



Privacy Preservation and Authentication Method in Cloud Computing

Mayank Kumar Jain, Ratan Singh, Shivani Shrivastri
CSE & RGPM, Bhopal,
India

Abstract— Cloud computing is providing a platform for sharing resources, services and information among the people and organizations across the globe. Cloud providers use virtualization technologies combined with self-service abilities for computing resources via network infrastructure. The recent developments in cloud computing technology show an increase in security, privacy and trust related issues, in many ways, which haven't been visaged by the ones who have been designing cloud environments. Still most of the organizations are not moving to cloud computing due to lack of trust on service provider. Privacy preserving has originated as an important concerns with the reference to success of cloud computing. Privacy preserving deals with protecting the privacy of individual data or sensitive knowledge without sacrificing the utility of the data. Current authentication methods, whether proxy based or biometrics based, are not user-centric and/or endanger users' (biometric) security and privacy. In this paper, a review on the concepts of cloud computing, trust and some open questions related to privacy preservation, authentication and security requirements in cloud computing environments will be discussed. The accounting for important elements of cloud environment that shapes the users trust and provides a way of evaluating each element's importance are also discussed.

Keywords— Cloud Computing, Cloud Security, Privacy Preservation, Authentication, Trust model.

I. BACKGROUND

A. Introduction: Cloud computing [1] can be defined as new computing that has focus on both industry and academia. Cloud computing is the result of evolution and adoption of existing technologies and paradigms. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources(e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models [1].Massive growth in digital data, changing data storage requirements, better broadband facilities and Cloud computing led to the emergence of cloud databases .Cloud Storage, Data as a service (DaaS) and Database as a service (DBaaS) are the different terms used for data management in the Cloud. They differ on the basis of how data is stored and managed. Cloud storage is virtual storage that enables users to store documents and objects. Drop box , iCloud etc. are popular cloud storage services. DaaS allows user to store data at a remote disk available through Internet. Cloud storage cannot work without basic data management services. So, these two terms are used interchangeably. DBaaS is one step ahead. It offers complete database functionality and allows users to access and store their database at remote disks anytime from any place through Internet. Amazon's Simple DB, Amazon RDS, Google's Big Table, Yahoo's Sherpa and Microsoft's SQL Azure Database are the commonly used databases in the Cloud [2].

B. Cloud Computing Services: There are main three types of cloud computing service models:

Software as a service (SaaS)-SaaS can be defined as the software that is deployed over the internet. A complete software is available over the cloud any customer can use that software on "pay-as-you-go" basis.[3] The SaaS provides on-demand access of software to the clients. One more characteristic of SaaS is that it delivers the software in "one too many" model. In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients.

Platform as a service (PaaS)-In platform as a service model, service provider provides hardware and software to the customer which is needed by him to database and web server. PaaS is a form help enterprise developers quickly develop software. In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming language execution environment and of cloud computing that holds web- potential write and test customer or employee facing application.

Infrastructure as a service (IaaS)-It is the most basic cloud service model. It provides computers physical or virtual machines and other resources. IaaS clouds often offer additional resources such as a virtual-machine. Disk image library, raw block storage, and file or object storage, firewalls, load balancers, IP addresses, virtual local area networks and software bundles [3].

Database As A Service (DAAS)-Cloud database is designed for virtualized computer environment. It is not as simple as taking relational database and deploying it over a cloud server.[4] Cloud database as a service has to fulfil all the characteristics of relational database as well as cloud database. There are two terms used for data storage in cloud DaaS (Data as a service) & DbaaS (Database as a service). In data as a service only a space is provided over the cloud to store the data but in database as a service client can store data as well as he can run queries over the data to alter them and get some useful information from the database. Cloud database is created over the service provider site. So security should be very high in the cloud database because client has to protect his data from the outsider as well as he has to protect the data from the service provider also. It might be possible that database has some harm from the cloud database provider.

Irrespective of the above mentioned service models [4]; cloud services can be deployed in four ways depending upon the customers' requirements:

Public Cloud: A cloud infrastructure is provided to many customers and is managed by a third party [5]. Multiple enterprises can work on the infrastructure provided, at the same time. Users can dynamically provision resources through the internet from an off-site service provider. Wastage of resources is checked as the users pay for whatever they use.

Private Cloud: Cloud infrastructure, made available only to a specific customer and managed either by the organization itself or third party service provider [5]. This uses the concept of virtualization of machines, and is a proprietary network.

Community cloud: Infrastructure shared by several organizations for a shared cause and may be managed by them or a third party service provider.

Hybrid Cloud: A composition of two or more cloud deployment models, linked in a way that data transfer takes place between them without affecting each other.

II. SECURITY ISSUES IN CLOUD DATABASE

Scalability-Cloud database should be scalable so that it can store as much data as possible when number of users in the cloud increases.

Heterogeneity-Cloud database should support all types of users i.e. users working on various platforms.

Data Intrusion: Data Intrusion [6] is another security risk that may occur with a cloud provider. Undesirable alteration of user data may commence due to intrusion. If any intruder can gain access to the account password, then he/she will be able to do any kind of unwanted changes to the account's private documents.

Data Integrity: The stored data in the cloud storage may suffer from enormous damage occurring during the transition operations from or to the cloud storage provider. It is very essential to maintain the integrity of data. The risk of attacks from both inside and outside the cloud provider exists and should be considered.

Non- Repudiation: It guarantees the transmission of message between parties and gives the assurance that someone cannot deny something.. It ensures that a party cannot deny the genuineness of their signature on a document or the sending of a message that they originated. Non-repudiation is a major concern for data security. Non-repudiation is often used for signatures, digital contracts, and email messages

Confidentiality: The data should be kept secured and should not be exposed to anyone at any cost. Confidentiality [6] of data is another security issue associated with cloud computing.. The users do not want their confidential data to be disclosed to any service provider. But it is not always possible to encrypt the data before storing it in cloud.

Access control

Access management [7] is one of the toughest issues facing cloud computing security. One of the fundamental differences between traditional computing and cloud computing is the distributed nature of cloud computing. Within cloud computing, access management must therefore be considered from a federated sense, where an identity and access management solution is utilized across multiple cloud services and potentially multiple CSPs. Access control can be separated into the following functions:

Authentication: An organization can utilize cloud services across multiple CSPs, and can use these services as an extension of its internal, potentially non-cloud services. It is possible for different cloud services to use different identity and credential providers, which are likely different from the providers used by the organization for its internal applications. The credential management system used by the organization must be consolidated or integrated with those used by the cloud services.

Authorization: Requirements for user profile and access control policy vary depending on whether the cloud user is a member of an organization, such as an enterprise, or as an individual. Access control requirements include establishing trusted user profile and policy information, using it to control access within the cloud service, and doing this in an auditable way. Once authentication is done, resources can be authorized locally within the CSP. Many of the authorization mechanisms that are used in traditional computing environments can be utilized in a cloud setting.

Confidentiality: In Cloud computing, confidentiality plays a major part especially in maintaining control over organizations data situated across multiple distributed databases across data centers.

Integrity: This attribute concerns with accuracy, completeness and validity of information in regard with business requirement policies and expectations. The integrity requirement lies in applying the ACID (atomicity, consistency, isolation and durability) within the cloud domain mainly when accessing data.

Non-repudiation: This attribute concerns with the ability to prevent users from refusing the responsibility of the actions performed.

Availability: This feature concerns with information being operational
Cloud Computing Concerns

Gartner has conducted an exploration regarding the information security issues that should be considered when dealing with Cloud computing. The following list contains seven security issues highlighted by Gartner [10]:

Privileged user access: Inquire about who has specialized access to data, and about the hiring and management of such administrators.

Regulatory compliance: Make sure that the vendor is willing to undergo external audits and/or security certifications.

Data location: Does the provider allow for any control over the location of data?

Data segregation: Make sure that encryption is available at all stages, and that these encryption schemes were designed and tested by experienced professionals.

Recovery: Find out what will happen to data in the case of a disaster. Do they offer complete restoration?

Investigative support: Does the vendor have the ability to investigate any inappropriate or illegal activity?

Long-term viability: What will happen to data if the company goes out of business?

III. LITERATURE SURVEY

A large portion of system breaches are caused by authentication failure, either during the login process or in the post authentication session; these failures are themselves related to the limitations associated with existing authentication methods. Current authentication methods, whether proxy based or biometrics based, are not user-centric and/or endanger users' (biometric) security and privacy. [1] propose a biometrics based user-centric authentication approach. This method involves introducing a reference subject (RS), securely fusing the user's biometrics with the RS, generating a Bio Capsule (BC) from the fused biometrics, and employing BCs for authentication. Such an approach is user friendly, identity bearing yet privacy-preserving, resilient, and revocable once a BC is compromised. It also supports "one-click sign-on" across systems by fusing the user's biometrics with a distinct RS on each system. Moreover, active and non-intrusive authentication can be automatically performed during post-authentication sessions. [9] Prove that the secure fusion based approach is secure against various attacks. Extensive experiments and detailed comparison with existing approaches show that its performance (i.e., authentication accuracy) is comparable to existing typical biometric approaches and the new BC based approach also possesses many desirable features such as diversity and revocability.

Emerging techniques for user authentication involve traditional biometric authentication, cognitive authentication, BCS, CB and the hybrid approach. Traditional biometrics binds users to their biological traits, either physiological traits (e.g., iris [2], palm print [3], sclera [4]) or behaviour traits (e.g., mouse dynamics [5], gait [6]). As indicated previously, a limitation of traditional biometrics is security, user privacy risk and irreplaceability.

Cognitive biometrics [7], [8] can be used to improve their vocabulary property. Cognitive biometrics represents a new approach which generates a "thought signature" of people using biological signals that characterize the brain's response to certain stimuli, giving a high degree of uniqueness to the individual. Revocability is provided by training a new thinking process and generating a new "thought signature" to replace the compromised one. However, catching brain signals requires special equipment. Also, the thinking process may change over time.

Biometric cryptosystems can be used for user authentication by matching the exactness of the outputted keys. The majority of BCSs require some biometric-dependent public information (known as helper data), which is not supposed to reveal much information about the biometrics; with the helper data, the cryptographic key is retrieved or extracted from the query biometrics. The helper data are either obtained by binding a chosen key to biometrics or derived only from biometrics. BCSs use different techniques to deal with biometric variance; for example, some schemes apply error correction codes [9], [10], while some others apply quantization [11]. The introduction of helper data, in some circumstances (e.g., when multiple copy of helper data extracted from the single biometrics are obtained) may create vulnerabilities [12], [13]. However, without using helper data it is believed that extracting a sufficiently long and revocable key is not feasible because of the information entropy limitation of most biometric characteristics [14].

Utilizing error-correction codes and cryptography, a concept secure sketch is generalized which allows error correction of a noisy input. Secure sketches can be used as primitives to build fuzzy extractors which extract a uniformly random string [15]. Secure sketches and fuzzy extractors, as primitive formalisms, have been used in concrete BCSs. Quantization has also been used frequently in BCSs [16], [17]. In the BCS using quantization techniques, several enrolment samples are trained to derive appropriate intervals for feature quantization. As in [17], the authors apply a context-based reliable component selection and construct intervals for the most reliable features of each subject. Such approaches require multiple samples from each subject to reliably extract helper data.

Cancellable biometrics applies a transformation on traditional biometrics and matches the biometrics in a transformed domain for authentication. Cancellable biometrics was first introduced by Ratha et al. in [18]. Pillai et al. presented a CB approach using random projections which embed biometrics from a higher dimensional space to a lower dimensional space [19]; however, it is shown that the system is less secure if an attacker obtains both the random projection parameters and the transformed patterns. Bio token was proposed by [20] to transform original biometric feature via scaling and translation into a transformed version; the transformed feature is then split into a stable part termed integer and unstable part. There are several questions associated with this approach, namely, how to design the function which separates biometric features into stable and unstable parts, and how to apply the approach to other biometrics.

Ouda et al.[21] proposed a token less cancellable biometrics. This approach extracts consistent bits from original iris codes by training a set of images from each subject. The consistent bits are mapped to another set of bits (system selected)

to constitute the protected Bio Code. This approach requires an enrolling user to provide enough training images to satisfy the “consistence”. The discriminative capability of the “consistent” sequence determines the performance; the length of a “consistent” sequence is critical to the security, which is not shown in the paper.

Some hybrid approaches using both BCS and CB are proposed. The bio hashing scheme [16], [31] operates as a key binding scheme but combined user-specific tokenized random numbers to generate a set of binary bit strings. Given the binary string, it is not feasible to recover biometric data.

Several works note that the improved performance of bio hashing could be achieved with subject-specific tokenized random numbers [14], [15], however if the token is stolen, the system accuracy deteriorates. Nanda Kumar et al. proposed a hardened fuzzy vault using a user-specific secret key or password [44]. Introducing user-specific information, however, has an impact on the usability of the biometric system. It was also pointed out that such a “stolen-key scenario” must be considered for system evaluation; otherwise biometrics is trivial since the system could rely on the key without any complications [38]. Introducing the additional factor, which is not intrinsically bound to the user, logically creates more vulnerability. It could suffer from the same issues of traditional proxy-based systems in that information can be stolen, lost or forgotten. The user-specific key is an additional factor correlated to each individual, which has the chance to reveal user-privacy. Further introducing a so-called user-specific key makes the identification under non-cooperative identification troublesome.

IV. CONCLUSION

Cloud computing is a computing in which large groups of remote servers are networked to allow centralized data storage and online access to computer services or resources. The major issue with the cloud database is that it requires a very high level security. Data are not always safe when they are stored inside cloud providers. In this paper we discuss the cloud database security. Privacy preserving deals with protecting the privacy of individual data or sensitive knowledge without sacrificing the utility of the data. Current authentication methods, whether proxy based or biometrics based, are not user-centric and/or endanger users’ (biometric) security and privacy. In this paper , a review on the concepts of cloud computing, trust and some open questions related to privacy preservation, authentication and security requirements in cloud computing environments will discussed.

REFERENCES

- [1] Yan Sui, Xukai Zou, Eliza Y. Du, Feng Li, “Design and Analysis of a Highly User-Friendly, Secure, Privacy-Preserving, and Revocable Authentication Method”, IEEE TRANSACTIONS ON COMPUTERS, VOL. 63, NO. 4, APRIL 2014, pp-902-916.
- [2] J. Daugman, “How Iris Recognition Works,” IEEE Trans. Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 21-30, Jan. 2004.
- [3] J. Dai, J. Feng, and J. Zhou, “Robust and Efficient Ridge-Based Palmprint Matching,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 34, no. 8, pp. 1618-1632, Aug. 2012.
- [4] Z. Zhou, E. Du, N. Thomas, and E. Delp, “A New Human Identification Method: Sclera Recognition,” IEEE Trans. Systems, Man, and Cybernetics, Part A: Systems and Humans, vol. 42, no. 3, pp. 571-583, May 2012.
- [5] A. Ahmed and I. Traore, “A New Biometric Technology Based on Mouse Dynamics,” IEEE Trans. Dependable and Secure Computing, vol. 4, no. 3, pp. 165-179, July/Sept. 2007.
- [6] J. Zhang, J. Pu, C. Chen, and R. Fleischer, “Low-Resolution Gait Recognition,” IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 40, no. 4, pp. 986-996, Aug. 2010.
- [7] L. Faria, V. Sa, and S. de Magalhaes, “Multimodal Cognitive Biometrics,” Proc. Sixth Iberian Conf. Information Systems and Technologies (CISTI), pp. 1-6, June 2011.
- [8] K. Revett and S. Tenreiro de Magalhes, “Cognitive Biometrics: Challenges for the Future,” Proc. Sixth Int’l Conf. Global Security, Safety, and Sustainability, pp. 79-86, 2010.
- [9] A. Juels and M. Sudan, “A Fuzzy Vault Scheme,” Designs, Codes and Cryptography, vol. 38, pp. 237-257, 2006.
- [10] A. Juels and M. Wattenberg, “A Fuzzy Commitment Scheme,” Proc. Sixth ACM Conf. Computer and Comm. Security (CCS ’99), pp. 28-36, 1999.
- [11] C. Vielhauer, R. Steinmetz, and A. Mayerhoefer, “Biometric Hash Based on Statistical Features of Online Signatures,” Proc. 16th Int’l Conf. Pattern Recognition (ICPR ’02), vol. 1, pp. 123-126, 2002.
- [12] T. Ignatenko and F.M.J. Willems, “Information Leakage in Fuzzy Commitment Schemes,” IEEE Trans. Information Forensics and Security, vol. 5, no. 2, pp. 337-348, June 2010.
- [13] K. Simoens, P. Tuyls, and B. Preneel, “Privacy Weaknesses in Biometric Sketches,” Proc. 30th IEEE Symp. Security and Privacy, pp. 188-203, May 2009.
- [14] C. Rathgeb and A. Uhl, “A Survey on Biometric Cryptosystems and Cancellable Biometrics,” EURASIP J. Information Security, vol. 2011, no. 3, pp. 1-25, 2011.
- [15] Y. Dodis, L. Reyzin, and A. Smith, “Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data,” Proc. Advances in Cryptology (Eurocrypt), vol. 3027, pp. 523-540, 2004.
- [16] C. Rathgeb and A. Uhl, “An Iris-Based Interval-Mapping Scheme for Biometric Key Generation,” Proc. Sixth Int’l Symp. Image and Signal Processing and Analysis, pp. 511-516, Sept. 2009.
- [17] C. Rathgeb and A. Uhl, “Privacy Preserving Key Generation for Iris Biometrics,” Proc. 11th IFIP TC 6/TC 11 Int’l Conf. Comm. And Multimedia Security, pp. 191-200, 2010.

- [18] N. Ratha, J. Connell, and R. Bolle, "Enhancing Security and Privacy in Biometrics-Based Authentication Systems," *IBM Systems J.*, vol. 40, pp. 614-634, 2001.
- [19] J. Pillai , V. Patel, R. Chellappa , and N. Ratha , "Secure and Robust Iris Recognition Using Random Projections and Sparse Representations," *IEEE Trans. Pattern Analysis and Machine Intelligence* ,vol. 33, no. 9, pp. 1877-1893, Sept. 2011.
- [20] T. Boulton, "Robust Distance Measures for Face-Recognition Supporting Revocable Biometric Tokens," *Proc. Seventh Int'l Conf. Automatic Face and Gesture Recognition*, pp. 560-566, Apr.2006.
- [21] O. Ouda, N. Tsumura, and T. Nakaguchi, "Bio Encoding: A Reliable Token less Cancellable Biometrics Scheme for Protecting Iris codes," *IEICE Trans. Information and Systems*, vol. E93-D, no. 7,pp. 1878-1888, July 2010.