PccP: A Model for Preserving Cloud Computing Privacy

Syed Mujib Rahaman
Associate Professor
Dept. of Computer Science Engineering
Dr.L.Bullayya College of Engineering for Women
Visakhapatnam, India.
rahaman.research@gmail.com

Mohammad Farhatullah
Assistant Professor
Dept. of Computer Science Engineering
Al-Ameer College of Engineering & I.T.
Visakhapatnam, India.
calmcview@gmail.com

***Abstract*— The widespread focus on the Cloud Computing has necessitated the corresponding mechanisms to ensure privacy and security. Various attempts have been made in the past to safeguard the privacy of the individual or agency trying to utilize the services being provided by the cloud. The most challenging task is to provide services to the users while also preserving the privacy of the user's information. In this paper a model that incorporates a three-level architecture, Preserving cloud computing Privacy (PccP) model is proposed which aims to preserve privacy of information pertaining to cloud users. The Consumer Layer deals with all the aspects which relate to enabling the user of the cloud to access the cloud services being provided by the cloud service provider. The Network Interface Layer creates an appropriate mapping between the original IP addresses of the users with a modified IP address, and thereby ensuring the privacy of the IP address of the users. The Privacy Preserved Layer utilizes the functionality of the Unique User Cloud Identity Generator for which an algorithm is proposed in this paper to generate an unique User Service Dependent Identity(USID) with privacy check by establishing mapping among the existing user identity(ID),if any to ID's available in a pool of User ID's to enhance the privacy of sensitive user information. A Privacy check method based on information privacy is being proposed which contributes significantly in maintaining user control over the generated user identities (USID's).**

***Keywords*—Cloud Computing, Privacy, Identity.**

I. INTRODUCTION

Cloud Computing is the collection of software, hardware and network resources which are capable of providing the required services to the users at an appropriate cost. The user of the cloud services need not have any knowledge of the nature and characteristics of the resources available. The responsibility of creating, maintaining and managing the resources is with the cloud service provider. Various types of services are currently being provided on the cloud which includes Software as a Service (SaaS), Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Component as a Service (CaaS) etc.

Cloud Computing aims to provide improved performance, reduced hardware equipment for end users, instant software updates, accessibility, better collaboration with peers, *Pay for what you Use* with the added benefits of flexibility and ease of maintenance. This will result in achieving better economies of scale as compared to the conventional models of computing.

In the achievement of the above objectives Cloud Service Providers are constrained by the creation, maintenance and operation of too many servers at multiple locations, time to transition among them based on certain legal and trust issues. However, the major issues of concern to Cloud Computing are the Security and Privacy issues.

If the responsibility of providing these services to the user is one task, the other important task is to ensure that these services are being provided while also ensuring the important features of privacy and security. Organizations and governments which try to implement their services using a cloud will have to implement it while considering the cloud as a social infrastructure [1].

As users try to interact with the cloud for gaining access to these cloud services being provided they will also have to provide personal information which is sensitive. Preserving the privacy of this information while also ensuring that this information is not misused is one of the main objectives of the current model being proposed

II. CLOUD COMPUTING PRIVACY ISSUES

A. From the perspective of Individuals connecting to the cloud:

Maximizing individual user control, creating anonymous services for individual users, creating facilities for the use of multiple identities and limiting identity information and authentication to high level transactions are the privacy issues which have to be guaranteed for an individual to feel that the privacy of information submitted to the cloud is ensured.

B. From the perspective of Cloud computing service providers:

Providing facilities for maintaining anonymity of personal information, encryption of data if it contains personal information, compartmentalizing data processing and storage,

978-1-4673-2149-5/12/$31.00 ©2012 IEEE

controlling unique identifiers, managing explicitly the privacy and security requirements between the cloud service providers are the major privacy issues from the point of view of the cloud service provider.

Furthermore, supporting the development of privacy enhancing technologies, using the Privacy Impact Assessment and coordinating privacy enforcement and compliance across jurisdictions are the other objectives of cloud computing privacy.

III. PROPOSED "PRESERVING CLOUD COMPUTING PRIVACY MODEL"(PCCP)

The proposed Preserving cloud computing Privacy(PccP) model has a three - layered architecture as given below:

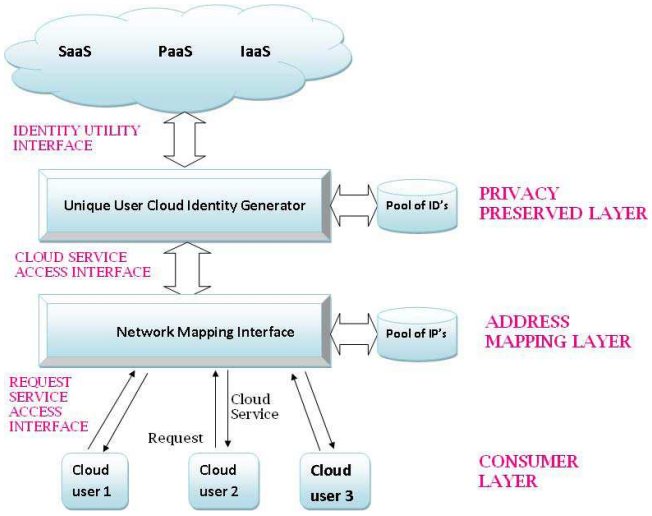

Figure 1. Preserving cloud computing Privacy (PccP) Model

The layers include the *Consumer layer*, the *Address mapping layer* and the *Privacy Preserving layer*. Any request for service by the cloud user will have to processed through these three layers and then accordingly cloud user request is serviced.

A. Consumer Layer: The user of the cloud service who may be an individual or an organization initially places a request for service to the address mapping layer.

B. Address Mapping Layer: The request for service arrives at the Request Service Access Interface(R-sa-I). The R-sa-I then interacts with the Network Mapping Interface (N-map-I) available at the Address mapping layer.

The N-map-I maintains a Pool of Internet Protocol (IP) Addresses. The N-map-I then creates an appropriate mapping between the original IP address of the consumer or user and a modified or Owned Translated IP address (OTIP). In this process the privacy of the original IP address of the user is being ensured. Once the address mapping has been done the Address Mapping Layer then forwards the request for service to the Privacy Preserved Layer. The request for service arrives at the Privacy Preserved Layer using the OTIP.

*C. Privacy Preserved Layer:*The request for service from the Addressing Mapping Layer arrives at the Cloud Service Access Interface(C-sa-I) of the Privacy Preserved Layer along with the OTIP address. The Privacy Preserved Layer has a Unique User Cloud Identity Generator with privacy check to generate a unique user identity which ensures that the privacy of the sensitive user information is guaranteed.

All communication from / to the Cloud at the Identity Utility Interface(Id-U-I) occurs using the OTIP address. Where as in general accessing will be done within the cloud with an unique identity generated by means of Unique User Cloud Identity Generator and cloud service may be provided through the N-map-I using the original IP addresses.

It is the functionality of the Unique User Cloud Identity Generator to generate a Unique Service-dependent identity (USID) by adopting the following measures.

(i) It should be unique to all the users using a particular type of service

(ii) The life span or scope of the identity is till the time the user is using that service, i.e., once the user has completed using the service the identity should be destroyed.

(iii)The user may be using different cloud services. In such a case the user will have to be allocated multiple identities based on the type of service. A list of all the different identities will have to be maintained and updated from time to time. The update may either result in discarding the identity in case the user has finished utilizing the service or it may also result in an extension in case the user wants to utilize the service for an extended period of time.

(iv)The Unique user identity that is being generated must be a function of the Modified IP address and time stamp.

The Unique Service dependent Identity (USID) so generated is cross checked in the pool of ID's whether it has been already generated using a matching logic. If a match is found then the ID is simply discarded and a list of discarded ID's is updated. In this paper an algorithm is proposed for the purpose which is discussed in section IV.

An earlier method based on data obfuscation and de-obfuscation was used to implement the Privacy Check mechanism. A new method based on information privacy is being proposed in section VI that would bring a significant shift in the user's perspective regarding the amount of control available to the user in the process of availing cloud services.

This USID generated at the Privacy Preserved Layer will then be used to access the Cloud Service at the Identity Utility Interface. The data which is being provided by the cloud as a result of the service is then sent back through the Privacy Preserved layer and the Address Mapping Layer to the Cloud user or consumer. Once a response arrives from the Cloud service provider, the list of USID's may have to be discarded/ extended. At the same time a translation also will have to be done from the OTIP addresses back to the original IP addresses at the Network Mapping Interface.

The functionalities of mapping the network addresses and user IDs ensures that the user is able to feel a greater amount of control over his identities. This enhanced control motivates the users to access the services being provided by the cloud service provider with greater amount of trust.

IV. USID GENERATION

Before proceeding to the understanding of process involved in USID generation there are the two aspects which has to be referred. They are (i) Proposed format of USID and (ii) Identified Data Structures needed to build an algorithm for Privacy preserved Match Logic [MaLog(PpC)]

A. Unique Service Dependent Identity (USID)Format:

USID that is generated and later allocated to each user should be *Concatenation* of three entities which can be given as

USID = Concat { OTip ,TS<req>,CS<Type>}

Where,

(i) Owned Translated IP reference *(OTip)*
(ii) Time Stamp of Cloud Service request *(TS <req>)*
(iii) Coded Type of Cloud Service *(CS <type>)*

The USID is a concatenation of the Owned Translated IP address, a time stamp and the coded type of service. OTIP has been incorporated into the USID in order to ensure location privacy i.e., to prevent any unauthorized third party from either knowing directly or deriving user information based on the user IP address. The main purpose of incorporating the coded Type of service is to ensure that the user identities which are being generated are service dependent. For different services being availed by the user different USID's will be generated.

B. Identified Data Structures:

The main functionality of Match logic is to ensure that the duplication of generated identities is avoided and that the USID's once generated are prevented from being allocated to any new user. The match logic requires the application of the following data structures:

PrC<area> : Area where Privacy to be checked USID is maintained.
PrK<area> : Area where key for privacy control is maintained ***POOL< id>:*** A place where all privacy preserved Identities of cloud users will be maintained.
ArG<id> : An identity of cloud user prepared as in section (4.1) for which privacy is under test.
PrCt<key> : A key for privacy control to verify the privacy.
PrL<id> : An identity with Privacy Leakage which should not be considered for further use.
PrL<area>: Area where all PrL<id> will be restored for future reference.
PrP<id > : An USID which has to be issued to the cloud user
PrP<area>: Area where all PrP<id> were restored for future reference.

V. ALGORITHM FOR MALOG(PPC)

Step:1 Place ArG<id> in PrC<area> for Privacy Preserving Check
Step: 2 Decide PrCt<key> and maintain in Prk<area>
Step: 3 Compare [ArG<id> , POOL<id>] based on PrCt<key> existing in Prk<area>
Step: 4 **If** MATCH { ArG<id> ,POOL<id>} is *successful*
then Send PrL<id> to PrL<area>
Update POOL<id> by removing
PrL <id> [matched Identities of ARG<id>]
Else
Send ARG<id> as PrP<id> to PrP<area>
Update POOL<id> by adding PrP<id>
Step :5 Issue PrP<id> to cloud user to avail Cloud Service from Cloud.

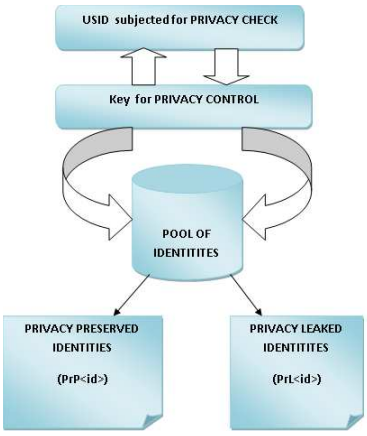

Figure 2. Functional diagram of USID generation process

The USID whose privacy has to be checked is denoted as the argument id, ArG<id>. This argument id is placed in the privacy check area, PrC<area> for privacy check. One of the essential of the privacy preserving methodology being proposed is that the current method enforces a certain amount of user control over the process of privacy preservation. This is done by allowing the user to specify a key, PrCt<key>. This key is chosen such that it is unknown to the cloud service provider and it is selected from the PrK<area>.The key is then used such that sensitive information can be shared by the cloud user with the cloud service provider without compromising the privacy of the inherent information. In case the information being shared is of a primordial nature then the user will be able to set preferences such that the use of the key is avoided.

The argument id is then compared with the pool of id's available. In case a match is found then the identity whose privacy has been leaked is transferred to the PrL <area>, i.e., a list of Privacy Leaked USID's is maintained. Furthermore, this privacy leaked USID is removed from the pool of unallocated USID's as represented in the *figure 2*. In a similar manner if the match is unsuccessful then the list of previously allocated whose privacy has been preserved is updated by placing it in the PrP <area>.

The current generated Privacy Preserved USID, PrP <id> is then allocated to the cloud user which enables the cloud user to access cloud services without any doubts regarding the privacy of the identities being used.

VI. USING INFORMATIONAL PRIVACY TO IMPLEMENT PRIVACY CHECK

The Privacy check mechanism is based on the concept of informational privacy. This method ensures that both the cloud user as well the cloud service provider have a sufficient amount of control over the amount of identity information which is being revealed during the process of accessing cloud service. The user specifies the amount of transparency that has to be incorporated as well as the possible users to whom this information has to be made available basing on Type of Service attribute available in the USID. Different Types of Service being provided by the user may require the active participation of different user attributes. The relevance as well as the selection of these attributes for each Type of service is subjective and it may depend on the service being provided the effect of a privacy leakage on the criticality of the application. These attributes are called Personalized Data Attributes (PDA). Basing on the Type of Service being provided to the user, a Boolean function called the Transparency Purpose in Cloud is being generated. The user specifies the transparency of an attribute as a Boolean function of these PDA's and it is called the Transparency Purpose in Cloud (TPC).

The Boolean function, TPC(PDA) may be defined as

TPC(PDA)=

1, ***if the PDA requires transparency***
0, ***if the PDA does not require transparency***
X, ***represents don't care condition***

When a user requests for a service the cloud service provider refers to the TPC to identify those attributes whose privacy needs to be protected, i.e., the Privacy Protected Attributes.

The specification of the transparencies by the user, the subsequent referral by the cloud service provider and controlled revelation of the identity attributes to the entities within the cloud makes this technique to be ideally suited for most privacy protection mechanisms.

This requires the creation and maintenance of a table which contains a list of PDA's as the fields. The tuples include the various TPC's along with the Boolean functions associated with the PDA's. Consider the scenario in which the cloud is assumed to be providing the services which related to Bank Loan, Insurance, Medical and Mailing applications. For example, a user may be considered to be having the PDA's of Name, age, sex, salary, designation and no. of children. All the PDA's may not be relevant to the same extent in the different TPC's. In the case of the Bank Loan except Age and Salary all other PDA's may have to be privacy protected. This is represented by mapping the Boolean function to a 1 in the age and sex fields and a 0 to the other fields of the Bank Loan TPC.

This may also be represented as

Bank Loan → {Salary ,Age}

Furthermore, in the case of the Insurance PDA except Age no other attribute needs to be privacy protected. However the attribute no. of children is a typical example of a don't care attribute as the TPC may map to either 0 or 1 depending upon the type of insurance.

For example, in the case of the Family Insurance the no. of children attribute will have to be maintained as Privacy protected whereas in the case of the Individual Insurance it will not be necessary to do so. For these attributes the TPC will map to a don't care (X) indicating that the privacy is not a major concern for the user irrespective of the purpose of utility.

Thus this informational privacy check considers the personalized details of Cloud User and maintain transparency of data attributes depending on their 'usage'/'Utility'/'Interest' or 'Concern' using Informational Privacy Protected Linear Table (IppL- Table) of the Cloud User leading to the Informational Privacy protection from 'others' (Anybody who are not suppose to be involved either directly and indirectly during cloud transaction).

Inputs to Build IppL-Table:

PersonalizedDataAttributes(PDA)={A1,A2,A3,A4,……..An}

Ex: A1 ={ Age} ; A2= {Sex} ; A3 ={Name}

Transparency Purpose in Cloud (TPC)={ P1, P2,P3,P4,……Pn}

Ex: P1 = {Banking} ; P2 ={Insurance} ; P3 = {Social}

TPC\PDA	A1	A2	A3	A4	--	An
P1	X					
P2		X				
P3				X		
P4						
--		X				
Pn	X			X		

Note: X - It Can be '0' or '1' or 'dc(Don't Care)'

Figure 3. Basic IppL table.

TPC\PDA	Name	Age	Sex	salary	Job title	No. of of Children
Bank Loan	0	1	0	1	0	0
Insurance	0	1	0	0	0	X
Medical	0	1	X	0	0	0
Mailing	0	0	0	0	0	0

Fig.4. IppL table for example cloud service scenario

By efficiently creating and maintaining the IppL table the user can be assured of Informational Privacy protection with good control on Personalized Data.

VII. CONCLUSIONS

The proposed model attempts to enhance the privacy of sensitive user information. The sensitive information relating the user's identity may be either directly provided by the cloud user in the process of trying to access the cloud service or it may also be derived based on secondary information such as IP addresses. The proposed model solves both the problems.

In the first case, the user provides sensitive information to the cloud service provider with the intention of accessing cloud service. If the information being provided by the user is sensitive then a mechanism has to be in place to prevent the cloud service provider from directly accessing it. On the other hand if the information being provided is not sensitive then the cloud service provider must be allowed to access it. This is made possible by applying the concept of data obfuscation and de-obfuscation available at the Unique User Identity Generator of the privacy preserved layer.

The proposed model attempts to solve the problem of preventing derivation of user identity based on secondary information such as IP addresses. It accomplishes this by mapping the original IP addresses to OTIP addresses. Using this mapping the cloud service provider is being prevented from accessing the original IP addresses. Rather it is able to access only the Owned Translated IP addresses which have been generated from an IP address pool. These OTIP's along with the time stamp are used to generate a USID from a pool of ID's using a mapping function. Comparison is made to a list of already created USID's so that the generation of duplicate USID's is prevented. For the purpose of checking for any duplicate USID's a match logic is employed. This process of verification is called Privacy Check.

A method based on information privacy is proposed to implement this Privacy Check. The proposed method takes into consideration the various Personal Data Attributes (PDA's) of the user and a boolean function called Transparency Purpose in Cloud (TPC) to determine which identity attributes have to be Privacy Protected i.e., whose visibility has to be made transparent. This specification of attributes by the user whose transparency is to be maintained and the later referral by the cloud service provider while allowing the user to access the cloud services brings about a significant shift in the user's perception regarding an increased amount of control.

An algorithm that may be used for performing the Match Logic has been proposed. The output of this Match Logic is that Privacy preserved and service dependent user identities are being generated. The Match logic mechanism also ensures that any privacy leaked identities are prevented from being allocated to the users. Furthermore, the pool of privacy preserved identities as well as the privacy leaked identities is updated. Finally the currently generated privacy preserved identity is allocated, thereby enabling the user to access cloud services.

The cloud service provider then provides services to the user based on the USID and the original user identities are masked away from the cloud service provider. As the USID is service dependent, there is a chance of multiple USID's being generated for the same user. A list of these is also maintained and the lifetime of the USID's may be extended or destroyed based on the whether the user likes to access continued service.

Thus this model prevents the direct access of user identification information or even the derivation of such information based on IP addresses. If this model is used then the users can be assured of the privacy of their identity information when they use the cloud service. This encourages more users to access the cloud services thereby resulting in economies of scale.

REFERENCES

[1] Hiroyuki Sato, Atushi Kanai, Shigeaki Tanimoto. "A Cloud Trust Model in a Security Aware Cloud". 2010 10th Annual International Symposium on Applications and the Internet.

[2] Robert Gellman, "Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing". World Privacy Forum. 2009.

[3] Hui (Wendy) Wang,"Ambiguity: Hide the Presence of Individuals and Their Privacy with Low Information Loss". International Conference on Management of Data, COMAD 2008.

[4] Miranda Mowbray, Siani Pearson, "A Client-Based Privacy Manager for Cloud Computing", COMSWARE'09, June 16-19,2009.

[5] V.S.Iyengar, "Transforming data to satisfy privacy constraints", In Proc. of SIGKDD, 2002.

[6] ISO, Information Security Management System, ISO/IEC 27001:2005, 2005.

[7] Mehmet Ercan Nergiz, Maurizio Atzori, and Chris Clifton, "Hiding the presence of individuals form shared databases", In Proc. of ACM Management of Data (SIGMOD), 2007.

[8] X.Xiao and Y.Tao, "Personalized privacy preservation", In Proc. of ACM Conference on management of data, 2006.

[9] PRIME, Privacy and Identity Management for Europe, 2008.

[10] Grossman, R.L., The case of Cloud Computing, IEEE - 2009.

[11] Maria, A.F., Fenu G., and Surcis, S. "An approach to Cloud Computing Network", IEEE-2008.