Intelligent Detecting Risk based on Privacy in Cloud Computing Environments

Saeromi Yang[1], Yeonwoo Lee[1], Seng-phil Hong[1*] and Sang-Yep Nam[2]

[1] Department *of Computer Science, Sungshin Women's University*
[2] *Department of Electrical Engineering, Kukje University*
{romy0315, yeonwoo57, philhong}@sungshin.ac.kr, synam58@gmail.com

Abstract

Cloud computing has been activated as a means of reducing costs and the effective utilization of IT infrastructure. Despite these changes, Service providers focus almost exclusively on the issues offering service availability without ensuring data privacy. The purpose of this paper is to analyze privacy threats in cloud computing and protect mobile user's personal information against from increasing privacy infringement attacks. So we suggested that privacy risk detection mechanism is the most appropriate means of protecting data privacy in cloud computing. The IDRP(Intelligent Detecting Risk based on Privacy) model is composed of these five mechanisms: 1) Information Verified Mechanism, 2) Path Checked Mechanism, 3) Malicious Coding Detecting Mechanism, 4) Semantic Analyzing Mechanism, and 5) Notice and Alert. With these security functions, the IDRP model will be an essential system to detect privacy-related threats and minimize them in cloud computing.

Keywords: *Privacy Protection, Cloud Computing, Personal Information Security*

1. Introduction

Over the last few years, cloud computing has emerged as one of the fastest growing technology and is changing a large part of the Information Technology industry. Cloud computing attracts the attention of customers who wish to acquire computing infrastructure without large up-front investment, particularly in cases where their demand may be variable and unpredictable [2]. Along with the benefits, however, there are a number of privacy issues associated with cloud computing. Personal information in cloud can be exposed to risks as a result of the ability of cloud computing services to collect and centrally store huge amounts of data. The following graph gives the number of privacy infringements occurred from 2007 to 2011 in Korea.

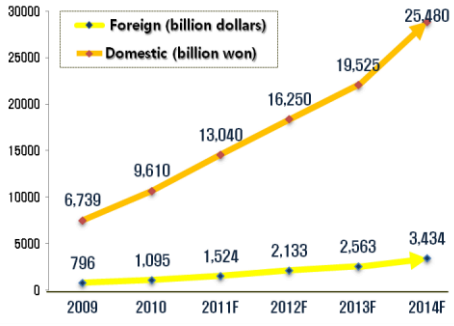

Figure 1. Cloud computing trend [1]

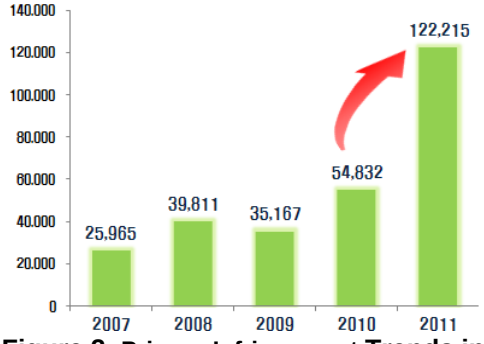

Figure 2. Privacy Infringement Trends in Korea

* Corresponding author : philhong@sungshin.ac.kr

According to KISA(Korea Information Security Agency), the total number of privacy infringement accidents in Korea increased by about 122percent from 54,000 in 2010 to 120,000 in 2011. And with the rise of the smart devices and new services, the number of privacy accidents by neglect such as spread of malicious applications is increasing every year. These accidents didn't gain media attention at first, but it becomes the most important social issue. Therefore, we propose the mechanism to analyze privacy threats and suggest the risk detection model in order to protect mobile user's personal information against from a variety of privacy infringement attacks in cloud computing.

2. Related Works

2.1 Security Threats of Cloud Computing

There are a variety of cloud computing definitions, but the following Gartner's definition is widely accepted. A definition of Cloud computing is a style of computing where scalable and elastic IT-related capabilities are provided 'as a service' to external customers using Internet technologies. And it is also described five defining attributes of cloud computing: service-based, scalable and elastic, shared, metered by use, uses Internet technologies.[4] As cloud computing technology has evolved, mobile services which are applied PC performance and mobile characteristic are also increased with spread of smart devices and new services. However, there are many security threats in cloud computing and these threats can be classified under three categories[5]. (See Table1.)

Table 1. Security Threats of Cloud Computing

Type	Threat
Mobile service	- Misuse and abuse of the mail and messages - Data loss and leakage - Service-based malicious code threats - Security systems threats
Wireless network	- Eavesdropping and wiretapping - Illegal authentication threats - Modulation and leakage of the network information - Denial of Service
Cloud computing service	- Misuse and abuse of the services - Data loss and leakage - Vulnerability in management system - Vulnerabilities in virtualization technology

And the rising demand for the use of personal information through smart devices, privacy-related threats are becoming the new issues of information technology industry. Of the above-mentioned security threats of cloud computing, we focus on protecting personal information by analyzing source code of application requested by user. The source code vulnerabilities of mobile application is input and output validation, API abuse, security features, errors, time and state, code quality and encapsulation[6].

2.2 Privacy Life Cycle in Threats

The development of automatic data processing has made it necessary to consider privacy protection in relation to personal information. There is a personal information life cycle in

order to help IT service providers to collect, store, use and destroy personal information securely and efficiently for their service. The following table shows the types of privacy violations in accordance with the personal information life cycle [7, 8, 9].

Table 2. Privacy Life Cycle related to Threats

Phase	Threats
Collect	- Personal information collecting without proper access authority - Illegal Privacy monitoring against the consent of the subject - Unnecessary personal information collecting to use for commercial
Store	- Unprotected database/system - Privacy leakage to third party - Privacy exposure via SMS, or SNS
Use	- Improper data analysis - Provide commercial or advertising information without the consent of the data subject
Destroy	- Insecure destruction of personal information - Destruction of personal information by data controller's mistake

3. Motivation

As cloud computing service has expanded rapidly, many kinds of mobile applications using the service are beginning to emerge. The unexpected growth, however, raises concerns about privacy and it can cause more complex and serious privacy-related problems than ever before. In this study, we are focusing on privacy-related issues for mobile user in cloud computing environments and here are the major problems [10].

- Data in cloud can be exposed to risks as a result of centrally collecting information and sharing information among un-trusted relationship.
- Privacy violations are increasing due to the undetectable collection, use and destruction of personal information.
- Malicious services are spreading easily because it is difficult to verify the quality of services in cloud computing environments.

In order to solve these problems, we suggest the IDRP(Intelligent Detecting Risk based on Privacy) model in cloud computing environments.

4. Intelligent Detecting Risk based on Privacy in Cloud Computing Environments

In this paper, we propose the IDRP(Intelligent Detecting Risk based on Privacy) model to analyze privacy threats in cloud computing environments and protect mobile user's personal information against from increasing privacy infringement attacks. The IDRP model is composed of 5 functions: 1) Information Verified Mechanism, 2) Path Checked Mechanism, 3) Malicious Coding Detecting, 4) Semantic Analyzing Mechanism, and 5) Notice and Alert Mechanism.

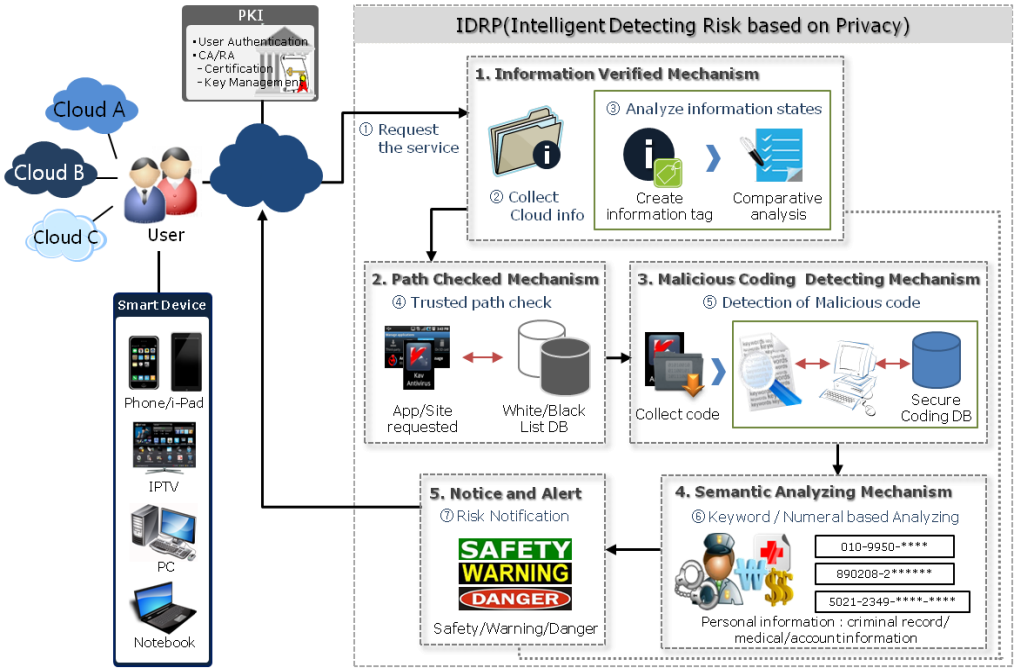

Figure 3. IDRP Architecture

4.1 Information Verified Mechanism

Cloud computing is a distributed processing of programs and data in virtual data center through the internet or the mobile internet. So it is difficult to confirm the ownership of information. We cannot be aware of physical location of services or other things related validation. As the first step in IDRP model, this function collects information and creates information tag that is composed of cloud provider information and data information requested by the user, and then analyzes the information states. In detail, this function checks created information tag by verifying the ownership, type, permission and rating of the data. If it cannot verify the validation of the service or includes unknown an executable file, the IDRP issues an alert and denies downloading an application or access to the information.

4.2 Path Checked Mechanism

When user downloads applications into their mobile device or accesses to websites via mobile web browser, this function confirms whether application is malicious or certificated by checking black list and white list from database.

Table 3. Items of Black list and White list databases

Database	Items
Blacklist	- Phishing sites and malicious applications reported to governmental organization or specialized agency - Services recognized as malicious by the result of IDRP
Whitelist	- Certified services that have certification marks such as e-Privacy mark - Services certified by certificate authority - Services recognized as safe by the result of IDRP

The black list consists of malicious application information and websites information, which can cause privacy violations or phishing, registered in governmental organization or specialized agency. On the other hand, the white list includes services certified by certificate authority. In this paper, certified services mean websites or applications which have safety marks such as e-Privacy mark, e-Trust mark or other certification marks related to privacy. And those databases update automatically when certain services are reported as malicious or certified by certificate authority and the result of IDRC model.

4.3 Malicious Coding Detecting Mechanism

If an application doesn't be registered in both black list and white list, the IDRP starts analyze the source codes of the application or website to detect malicious codes. We detect malicious codes based on the mobile application vulnerabilities which explained in related works. The IDRP issues an alert to user informing a result of analyzing source code. And then updates information of application into the black list when that has insecure codes that can lead to exploitable vulnerabilities.

4.4 Semantic Analyzing Mechanism

When the application and the web site collects data includes the personal information, consent of data's owner and notification should be required. If the service requires identifiable information like resident registration number or sensitive information like account information, this function issues an alert to user indicating that the private data is required. Then Semantic Analyzing Mechanism help user prevent from collecting unapproved data and minimize privacy violations. In this function, we propose two analyzing methods to prevent collecting sensitive or identifiable information.

(1) *Keyword based Analysis*: This method is to examine whether the application requires personally identifiable information and Sensitive information.
- *Personally identifiable information* : any information that could be used to identify or locate an individual(e.g. name, address) or information that can be correlated with other information to identify an individual(e.g. postal code, Internet Protocol(IP) address)
- *Sensitive information*: information on religion or race, health, sexual orientation, union membership or other information that is considered private.

(2) *Numeric Combination based Analysis*: This method detects a series of number that the service requires such as following number combination.
- *Sixteen-digit numbers* : It could require user's credit card number composed of sixteen figures
- *Thirteen-digit numbers*: It could require user's resident registration number composed of thirteen figures.

4.5 Notice and Alert Mechanism

The IDRP makes an alert to inform user the result of inspection. In this function, making an alert can let subject know how the application affect privacy. The alert level can be divided into three level; safety, warning, and danger. When the IDRP detects malicious codes in application, it recognizes that the application is danger, and issues an alert to user notifying that the application could lead to serious security problems. In the same manner, the IDRP issues an alert depending on the result of application inspection.

Table 4. Alert levels

Level	Description
Safety	- Application or website is certified (white list) - Cannot found malicious codes and don't request any personal information
Warning	- Require the low level of information. (e.g. name, id, phone number) - Cannot found malicious codes but requiring critical information - Cannot analyze the code and don't know whether they collect the personal information or not
Danger	- Detect malicious codes - Download unknown an executable file - Require sensitive information or identifiable information. (e.g. resident registration number, account number)

5. Prototype implementation

5.1 Algorithm

Here is the simple algorithm of the IDRP.

Algorithm

```
1:  ReadUserInfo(u_id, u_name)   // information verified mechanism
2:  ReadCspInfo(s_name, c_securitylevel, c_organizationlevel) //csp: cloud service provider
3:  ReadDataInfo(d_owership, d_type,  d_permission, d_rating) ← SearchData(data_requiredbyuser)
4:  CreateTag()
5:  vic_result ← CheckVarifedTag(user_tag, csp_info, data_tag) //vic: verified information check
6:  if (vic_result = danger) then
7:     Notice(Alert_level ← danger, mechanism)
8:  end if
9:  function AlertLevel(result) {   // notice and alert mechanism
10:    if (result = danger) then  UpdateBlackList(access_path, path)
11:    else if (result = safety) then  UpdateWhiteList(access_path,path)
12: end if
13: Notice(Alert_level ←result, mechanism) }
14: switch(access_path) //  path check mechanism
15:    case App:  path ← GetAppPath(app_name, app_csign)  break
16:    case Web:  path ← GetSitePath(site_name, site_url)  break
17: end switch
18: tpc_result ← CheckTrustedPath(trusted_pathDB, path) //tpc: trusted path check
19: AlertLevel(tcp_result)
20: code_info ← ReadCode(access_path, path)   // malicious coding detection mechanism
21: dmc_result ← DetectMaliciousCode(secure_codingDB, code_info) //dmc: detection of malicious code
22: AlertLevel(dmc_result)
23: end if
24: Input_data ← CheckRequiredData()    // semantic analyzing mechanism
25:  if (input_data ≠ null) then
26:     if (format(data_required) = keyword) then
27:        ksa_result ← AnalyzeBasedKeyword() //ksa: keyword semantic analyzing
28:        AlertLevel(ksa_result)
29:     else if (format(data_required) = numeric) then
30:        nsa_result ← AnalyzeBasedNumericCombination() //nsa: numeric semantic analyzing
31:        AlertLevel(nsa _result)
```

```
32:          else (format(data_required) = unknown) then
33:            Alert_level ← warning
34:        end if
35:    end if
```

5.2 Prototype

As a feasible approach, we developed the prototype of the IDRP to apply the real system environments. The application was developed using the following tools: MySQL 5.x - database, JAVA - application development language, Apache tomcat 5.5 - web server. The following screen capture images are the prototype of the IDRP in mobile devices. We used programming tool named Eclipse to make this application and it is based on Android 2.2 in Windows 7.

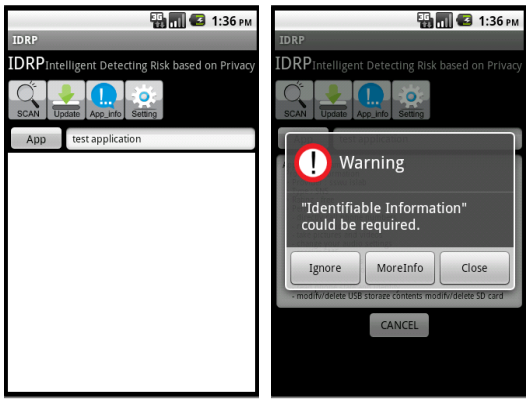

Figure 4. Prototype of the IDRP for app

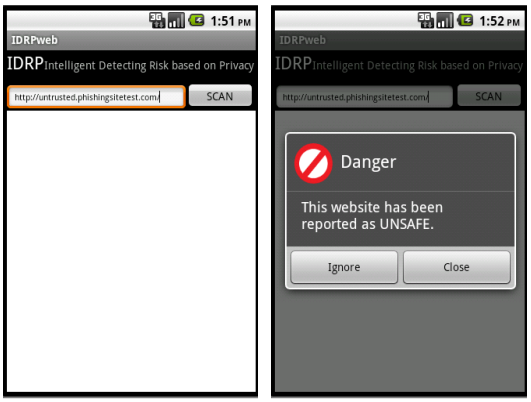

Figure 5. Prototype of the IDRP for web

6. Conclusion and Future Work

The development of Information Technology has made it possible to process a large quantity of personal information automatically. Moreover as cloud computing services have increased, there are a great variety of mobile applications. However, the development of new technology raises new concerns about privacy. In this paper, we suggested the model to manage the personal information threats in cloud computing. The IDRP will be an essential system to protect personal information from increasing privacy-related threats in cloud computing. And we plan to apply a systematic algorithm to the Semantic Analyzing Mechanism and extend the research on risk management based on the result of analyzed data.

Acknowledgements

This work was supported by the Sungshin Women's University Research Grant of 2012.

References

[1] J.-Y. Lee, “Features of Cloud Computing and Status of the service provided by operators”, Trend data report of the Institute of Information and Communication Policy, vol. 22, no. 6, **(2010)**, pp. 1-22.
[2] S. Bradshaw, C. Millard, I. Walden, “Contracts for clouds: comparison and analysis of the Terms and Conditions of cloud computing services”, International Journal of Law and Information Technology, vol. 19, no. 3, **(2010)**, pp. 187-223.

[3] Privacy Infringement Trends in Korea, e-National indicators, http://www.index.go.kr/ **(2011)**.
[4] Sonian, "What is cloud computing?", http://www.sonian.com/about/faq/technology/what-is-cloud-computing/.
[5] E.-Y. Jang, H.-Jo. Kim, C.-S. Park, J.-Y. Kim and J. Lee, "The study on a threat countermeasure of mobile cloud services", Journal of Information Security and Cryptology, vol. 21, no. 6, **(2011)**, pp. 177-186.
[6] Android-JAVA Secure Coding Guide, Ministry of Public Administration and Security, **(2011)**.
[7] Security Management Model for Life Cycle of Personal Information, http://www.tta.or.kr/data/ttas_view.jsp?rn=1&pk_num=TTAS.KO-12.0053.
[8] Study on leakage/exposure prevention and development of privacy framework in mobile environment. Korea Internet & Security Agency, (**2010)**.
[9] OECD Guidelines on the Protection of Privacy and Trans-border Flows of Personal Data, http://www.oecd.org.
[10] S. Pearson, "Taking account of privacy when designing cloud computing services", HP Laboratories, **(2009)**.

Authors

Saeromi Yang

Saeromi Yang received her B.S. degree in Computer Science from Sungshin Women's University, Seoul, Korea, in 2011. Currently she is studying for her M.S degree in Information Security at the same university. Her research interests include cloud computing and privacy.

Yeonwoo Lee

Yeonwoo Lee received her B.S. degree in Computer Science from Sungshin Women's University, Seoul, Korea, in 2011. Currently she is studying for her M.S degree in Information Security at the same university. Her research interests include security architecture, mobile security and privacy protection

Seng-phil Hong

Professor Seng-phil Hong received his BS degree in Computer Science from Indiana State University, and MS degree in Computer Science from Ball State University at Indiana, USA. He researched the information security for PhD at Illinois Institute of Technology from 1994 to 1997, He joined the Research and Development Center in LG-CNS Systems, Inc since 1997, and he received Ph.D. degree in computer science from Information and Communications University in Korea. He is actively involved in teach and research in information security at The Sungshin Women's University, Korea. His research papers appeared in a number of journals such as ACM Computing, Springer-Verlag's Lecture Notes in Computer Science, etc. His research interests include access control, security architecture, Privacy, and e-business security.