

Risk of Privacy Loss Assessment of Cloud Storage Services

Chatchawan Kamwan, Twittie Senivongse

Computer Science Program, Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University
254 Phayathai Road, Pathumwan, Bangkok, 10330, Thailand
chatchawan.k@student.chula.ac.th, twittie.s@chula.ac.th

Abstract— Cloud storage services have promising characteristics for personal and corporate use as consumers' data can be stored on shared pools of storage hosted by cloud providers, while consumers can access and manage their data anywhere via an Internet connection. Nevertheless, protection of data privacy is one of the major issues, and at the time of storage service selection, prospective consumers are concerned with how the cloud storage providers handle their personal data. To help a consumer with storage service selection, this paper proposes a methodology to assess the risk of privacy loss when consumer data are stored with a particular cloud storage service. The risk of privacy loss is viewed from two aspects. First, sensitivity of the personal data to be stored in the cloud and sensitivity of consumers' personal data requested at the time of service registration can contribute to the risk. Second, lack of privacy control transparency can be a risk factor if the cloud storage provider inadequately show the necessary privacy principles that are practiced. The proposed methodology can assist the consumers when determining the risk levels of privacy loss of different cloud storage services. We present the application of the methodology to a case of an organization selecting a cloud storage service to host its corporate data.

Keywords— privacy; sensitivity; personal data; cloud storage; assessment

I. INTRODUCTION

A cloud storage service provided by a service-hosting company maintains and manages its customers' data that are stored on large shared pools of storage [1]. The service is available with minimal procurement and configuration effort, offers elastic capacity, and bills the consumers only for the storage they use. Nevertheless, without knowledge of the physical location of the stored data and of how they are processed, prospective consumers may feel threatened by the potential loss of privacy and control over their data. As different kinds of data may require different treatment from storage providers, while different storage providers may handle stored consumers' data differently, a prospective consumer needs to determine which storage provider is right for its use.

A person, either an individual or a corporate body, may have several kinds of data put in the cloud. Those data are personal data when they are "produced by any services (not only software and Web-based services, but also human and

device-based services) around a person during the interactions between the person and the services, on condition that if the person's identification is removed from the data, the data will become meaningless." [2]. Personal data can be classified from different dimensions [2]. For example, on the format dimension, personal data are in the format of documents, multimedia, Web pages and fragments, emails, and business data stored in domain-specific services or databases. On the semantics and functions dimension, personal data include preference data, communication records, Web footprints, personal profiles, transaction records of service consumptions, and public service records.

A prospective consumer should be able to determine the risks associated with each kind of personal data that will be stored and processed in the cloud. Those risks depend on sensitivity of the personal data. Data sensitivity is concerned with the adverse effect to individual(s) or organization(s) as a result of the loss of confidentiality, integrity, and availability of the data [3]. Data sensitivity is generally classified into three levels if the data are lost or compromised [3][4]: 1) high sensitivity if a severe or catastrophic adverse effect will result, 2) moderate sensitivity if a serious but not severe adverse effect will result, and 3) low sensitivity if a limited or minor adverse effect will result. However, the risks are associated not only with the personal data but also with the way the cloud provider handles privacy of the data. Privacy [5] refers to "the right to self-determination, that is, the right of individuals to 'know what is known about them', be aware of stored information about them, control how that information is communicated, and prevent its abuse." The person has the right to know and control how the cloud storage provider protects personal data. For a prospective consumer, this raises questions of what privacy principles are followed by the provider and how transparent the provider is with regard to disclosure of privacy protection information.

This paper proposes a methodology for a prospective consumer, especially an organization, to determine the risk of privacy loss if the organization moves data to a cloud storage service. Organizational data can be considered as personal data as they can identify the organization as well as individuals and other organizations that are related one way or another to the organization, e.g. employees, customers, business partners. The risk of privacy loss is then viewed from

two aspects. First, sensitivity of the personal data to be stored and processed in the cloud will be considered. Here the personal data refer to the data that the organization plans to move to the cloud together with the data about the organization which are requested at the time of service registration. Second, lack of privacy control transparency can also be a risk factor if the cloud storage provider does not follow privacy protection guidelines, or does follow the guidelines but does not publish sufficiently the information about its privacy protection practices. The proposed methodology introduces a risk of privacy loss assessment model that relies on data sensitivity assessment and privacy control transparency assessment. We will assess sensitivity of personal data based on the impact of different data categories on the organization as proposed by NIST 800-60 [6] together with the approach by Jang and Yoo [7]. To assess privacy control transparency, we mainly follow the guidelines for privacy controls as outlined in CSA privacy level agreement [8], and also consider additional information from the Australian privacy principles [9], and privacy controls in NIST 800-53 [10]. An organization can then assess the risk associated with its personal data if they are stored and processed by a particular cloud storage service. The assumption is that, if highly sensitive data will be stored with a cloud storage service that is not very transparent, or does not provide sufficient information, about its compliance with the privacy control guidelines, the organization should note that there could be a high risk of privacy loss.

Section II of the paper discusses related work. Section III presents sensitivity assessment of personal data and Section IV assesses privacy control transparency of cloud storage providers. Then Section V proposes the overall model to assess the risk of privacy loss. The application of the methodology to a case of an advertising company in Thailand will be shown as an example. The paper discusses an experiment and concludes in Section VI.

II. RELATED WORK

We focus here only on the related research in quantitative assessment of data sensitivity and in data privacy guidelines.

On data sensitivity, Jang and Yoo [7] classify sensitivity of a personal data attribute based on the degree of relatedness it has with other data attributes, whether it is an identity or a private attribute, and whether it can derive other attributes. Chaiwongsa and Senivongse [11] apply this approach to measure sensitivity of data exchanged with Web services in order to select a service that exchanges less sensitive data.

On data privacy guidelines, Pearson [12] outlines privacy requirements for cloud providers such as notify the users of what data to collect, how they will be processed, and put security safeguards and audit functions in place. Tancock et al. [13] present a conceptual design of a privacy impact assessment tool that has a knowledge base of rules for privacy and data protection legislations, and can highlight privacy risks and compliance issues that the providers or organizations are to be taken into account. Pearson and Benameur [14] suggest an organization considering using cloud services to

classify its data assets to identify which ones are confidential, and make sure that the agreement with the selected provider explicitly states the requirement to comply with applicable data protection and privacy laws. Lar et al. [15] propose security-privacy mechanisms such as manage digital identity including the use of pseudonyms to preserve privacy via anonymity, and provide data-centric security approach that allows the data owner to have full control over who has the right to use the data and what they are allowed to do with it once they gain access. Online Trust Alliance [16] proposes a list of questions for cloud providers or organizations to self-assess security and privacy risks of their infrastructure and privacy practices. Another list of questions is also provided for the organizations to assess the risks of outsourcing to cloud providers. These privacy controls and mechanisms by related researches are taken into account in our assessment of privacy control transparency. Unlike them, as compared in Table 1, our work focuses on quantitative assessment and tackles both aspects of sensitivity of stored data and provider conduct of privacy controls.

TABLE 1. COMPARISON OF APPROACHES TO DATA PRIVACY IN CLOUD

Approach	Privacy Controls	Privacy Controls-Based Assessment	Sensitivity Assessment	Quantitative Assessment
[12], [14], [15]	x			
[13], [16]	x	x		
This work	x	x	x	x

III. DATA SENSITIVITY ASSESSMENT

The first step of the methodology for an organization, considering using a cloud storage service, is the assessment of sensitivity of its personal data. Here personal data refer to organizational data that will be stored and processed in the cloud. In addition, the organization will have to supply certain personal profile data when registering to use a cloud storage service.

A. Sensitivity of Data to Be Stored in the Cloud

Since sensitivity is concerned with the adverse effect to the organization as a result of the loss of confidentiality, integrity, and availability of the data [3], we adopt the impact of different information types as defined by NIST SP 800-60 [6] to determine sensitivity of organizational data that will be stored in the cloud. NIST develops standards for US Federal agencies to categorize information types that are related to 1) specific missions of the agencies, and 2) administrative and business services in general. The information is categorized into information groups and each group has associated information types. This paper focuses on the information types that are related to administrative and business services in general. There are 77 of them in total with some example shown in Table 2. For each information type, NIST defines the default level of impact on the organization if such an

information type loses its confidentiality, integrity, and availability. The impact level is either high (H), moderate (M), low (L), or N/A (i.e. the information type does not require its own impact assessment). Note that NIST allows reassignment of the impact levels of some information types (shown in *italic*) if special factors occur, by which the impact levels can be increased [6]. For example, the information type Accounting can have its loss-of-confidentiality impact changed from the default low to high if unauthorized disclosure can pose a threat to human life or a loss of major assets. It can also be assigned a moderate impact if the loss of its confidentiality violates non-disclosure agreements or if it is involved in fraud or criminal investigations. Based on these impact levels and the definition of sensitivity [3], we derive the sensitivity level of an information type using the high water mark, i.e. the highest level of all impact levels assigned to that information type. Table 2 also gives some example of the sensitivity levels that we derive for different information types.

TABLE 2. EXAMPLE OF INFORMATION TYPES AND SENSITIVITY LEVELS DERIVED FROM NIST IMPACT LEVELS

Information Type	Loss-of-Confidentiality Impact	Loss-of-Integrity Impact	Loss-of-Availability Impact	Sensitivity
Financial Management				
Asset and Liability Mgmt	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Reporting and Information	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Funds Control	<i>M</i>	<i>M</i>	<i>L</i>	<i>M</i>
Accounting	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Payments	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Collections and Receivables	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Cost Accounting/ Performance Measurement	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Human Resource Management				
Staff Acquisition	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Organization & Position Mgmt	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Employee Performance Mgmt	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Employee Relations	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Labor Relations	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Human Resources Development	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
...	...	...	...	...
Information and Technology Management				
System Maintenance	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
Record Retention	<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
Information Management	<i>L</i>	<i>M</i>	<i>L</i>	<i>M</i>
System and Network Monitoring	<i>M</i>	<i>M</i>	<i>L</i>	<i>M</i>
...	...	...	...	...

Remark: Impact levels shown in *italic* can be increased if special factors occur.

The organization has to specify which organizational data, i.e. which information types, to put in the cloud. Based on the sensitivity level of each information type to be put in the cloud

and, again, on the high water mark, the overall sensitivity level S_{Stored} of all organizational data that will be stored together in the cloud is represented by the highest sensitivity level of all stored information types and is quantified by

$$S_{Stored} = \begin{cases} 1 & \text{if highest sensitivity level of stored information types is H} \\ 0.67 & \text{if highest sensitivity level of stored information types is M} \\ 0.33 & \text{if highest sensitivity level of stored information types is L} \\ 0 & \text{if highest sensitivity level of stored information types is N / A} \end{cases} \quad (1)$$

For example, the Financial Department (FN) of an advertising company plans to use a cloud storage service to store all seven information types in the Financial Management information group, as in Table 2, in the cloud. So the highest sensitivity level of the seven information types to be stored is M and, by (1), $S_{StoredFN} = 0.67$.

B. Sensitivity of Data Requested by Cloud Provider

To register for a cloud storage service, a cloud consumer will be requested to provide some personal profile data. To assess their sensitivity, we propose an approach adapted from Jang and Yoo [7]. First, a cross table that consists of different concepts, covering all information that is the subject of sensitivity assessment, is defined. Each concept is associated with a number of attributes that belong to that concept. Table 3 gives an example of a cross table of personal data requested by a cloud provider P1. Four concepts are defined and each requested data attribute is assigned (with an x mark) to the concept(s). The attributes are basic data, private data (if they are considered private and should not be disclosed), identity data (if they can identify a data subject or a person), and finance data. Each organization may define a cross table differently, i.e. concepts and assignment of attributes to concepts may be different. For example, the private concept may be subjective to each consumer, as one organization may see Mobile Phone as private information whereas it may be not for others. In addition, if the organization considers using storage services from different cloud providers, it has to define a specific cross table for each provider since different providers may request different sets of personal profile data.

TABLE 3. EXAMPLE OF CROSS TABLE FOR PERSONAL PROFILE DATA REQUESTED BY PROVIDER P1

Concept/Attribute	First Name	Last Name	Company Name	Email	Work Phone	Mobile Phone	Credit Card Number	Security Code	Expiration Date
Basic	x	x	x	x	x	x			
Private						x	x	x	
Identity		x	x				x		
Finance	x	x	x		x	x	x	x	x

Given the cross table of data attributes requested by a provider, sensitivity of a data attribute submitted to that provider is calculated by the following equations.

1) Degree of Conjunction: Degree of conjunction of an attribute d is derived from the number of concepts which the attribute d describes. This means the knowledge of d may give a chance to link to other requested attributes in the same concept(s) as d . The degree of conjunction of d or $D_C(d)$ is determined by

$$D_C(d) = \frac{\text{number of concepts to which } d \text{ belongs}}{\text{total number of concepts}}. \quad (2)$$

For example, the attribute First Name is a member of 2 out of 4 concepts, i.e. Basic and Finance. Therefore $D_C(\text{First Name}) = 2/4 = 0.5$.

2) Principle of Identity: Principle of identity of an attribute d indicates that d is an identity attribute, i.e. d can be used as a key to access other requested attributes within the same concept(s) as d . Disclosure of d will lead to disclosure of other attributes. The principle of identity of d or $I_A(d)$ is determined by

$$I_A(d) = \begin{cases} 0 & \text{if } d \text{ is not identity attribute} \\ \frac{\text{number of attributes in the concepts}}{\text{total number of attributes}} & \text{if } d \text{ is identity attribute for the concepts.} \end{cases} \quad (3)$$

For example, the attribute Company Name is an identity attribute (i.e. it belongs to the concept *Identity*) for the concepts Basic and Finance. There are 9 attributes associated with Basic and Finance and also there are 9 attributes in total. Therefore $I_A(\text{Company Name}) = 9/9 = 1$. For the attribute Work Phone, it is not an identity and $I_A(\text{Work Phone}) = 0$.

3) Principle of Privacy: Principle of privacy of an attribute d indicates that d is private information. The principle of privacy of d or $P_A(d)$ is determined by

$$P_A(d) = \begin{cases} 0 & \text{if } d \text{ does not belong to the concept Private} \\ 1 & \text{if } d \text{ belongs to the concept Private.} \end{cases} \quad (4)$$

For example, the attribute Credit Card Number is a private attribute and $P_A(\text{Credit Card Number}) = 1$, whereas the attribute Work Phone is not and $P_A(\text{Work Phone}) = 0$.

4) Value of Analogism: Value of analogism of an attribute d indicates whether the knowledge of d can derive other requested attributes. The value of analogism of d or $A_A(d)$ is determined by

$$A_A(d) = \begin{cases} 0 & \text{if } d \text{ cannot derive other attributes} \\ 1 & \text{if } d \text{ can derive other attributes.} \end{cases} \quad (5)$$

For example, if the attribute Email can derive the attribute Last Name or Company Name, $A_A(\text{Email}) = 1$. For the attribute Expiration Date, it cannot derive other attributes and $A_A(\text{Expiration Date}) = 0$.

Using (2)-(5), we can compute the sensitivity level of an attribute d by

$$s(d) = D_C(d) + I_A(d) + P_A(d) + A_A(d). \quad (6)$$

Then, the sensitivity level of the data requested by the cloud provider is computed by

$$S_{\text{Requested}} = \frac{\sum_{i=1}^k s(d_i)}{4k} \quad (7)$$

where $S_{\text{Requested}}$ = sensitivity level of data requested by cloud provider; $S_{\text{Requested}}$ is in (0, 1],

$s(d_i)$ = sensitivity level of data attribute i requested by cloud provider, and

k = number of data attributes requested by cloud provider.

For example, the FN department plans to use cloud storage of a provider P1. Using (6)-(7), it calculates the sensitivity of the personal data requested by P1 as in Table 4.

TABLE 4. SENSITIVITY OF FN DATA REQUESTED BY PROVIDER P1

Attribute d	$D_C(d)$	$I_A(d)$	$P_A(d)$	$A_A(d)$	$s(d)$
First Name	2/4	0	0	0	0.5
Last Name	3/4	9/9	0	0	1.75
Company Name	3/4	9/9	0	0	1.75
Email	1/4	0	0	1	1.25
Work Phone	2/4	0	0	0	0.5
Mobile Phone	3/4	0	1	0	1.75
Credit Card Number	3/4	8/9	1	0	2.64
Security Code	2/4	0	1	0	1.50
Expiration Date	1/4	0	0	0	0.25
Total					11.89
$S_{\text{RequestedFN-P1}}$					0.33

C. Sensitivity of Consumer Data

Based on S_{Stored} from (1) and $S_{\text{Requested}}$ from (7), the sensitivity level of all the consumer's personal data is computed by

$$S_{\text{Data}} = (W_{\text{Stored}} * S_{\text{Stored}}) + (W_{\text{Requested}} * S_{\text{Requested}}) \quad (8)$$

where S_{Data} = sensitivity level of consumer data; S_{Data} is in (0, 1],

W_{Stored} = Weight given to data to be stored in the cloud,

$W_{\text{Requested}}$ = Weight given to data requested by cloud provider, and

$$W_{\text{Stored}} + W_{\text{Requested}} = 1.$$

The Financial Department (FN) may, for example, assign $W_{\text{Stored}} = 0.8$ and $W_{\text{Requested}} = 0.2$ as it sees that the data to be stored in the cloud relatively are more important than the data requested by the provider. The sensitivity level of all FN personal data when stored with the provider P1 $S_{\text{DataFN-P1}}$ is $(0.8 * 0.67) + (0.2 * 0.33) = 0.6$.

IV. PRIVACY CONTROL TRANSPARENCY ASSESSMENT

In addition to the degree of sensitivity of data in the cloud, privacy controls that a provider practices contribute to the risk of privacy loss. We present a Cloud Privacy Control Questionnaire (CPCQ) that a prospective consumer may use for checking the provider's compliance with the privacy guidelines and best practices. The CPCQ is derived mainly from the CSA privacy level agreement [8], with some additional information from the Australian privacy principles [9], and privacy controls in NIST 800-53 [10]. The privacy guidelines suggested by the related work in Section II address similar issues to these three privacy principles. There are 42 questions grouped under 18 controls. Due to space limitation, Table 5 shows an example of the questions under each control.

TABLE 5. PART OF CLOUD PRIVACY CONTROL QUESTIONNAIRE

Code	Cloud Privacy Control	Cloud Privacy Control Question
CPC1	Identity of the CSP (and of Representative, as applicable), its role, and the contact information for the data protection officer and information security officer	1. Does the provider specify its name, its representative, and address? 2. Does the provider specify its role in processing of consumer data (i.e. controller, joint-controller, processor or subprocessor)?
CPC2	Categories of personal data that the customer is prohibited from sending to or processing in the cloud	1. Does the provider clearly specify which data categories are uploadable or prohibited from uploading to cloud?
CPC3	Ways in which the data will be processed	1. Does the provider specify locations of data centers where consumer data may be processed (e.g. collected, stored, used, disseminated, changed, and erased)?
CPC4	Data transfer	1. Does the provider indicate whether consumer data will be transferred, backed up, and recovered across borders in regular operations or emergency?
CPC5	Data security measures	1. Does the provider describe the processes and measures to ensure availability (e.g. backup network links, redundant storage, data backup and restore)?
CPC6	Monitoring	1. Does the provider specify whether the consumer can monitor or audit, on an ongoing basis, to see if privacy and security measures are met, and if so, how (e.g. reporting, audit)?
CPC7	Third-party audits	1. Does the provider specify whether and what third-party audit reports will be provided to the consumer?
CPC8	Personal data breach notification	1. Does the provider specify whether and how the consumer will be informed of security breach (e.g. accidental or unlawful destruction, loss, alteration, unauthorized access) in consumer data?
CPC9	Data portability, migration, and transfer-back assistance	1. Does the provider specify the supported formats, preservation of logical relations, and costs for porting consumer data?
CPC10	Data retention, restitution, and deletion	1. Does the provider indicate for how long consumer data will be retained?

CPC11	Accountability	1. Does the provider describe policies and procedures that demonstrate accountability (e.g. documenting all processing operations, provider monitoring and logging)?
CPC12	Cooperation	1. Does the provider specify how it will cooperate with the consumer to ensure data protection compliance?
CPC13	Law enforcement access	1. Does the provider describe the process to respond to requests for disclosure of consumer data by law enforcement authorities?
CPC14	Remedies	1. Does the provider indicate what remedies are available to the consumer on contractual obligation breaches by itself or by subcontractor (e.g. compensation, service credits, financial penalties, ability to sue)?
CPC15	Complaint and dispute resolution	1. Does the provider provide contract detail of the person who will receive questions and complaints?
CPC16	Cloud service provider insurance policy	1. Does the provider describe its cyber-insurance policy regarding data protection and security breaches?
CPC17	Intended use of service	1. Does the provider indicate that consumer data will not be processed in a way that is not intended by the use of service, e.g. no transfer of data for marketing purpose?
CPC18	Dissemination of Privacy Program Information	1. Does the provider make its privacy program information available and easily accessible?

To answer the CPCQ, the prospective consumer may look at available information (e.g. Web pages, privacy policy, terms of use, published documents) of a storage provider with regard to its data privacy protection. The assessment, as a result, is preliminary and is a measure of how transparent the provider is in terms of privacy protection, since it relies on the information available to the prospective consumer and public. A score of 1 is given to Yes answer and 0 to No. If some evidence (e.g. report, certificate) is present to support the control, the provider can get one extra score for the question. The privacy control transparency level of a provider is calculated by

$$T_{Provider} = \frac{\sum_{i=1}^n (C_i + e_i)}{2n} \quad (9)$$

where $T_{Provider}$ = privacy control transparency level of provider; $T_{Provider}$ is in $[0,1]$,

C_i = compliance score for privacy control question i ;
1 = yes and 0 = no,

e_i = evidence score for privacy control question i ;
1 = evidence present and 0 = no evidence or not compliant,

n = number of privacy control questions; $n = 42$.

Using (9), the Financial Department (FN) gets the privacy control transparency level of the provider P1 $T_{ProviderP1} = 0.81$.

V. RISK OF PRIVACY LOSS ASSESSMENT

Based on S_{Data} from (8) and $T_{Provider}$ from (9), the risk level of privacy loss for a consumer using a cloud storage service is calculated by

$$R_{Privacy} = (1 - T_{Provider}) * S_{Data} \quad (10)$$

where $R_{Privacy}$ = risk level of privacy loss when storing data with cloud provider; $R_{Privacy}$ is in [0, 1].

For example, using (10), the Financial Department (FN) gets the risk level of privacy loss when storing data with the provider P1 $R_{PrivacyFN-P1} = (1-0.81)*0.6 = 0.11$.

VI. DISCUSSION ON EXPERIMENT AND CONCLUSION

Apart from the Financial Department (FN) planning to store their data with the provider P1, the advertising company plans to move the data of the Human Resource (HR) and Information Technology (IT) departments to the cloud as well. The HR and IT information types that are to be stored are listed in Table 2. In addition, other two providers P2 and P3 are considered. (We leave out the provider names here.) In the experiment, we select these three providers by first picking ten providers from the CSA STAR repository [17] at random, studying their self-assessments of the security controls that they have in place, and finally ranking them. The providers P1, P2, P3 come 1st, 5th, and 10th in the security ranking. Then we use the proposed methodology to assess the risk of privacy loss for the three departments if the services of the three providers are used. The risks in Figure 1 show a similar pattern. That is, for each department, the risk is lowest when using the service of the provider P1 and highest in the case of the provider P3. This happens to coincide with the security ranking that we observe from the providers' self-assessments in the CSA STAR repository. In addition, different information types that different departments store with a particular cloud service may be subject to different risk levels.

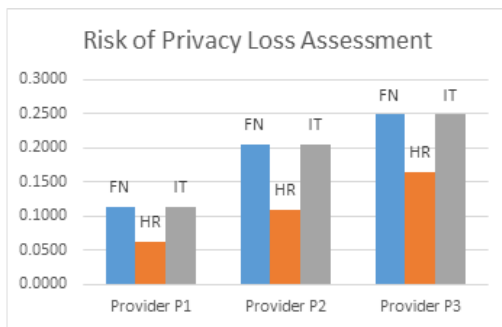


Figure 1. Comparison of risk levels for an advertising company

It is important to note that the assessed risk level should not be considered as an absolute risk score of a provider since our methodology takes into account only data sensitivity and providers' published information. Unfortunately, for a prospective consumer, only information that is open to public

can be accessed, while there might be other internal risks that are not exposed. Also, the risk of privacy loss is not the only factor for a prospective consumer to select a provider. Other quality factors, such as cost, performance, security, usability, and customer support, are generally considered, and so the assessed risk level should rather be considered as one factor to support provider selection. To make the proposed methodology more automated, a tool to help with data sensitivity assessment and with the analysis of the published privacy information for assessing privacy control transparency is under development.

REFERENCES

- [1] N. Beagrie, A. Charlesworth, and P. Miller, "Guidance on cloud storage and digital preservation," Charles Beagrie Ltd., April 2014.
- [2] J. Wang and Z. Wang, "A survey on personal data cloud," *The Scientific World J.*, vol. 2014, 13 p., 2014.
- [3] *Identify Data Sensitivity*, DataONE. Available: <https://www.dataone.org/best-practices/identify-data-sensitivity>
- [4] F. Simorjay, "Data classification for cloud readiness," Microsoft, 2014. Available: <http://download.microsoft.com/download/0/A/3/0A3BE969-85C5-4DD2-83B6-366AA71D1FE3/Data-Classification-for-Cloud-Readiness.pdf>
- [5] ITU, "Privacy in cloud computing," ITU-T Technology Watch Report, March 2012. Available: <http://www.itu.int/en/ITU-T/techwatch/Pages/cloud-computing-privacy.aspx>
- [6] *Information Security*, NIST Special Publication 800-60 Volume II, August 2008.
- [7] I. Jang and H.S. Yoo, "Personal information classification for privacy negotiation," *Proc. 4th Int. Conf. Computer Sciences and Convergence Information Technology*, 2009, pp. 1117-1122.
- [8] Privacy Level Agreement Working Group, "Privacy level agreement outline for the sale of cloud services in the European union," Cloud Security Alliance, February 2013. Available: <https://cloudsecurityalliance.org/download/privacy-level-agreement-pla-outline-for-the-sale-of-cloud-service-providers-providing-services-in-the-european-union>
- [9] "Privacy fact sheet 17 Australian privacy principles," Australian Government, January 2014. Available: <http://www.oaic.gov.au/privacy/privacy-resources/privacy-fact-sheets/other/privacy-fact-sheet-17-australian-privacy-principles>
- [10] *Security and Privacy Controls for Federal Information Systems and Organizations*, NIST Special Publication 800-53, April 2013.
- [11] P. Chaivongsa and T. Senivongse, "Web services privacy measurement based on privacy policy and sensitivity level of personal information," *Proc. 8th Int. Conf. Computing and Information Technology (IC2IT 2012)*, May 2012, pp. 77-85.
- [12] S. Pearson, "Taking account of privacy when designing cloud computing services," *Proc. ICSE Workshop Software Engineering Challenges of Cloud Computing*, September 2009, pp. 44-52.
- [13] D. Tancock, S. Pearson, and A. Charlesworth, "A privacy impact assessment tool for cloud computing," *Proc. 2nd IEEE Int. Conf. Cloud Computing Technology and Science*, November 2010, pp. 667-676.
- [14] S. Pearson and A. Benameur, "Privacy, security and trust issues arising from cloud computing," *Proc. 2nd IEEE Int. Conf. Cloud Computing Technology and Science*, November 2010, pp. 693-702.
- [15] S.-U. Lar, X. Liao, and S.A. Abbas, "Cloud computing privacy & security: Global issues, challenges, & mechanisms," *Proc. 6th Int. ICST Conf. Communications and Networking in China (CHINACOM)*, August 2011, pp. 1240-1245.
- [16] "Security & privacy risk assessment," Online Trust Alliance, 21 January 2015. Available: <https://otalliance.org/resources/security-privacy-risk-assessment>
- [17] *CSA Security, Trust & Assurance Registry (STAR)*, CSA. Available: <https://cloudsecurityalliance.org/star/>



C.Kamwan was born in Khon Kaen, Thailand, in 1981. He received the B.E. degree in Information Technology and Computer form North-Eastern University, Khon Khaen Thailand, in 2008, and join M. Computer Science from Chulalongkorn University, Bangkok, Thailand, in 2013. In 2013, he join computer science of Chulalongkorn University as a student.

In 2013, he join computer science of Chulalongkorn University as a student. His interest research in evaluate, assessment of privacy on cloud, evaluate, assessment of security on cloud.

He submitted research “Risk of Privacy Loss Assessment of Cloud Storage Services” in International Conference on Advanced Communications Technology (ICACT2016) on Jan, 31 – Feb, 03 2016.