support privacy protection, while the privacy policy could be encoded in the privacy policy. The design of privacy reference monitor follows the same design philosophy of the security reference monitor with the aim to have a simple monitor that can be verified and be able to check the data privacy protection requirements correctly. With a small privacy reference monitor, it should be possible to verify the privacy protection principles are strictly followed.

A. *Privacy Rules*

The privacy reference monitor consists of the privacy checking function and the privacy policy defined by the data controller. The privacy policy is a set of privacy rule with the following form:

Data (D) belongs to data subject (S) can be used by application software (A) owned and executed by the data controller (C) within the retention period (T).

The application software A can be used by users other than the data controller and the usage of A should be controlled by suitable access control policy according to the security safeguard principle.

The entities in the policy rule can be divided into the following two groups:

1. Data group (D,S,T): personal data D belongs to subject S with retention period T
2. User group (A,C): application software A owned and executed by data controller C

Each rule can then be represented in the following forms:

- (D1,S1,T1) , (A1,C1) : *Allowed*
- (D2,S2,T2) , (A2,C2) : *Denied*

The privacy checking function will take input parameters (D,S,T) and (A,C) and then check against the all defined privacy policy rules. If the input parameters match any rule, it will return the rule setting, either *Allowed* or *Denied*. If the input parameters do not match any rule, it should return *Denied* to avoid potential leakage of personal data through undefined rule.

B. Issues

The design in the previous section is a relaxed version of the requirements in the Data Protection Directives since the Directive specifies personal data should be destroyed after the retention period while the above checking only ensure personal data cannot be used after the retention period. Additional housekeeping tools should be implemented by the data controller to destroy the data after the retention period.

Moreover, the data dissimilation requirement is not considered in the above design as it is difficult to handle. On the other hand, most data leakage incidents are due to uncontrolled data dissimilation. There are two possible approaches that can be used to tackle the data dissimilation problem: adding metadata to control the dissimilation and adding metadata to fingerprint the source.

According to the Directive, the dissimilation of personal data should be done in a controlled manner with proper acknowledgement from the data subject. Moreover, during data collection, the data controller should be aware of the necessity of data dissimilation to other parties and how far away the personal data will be propagated. To address this requirement, the metadata "dissimilation level" is added to each personal data object which controls whether the data can be further dissimilated or not.

In order to support the "dissimilation level" metadata, the privacy reference monitor has to support the personal data dissimilation function. The processing of the dissimilation function is as follow:

1. If the dissimilation level of the personal data is 0, the function should return "*Denied*".
2. If the dissimilation level of the personal data is N (N>0), the function should return the personal data with the dissimilation level set to N – 1.

When there is personal data leakage, the objective of investigation is to find the source of the leakage so as to stop further leakage. To help identifying the data leakage source, the "forensic trace" metadata is added to each personal data, which is a cryptographic token associated with the personal data object and the data controller. When there is data leakage, one can then extract the "forensic trace" of the leaked data object and determine the source of the leakage since each forensic trace is uniquely identified by its data controller. The cryptographic algorithm used to construct the "forensic trace" requires further research and will not be discussed in this paper.

With the metadata "dissimilation level" and "forensic trace", the accountability of the data controller is greatly enhanced. When there is data leakage incident, the leaking path can then be identified.

VI. Conclusion

In the paper, we have proposed a privacy reference monitor that supports the data protection principles. With the above analysis, we have argued that the purpose and binding principles of the Data Protection Directive can be established. Our next step is to implement a prototype privacy reference monitor to demonstrate such concept is practical. One possible implementation is to add the privacy reference monitor to the JVM since Java provides sufficient details to programmer on how to define his own security manager

References

[1] G. Bahadur, W. Chan and C. Weber, Privacy Defended: Protecting yourself oneline, Que, 2002.

[2] F. Xu, J. He, X. Wu and J. Xu, "A privacy-enhanced access control model", 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing, IEEE, 2009, pp. 703-706.

[3] R. Everett-Church, Privacy law and the Internet, in H. Bidgoli (Ed), Global Perspectives in Information Security, Wiley, 2009, pp408-433.

[4] D.S. Herrmann, Complete Guide to Security and Privacy Metrics, Auerbach Publications, 2007.

[5] OECD, Guidelines governing the protection of privacy and transborder flows of personal data, 23 September 1980.

[6] European parliament and Council, 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, November 1995.

[7] C. Wright, W. McAuliffe and A. Gamvros, Internet Law in Hong Kong, Thomson, Hong Kong, 2003.

[8] W3C, Platform for Privacy Preferences (P3P) Project, available at http://www.w3.org/P3P.

[9] P. Ashley, S. Hada, G. Karjoth, C. Powers and M. Schunter, Enterprise Privacy Authorization Language 1.2 (EPAL 1.2), W3C Member Submission, Nov 2003.

[10] Organization for the Advancement of Structured Information Standards (OASIS), Extensible Access Control Markup Language (XACML), Ver. 1.1, July 2003.

[11] D. Gollmann, Computer Security, John Wiley & Sons, New York, 1999.

[12] Q. Ni, D. Lin, E. Bertino and J. Lobo, “Conditional privacy-aware role based access control”, Proc. 12th European Symposium on Research in Computer Security, 2007, 72-89.

[13] S. Al-Fedaghi, “Beyong purpose-based privacy access control”, Proc. 18th Australasian Database Conference (ADC 2007), Ballarat, Australia, 2007.

[14] C.P. Pfleeger and S.L. Pfleeger, Security in Computing (4th Ed), Prentice Hall, 2007.

[15] G.V. Lioudakis, E.A. Koutsoloukas, N.K. Dellas, N. Tselikas, S. Kapellaki, G.N. Prezerakos, D.I. Kaklamani, I.S. Venieris, “A middleware architecture for privacy protection”, Computer Networks (51), Elsevier, 2007, pp. 4679-4696.

[16] J. Pieprzyk, T. Hardjono, and J. Seberry, Fundamentals of Computer Security, Springer, 2003.

[17] E.S. Siougle and V.C. Zorkadis, “A model enabling law compliant privacy protection through selection and evaluation of appropriate security controls”, InfraSec 2002, LNCS 2437, Springer-Version Berlin, 2002.

[18] S. Fischer-Hubner, IT – security and privacy, design and use of privacy-enhancing security mechanisms, LNCS 1958, Springer, 2001.

[19] G. Canfora, E. Costane, I. Pennino, C.A. Visaggio, “A three-layered model to implement data privacy policies”, Computer Standards & Interfaces (30), Elsevier, 2008, pp. 398-409.