Check for updates

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2014, Article ID 190903, 9 pages
http://dx.doi.org/10.1155/2014/190903

Review Article

Data Security and Privacy in Cloud Computing

Yunchuan Sun,[1] Junsheng Zhang,[2] Yongping Xiong,[3] and Guangyu Zhu[4]

[1] *Business School, Beijing Normal University, Beijing 100875, China*
[2] *IT Support Center, Institute of Scientific and Technical Information of China, Beijing 100038, China*
[3] *State Key Lab of Networking and Switching Tech., Beijing University of Posts and Telecommunications, Beijing 100876, China*
[4] *Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Junsheng Zhang; zhangjs@istic.ac.cn

Received 25 April 2014; Accepted 26 June 2014; Published 16 July 2014

Academic Editor: Zhangbing Zhou

Copyright © 2014 Yunchuan Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data security has consistently been a major issue in information technology. In the cloud computing environment, it becomes particularly serious because the data is located in different places even in all the globe. Data security and privacy protection are the two main factors of user's concerns about the cloud technology. Though many techniques on the topics in cloud computing have been investigated in both academics and industries, data security and privacy protection are becoming more important for the future development of cloud computing technology in government, industry, and business. Data security and privacy protection issues are relevant to both hardware and software in the cloud architecture. This study is to review different security techniques and challenges from both software and hardware aspects for protecting data in the cloud and aims at enhancing the data security and privacy protection for the trustworthy cloud environment. In this paper, we make a comparative research analysis of the existing research work regarding the data security and privacy protection techniques used in the cloud computing.

1. Introduction

Cloud computing has been envisioned as the next generation paradigm in computation. In the cloud computing environment, both applications and resources are delivered on demand over the Internet as services. Cloud is an environment of the hardware and software resources in the data centers that provide diverse services over the network or the Internet to satisfy user's requirements [1].

The explanation of "cloud computing" from the National Institute of Standards and Technology (NIST) [2] is that *cloud computing enables 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.* According to the explanation, cloud computing provides a convenient on-demand network access to a shared pool of configurable computing resources. Resources refer to computing applications, network resources, platforms, software services, virtual servers, and computing infrastructure.

Cloud computing can be considered as a new computing archetype that can provide services on demand at a minimal cost. The three well-known and commonly used service models in the cloud paradigm are *software as a service* (SaaS), *platform as a service* (PaaS), and *infrastructure as a service* (IaaS). In SaaS, software with the related data is deployed by a cloud service provider, and users can use it through the web browsers. In PaaS, a service provider facilitates services to the users with a set of software programs that can solve the specific tasks. In IaaS, the cloud service provider facilitates services to the users with virtual machines and storage to improve their business capabilities.

Cloud computing is closely related to but not the same as grid computing [3]. Grid computing integrates diverse resources together and controls the resources with the unified operating systems to provide high performance computing services, while cloud computing combines the computing and storage resources controlled by different operating systems to provide services such as large-scaled data storage and high performance computing to users. The overall picture of grid computing has been changed by cloud computing.

Distribution of data is in a new way of cloud computing comparing with the grid computing.

Cloud computing will enable services to be consumed easily on demand. Cloud computing has the characteristics such as on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing, and transference of risk. These merits of cloud computing have attracted substantial interests from both the industrial world and the academic research world. Cloud computing technology is currently changing the way to do business in the world.

Cloud computing is very promising for the IT applications; however, there are still some problems to be solved for personal users and enterprises to store data and deploy applications in the cloud computing environment. One of the most significant barriers to adoption is data security, which is accompanied by issues including compliance, privacy, trust, and legal matters [4, 5]. The role of institutions and institutional evolution is close to privacy and security in cloud computing [6].

Data security has consistently been a major issue in IT. Data security becomes particularly serious in the cloud computing environment, because data are scattered in different machines and storage devices including servers, PCs, and various mobile devices such as wireless sensor networks and smart phones. Data security in the cloud computing is more complicated than data security in the traditional information systems.

To make the cloud computing be adopted by users and enterprise, the security concerns of users should be rectified first to make cloud environment trustworthy. The trustworthy environment is the basic prerequisite to win confidence of users to adopt such a technology. Latif et al. discussed the assessment of cloud computing risks [7].

Before the data security issues are discussed, the functions of cloud computing are analyzed first. Cloud computing is also known as on-demand service. In the cloud computing environment, there is a cloud service provider that facilitates services and manages the services. The cloud provider facilitates all the services over the Internet, while end users use services for satisfying their business needs and then pay the service provider accordingly.

Cloud computing environment provides two basic types of functions: *computing* and *data storage*. In the cloud computing environment, consumers of cloud services do not need anything and they can get access to their data and finish their computing tasks just through the Internet connectivity. During the access to the data and computing, the clients do not even know where the data are stored and which machines execute the computing tasks.

Coming to data storage, data protection and security are the primary factors for gaining user's trust and making the cloud technology successfully used. A number of data protections and data security techniques have been proposed in the research field of cloud computing. However, data protection related techniques need to be further enhanced.

Services of cloud computing are provided across the entire computing spectrum. Nowadays, organizations and companies are moving and extending their business by adopting the cloud computing to lower their cost. This can contribute to free more man-powers to focus on creating strategic differentiation and business division of labor is clearer.

The cloud is growing continuously because it could provide high performance computational services at cheaper rates. Famous IT companies such as Microsoft (http://azure.microsoft.com/), Amazon (http://aws.amazon.com/), Google (https://cloud.google .com/), and Rakespace (http://www.rackspace.com/) have provided cloud service on the Internet.

The concept of cloud has a number of implementations based on the services from service providers. For example, Google Apps Engine, Microsoft Azure, and Amazon Stack are popular implementations of cloud computing provided by cloud service providers, that is, Google, Microsoft, and Amazon companies. Besides, the ACME enterprise implemented VMware based v-Cloud for permitting multiple organizations to share computing resources.

According to the difference of access scope, cloud can be divided into three types: *public cloud*, *private cloud,* and *hybrid cloud*. Public cloud is as the property of service provider and can be used in public, private cloud refers to being the property of a company, and hybrid cloud is the blends of public and private cloud. Most of the existing cloud services are provided by large cloud service companies such as Google, Amazon, and IBM. A private cloud is a cloud in which only the authorized users can access the services from the provider. In the pubic cloud anybody can use the cloud services whereas the hybrid cloud contains the concept of both public and private clouds.

Cloud computing can save an organization's time and money, but trusting the system is more important because the real asset of any organization is the data which they share in the cloud to use the needed services by putting it either directly in the relational database or eventually in a relational database through an application.

Cloud computing brings a number of attributes that require special attention when it comes to trusting the system. The trust of the entire system depends on the data protection and prevention techniques used in it. Numerous different tools and techniques have been tested and introduced by the researchers for data protection and prevention to gain and remove the hurdle of trust but there are still gaps which need attention and are required to be lined up by making these techniques much better and effective.

The meaning of security is plentiful. Security is the combination of confidentiality, the prevention of the unauthorized disclosure of information, integrity, the prevention of the unauthorized amendment or deletion of information, and availability, the prevention of unauthorized withholding of information [8].

The major issues in the cloud computing include resource security, resource management, and resource monitoring. Currently, there are no standard rules and regulations to deploy applications in the cloud, and there is a lack of standardization control in the cloud. Numerous novel techniques had been designed and implemented in cloud; however, these

techniques fall short of ensuring total security due to the dynamics of the cloud environment.

The inherent issues of data security, governance, and management with respect to control in the cloud computing are discussed in [9]. Sun et al. [10] highlighted the key security, privacy, and trust issues in the existing environment of cloud computing and help users to recognize the tangible and intangible threats related to its use. According to the authors, there are three major potential threats in cloud computing, namely, *security*, *privacy,* and *trust.* Security plays a critical role in the current era of long dreamed vision of computing as a utility. It can be divided into four subcategories: *safety mechanisms*, *cloud server monitoring or tracing*, *data confidentiality,* and *avoiding malicious insiders' illegal operations and service hijacking.*

A data security framework for cloud computing networks is proposed [11]. The authors mainly discussed the security issues related to cloud data storage. There are also some patents about the data storage security techniques [12]. Younis and Kifayat give a survey on secure cloud computing for critical infrastructure [13]. A security and privacy framework for RFID in cloud computing was proposed for RFID technology integrated to the cloud computing [14], which will combine the cloud computing with the Internet of Things.

In short, the foremost issues in cloud data security include data privacy, data protection, data availability, data location, and secure transmission. The security challenges in the cloud include threats, data loss, service disruption, outside malicious attacks, and multitenancy issues [15]. Chen and Zhao [16] analyzed privacy and data security issues in the cloud computing by focusing on privacy protection, data segregation, and cloud security. Data security issues are primarily at SPI (SaaS, PaaS, and IaaS) level and the major challenge in cloud computing is data sharing.

In this paper, we will review different security techniques and challenges for data storage security and privacy protection in the cloud computing environment. As Figure 1 shows, this paper presents a comparative research analysis of the existing research work regarding the techniques used in the cloud computing through data security aspects including data integrity, confidentiality, and availability. Data privacy issues and technologies in the cloud are also studied, because data privacy is traditionally accompanied with data security. Comparative studies on data security and privacy could help to enhance the user's trust by securing data in the cloud computing environment.

2. Data Integrity

Data integrity is one of the most critical elements in any information system. Generally, data integrity means protecting data from unauthorized deletion, modification, or fabrication. Managing entity's admittance and rights to specific enterprise resources ensures that valuable data and services are not abused, misappropriated, or stolen.

Data integrity is easily achieved in a standalone system with a single database. Data integrity in the standalone system is maintained via database constraints and transactions, which is usually finished by a database management system (DBMS). Transactions should follow ACID (atomicity, consistency, isolation, and durability) properties to ensure data integrity. Most databases support ACID transactions and can preserve data integrity.

Authorization is used to control the access of data. It is the mechanism by which a system determines what level of access a particular authenticated user should have to secure resources controlled by the system.

Data integrity in the cloud system means preserving information integrity. The data should not be lost or modified by unauthorized users. Data integrity is the basis to provide cloud computing service such as SaaS, PaaS, and IaaS. Besides data storage of large-scaled data, cloud computing environment usually provides data processing service. Data integrity can be obtained by techniques such as RAID-like strategies and digital signature.

Owing to the large quantity of entities and access points in a cloud environment, authorization is crucial in assuring that only authorized entities can interact with data. By avoiding the unauthorized access, organizations can achieve greater confidence in data integrity. The monitoring mechanisms offer the greater visibility into determining who or what may have altered data or system information, potentially affecting their integrity. Cloud computing providers are trusted to maintain data integrity and accuracy. However, it is necessary to build the third party supervision mechanism besides users and cloud service providers.

Verifying the integrity of data in the cloud remotely is the perquisite to deploy applications. Bowers et al. proposed a theoretical framework "Proofs of Retrievability" to realize the remote data integrity checking by combining error correction code and spot-checking [17]. The HAIL system uses POR mechanism to check the storage of data in different clouds, and it can ensure the redundancy of different copies and realize the availability and integrity checking [18]. Schiffman et al. proposed trusted platform module (TPM) remote checking to check the data integrity remotely [19].

3. Data Confidentiality

Data confidentiality is important for users to store their private or confidential data in the cloud. Authentication and access control strategies are used to ensure data confidentiality. The data confidentiality, authentication, and access control issues in cloud computing could be addressed by increasing the cloud reliability and trustworthiness [20].

Because the users do not trust the cloud providers and cloud storage service providers are virtually impossible to eliminate potential insider threat, it is very dangerous for users to store their sensitive data in cloud storage directly. Simple encryption is faced with the key management problem and cannot support complex requirements such as query, parallel modification, and fine-grained authorization.

3.1. Homomorphic Encryption. Encryption is usually used to ensure the confidentiality of data. Homomorphic encryption is a kind of encryption system proposed by Rivest et al. [21].

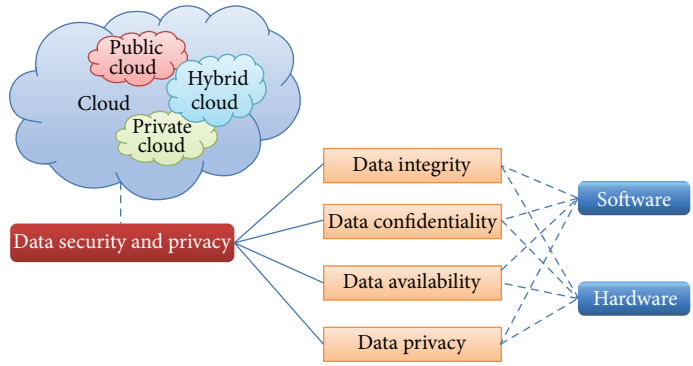

Figure 1: Organization of data security and privacy in cloud computing.

It ensures that the cipher text algebraic operation results are consistent with the clear operation after encryption results; besides, the whole process does not need to decrypt the data. The implementation of this technique could well solve the confidentiality of data and data operations in the cloud.

Gentry firstly proposed the fully homomorphic encryption method [22], which can do any operation that can be performed in clear text without decrypting. It is an important breakthrough in the homomorphic encryption technology. However, the encryption system involves very complicated calculation, and the cost of computing and storage is very high. This leads to the fact that the fully homomorphic encryption is still far from real applications.

A cryptographic algorithm named Diffie-Hellman is proposed for secure communication [23], which is quite dissimilar to the key distribution management mechanism.

For more flexibility and enhanced security, a hybrid technique that combines multiple encryption algorithms such as RSA, 3DES, and random number generator has been proposed [24]. RSA is useful for establishing secure communication connection through digital signature based authentication while 3DES is particularly useful for encryption of block data. Besides, several encryption algorithms for ensuring the security of user data in the cloud computing are discussed [25].

3.2. Encrypted Search and Database. Because the homomorphic encryption algorithm is inefficient, researchers turn to study the applications of limited homomorphic encryption algorithm in the cloud environment. Encrypted search is a common operation.

Manivannan and Sujarani [26] have proposed a lightweight mechanism for database encryption known as transposition, substitution, folding, and shifting (TSFS) algorithm. However, as the numbers of keys are increased, the amount of computations and processing also increases.

In-Memory Database encryption technique is proposed for the privacy and security of sensitive data in untrusted cloud environment [27]. A synchronizer exists between the owner and the client for seeking access to the data. Client would require a key from the synchronizer to decrypt the encrypted shared data it receives from the owner. The synchronizer is utilized to store the correlated shared data and the keys separately. A shortcoming of this technique is that the delays occur due to the additional communication with the central synchronizer. However, this limitation can be mitigated by adopting group encryption and through minimizing communication between nodes and synchronizer.

Huang and Tso [28] proposed an asymmetric encryption mechanism for databases in the cloud. In the proposed mechanism, the commutative encryption is applied on data more than once and the order of public/private key used for encryption/decryption does not matter. Reencryption mechanism is also used in the proposed scheme which shows that the cipher-text data is encrypted once again for duality. Such schemes are very useful in the cloud applications where privacy is a key concern.

A privacy-preserving multikeyword ranked search approach over encrypted cloud data was proposed [29], which can search the encrypted cloud data and rank the search results without leakage of the user's privacy.

3.3. Distributive Storage. Distributive storage of data is also a promising approach in the cloud environment. AlZain et al. [30] discussed the security issues related to data privacy in the cloud computing including integrity of data, intrusion, and availability of service in the cloud. To ensure the data integrity, one option could be to store data in multiple clouds or cloud databases. The data to be protected from internal or external unauthorized access are divided into chunks and Shamir's secret algorithm is used to generate a polynomial function against each chunk. Ram and Sreenivaasan [31] have proposed a technique known as security as a service for securing cloud data. The proposed technique can achieve maximum security by dividing the user's data into pieces. These data chunks are then encrypted and stored in separated databases which follow the concept of data distribution over cloud. Because each segment of data is encrypted and separately distributed in databases over cloud, this provides enhanced security against different types of attacks.

Arfeen et al. [32] describe the distribution of resources for cloud computing based on the tailored active measurement. The tailored measurement technique is based on the network design and the specific routes for the incoming and outgoing traffic and gradually changing the resources according to the user needs. Tailored measurement depends on the computing

resources and storage resources. Because of the variable nature of networks, the allocation of resources at a particular time based on the tailored active method does not remain optimal. The resources may increase or decrease, so the system has to optimize changes in the user requirement either offline or on-line and the resource connectivity.

3.4. Hybrid Technique. A hybrid technique is proposed for data confidentiality and integrity [33], which uses both key sharing and authentication techniques. The connectivity between the user and the cloud service provider can be made more secure by utilizing powerful key sharing and authentication processes. RSA public key algorithm can be used for secure distribution of the keys between the user and cloud service providers.

A three-layered data security technique is proposed [34]: the first layer is used for authenticity of the cloud user either by one factor or by two factor authentications; the second layer encrypts the user's data for ensuring protection and privacy; and the third layer does fast recovery of data through a speedy decryption process.

An event-based isolation of critical data in the cloud approach is proposed [35], TrustDraw, a transparent security extension for the cloud which combines virtual machine introspection (VMI) and trusted computing (TC).

3.5. Data Concealment. Data concealment could also be used to keep the data confidentiality in the cloud. Delettre et al. [36] introduced a concealment concept for databases security. Data concealment approaches merge real data with the visual fake data to falsify the real data's volume. However, authorized users can easily differentiate and separate the fake data from the real data. Data concealment techniques increase the overall volume of real data but provide enhanced security for the private data. The objective of data concealment is to make the real data safe and secure from malicious users and attackers. Watermarking method can serve as a key for the real data. Only the authorized users have key of watermarking, so the authentication of users is the key to ensure the true data to be accessible for right users.

3.6. Deletion Confirmation. Deletion confirmation means that data could not be recovered when users delete their data after the deletion confirmation. The problem is very serious, because more than one copy exists in the cloud for the security and convenience of data recovery. When users delete their data with confirmation, all the copies of data should be deleted at the same time. However, there are some data recovery technologies that could recover the data deleted by users from the hard disks. So the cloud storage providers should ensure that the deleted data of users could not be recovered and used by other unauthenticated users.

To avoid the data be recovered and unauthenticatedly used, a possible approach is to encrypt the data before uploading to the cloud storage space. FADE system [37] is based on technologies such as Ephemerizer. In the system, data are encrypted before they are uploaded to the cloud storage. When users decide to delete their data, the system just to apply the specific strategy to all the storage space could be covered with new data for replacing the deletion operation.

4. Data Availability

Data availability means the following: when accidents such as hard disk damage, IDC fire, and network failures occur, the extent that user's data can be used or recovered and how the users verify their data by techniques rather than depending on the credit guarantee by the cloud service provider alone.

The issue of storing data over the transboarder servers is a serious concern of clients because the cloud vendors are governed by the local laws and, therefore, the cloud clients should be cognizant of those laws. Moreover, the cloud service provider should ensure the data security, particularly data confidentiality and integrity. The cloud provider should share all such concerns with the client and build trust relationship in this connection. The cloud vendor should provide guarantees of data safety and explain jurisdiction of local laws to the clients. The main focus of the paper is on those data issues and challenges which are associated with data storage location and its relocation, cost, availability, and security.

Locating data can help users to increase their trust on the cloud. Cloud storage provides the transparent storage service for users, which can decrease the complexity of cloud, but it also decreases the control ability on data storage of users. Benson et al. studied the proofs of geographic replication and succeeded in locating the data stored in Amazon cloud [38].

4.1. Reliable Storage Agreement. The most common abnormal behavior of untrusted storage is that the cloud service providers may discard part of the user's update data, which is hard to be checked by only depending on the simple data encryption. Additionally, a good storage agreement needs to support concurrent modification by multiple users.

Mahajan et al. proposed Depot which can guarantee Fork-Join-Causal-Consistency and eventual consistency [39]. It can effectively resist attacks such as discarding and it can support the implementation of other safety protections in the trusted cloud storage environment (such as Amazon S3).

Feldman et al. proposed SPORC [40], which can implement the safe and reliable real-time interaction and collaboration for multiple users with the help of the trusted cloud environment, and untrusted cloud servers can only access the encrypted data.

However, operation types supported by reliable storage protocol support are limited, and most of the calculations can only occur in the client.

4.2. Reliability of Hard-Drive. Hard-drive is currently the main storage media in the cloud environment. Reliability of hard disks formulates the foundation of cloud storage. Pinheiro et al. studied the error rate of hard-drives based on the historical data of hard-drive [41]. They found that the error rate of hard-drives is not closely relevant to the temperature and the frequency to be used, while the error rate of hard-drives has the strong clustering characteristics.

Current SMART mechanism could not predict the error rate of hard disks. Tsai et al. studied the correlation between the soft error and hard error of hard disks, and they also found that the soft error could not predict the hard errors of hard-drives precisely [42], only about 1/3 probability that hard errors follow the soft errors.

5. Data Privacy

Privacy is the ability of an individual or group to seclude themselves or information about themselves and thereby reveal them selectively [43]. Privacy has the following elements.

(i) When: a subject may be more concerned about the current or future information being revealed than information from the past.

(ii) How: a user may be comfortable if his/her friends can manually request his/her information, but the user may not like alerts to be sent automatically and frequently.

(iii) Extent: a user may rather have his/her information reported as an ambiguous region rather than a precise point.

In commerce, consumer's context and privacy need to be protected and used appropriately. In organizations, privacy entails the application of laws, mechanisms, standards, and processes by which personally identifiable information is managed [44].

In the cloud, the privacy means when users visit the sensitive data, the cloud services can prevent potential adversary from inferring the user's behavior by the user's visit model (not direct data leakage). Researchers have focused on Oblivious RAM (ORAM) technology. ORAM technology visits several copies of data to hide the real visiting aims of users. ORAM has been widely used in software protection and has been used in protecting the privacy in the cloud as a promising technology. Stefanov et al. proposed that a path ORAM algorithm is state-of-the-art implementation [45].

The privacy issues differ according to different cloud scenarios and can be divided into four subcategories [44, 46, 47] as follows:

(i) how to enable users to have control over their data when the data are stored and processed in cloud and avoid theft, nefarious use, and unauthorized resale,

(ii) how to guarantee data replications in a jurisdiction and consistent state, where replicating user data to multiple suitable locations is an usual choice, and avoid data loss, leakage, and unauthorized modification or fabrication,

(iii) which party is responsible for ensuring legal requirements for personal information,

(iv) to what extent cloud subcontractors are involved in processing which can be properly identified, checked, and ascertained.

5.1. Service Abuse. Service abuse means that attackers can abuse the cloud service and acquire extra data or destroy the interests of other users.

User data may be abused by other users. Deduplication technology has been widely used in the cloud storage, which means that the same data often were stored once but shared by multiple different users. This will reduce the storage space and cut down the cost of cloud service providers, but attackers can access the data by knowing the hash code of the stored files. Then, it is possible to leak the sensitive data in the cloud. So proof of ownership approach has been proposed to check the authentication of cloud users [48].

Attackers may lead to the cost increase of cloud service. Fraudulent resource consumption is a kind of attack on the payment for cloud service. Attackers can consume the specific data to increase the cost for cloud service payment. Idziorek et al. proposed this question and researched on the detection and identification of fraud resource consumption [49].

5.2. Averting Attacks. The cloud computing facilitates huge amount of shared resources on the Internet. Cloud systems should be capable of averting Denial of Service (DoS) attacks.

Shen et al. analyzed requirement of security services in cloud computing [50]. The authors suggest integrating cloud services for trusted computing platform (TCP) and trusted platform support services (TSS). The trusted model should bear characteristics of confidentiality, dynamically building trust domains and dynamic of the services. Cloud infrastructures require that user transfers their data into cloud merely based on trust. Neisse et al. analyzed indifferent attacks scenarios on Xen cloud platform to evaluate cloud services based on trust. Security of data and trust in cloud computing is the key point for its broader adoption [51].

Yeluri et al. focused on the cloud services from security point of view and explored security challenges in cloud when deploying the services [52]. Identity management, data recovery and management, security in cloud confidentiality, trust, visibility, and application architecture are the key points for ensuring security in cloud computing.

5.3. Identity Management. Cloud computing provides a podium to use wide range of Internet-based services [53]. But besides its advantages, it also increases the security threat when a trusted third party is involved. By involving a trusted third party, there is a chance of heterogeneity of users which affects security in the cloud. A possible solution to this problem could be to use a trusted third party independent approach for Identity Management to use identity data on untrusted hosts.

Squicciarini et al. focused on problems of data leakage and loss of privacy in cloud computing [54]. Different levels of protections can be used to prevent data leakage and privacy loss in the cloud. Cloud computing provides new business services that are based on demand. Cloud networks have been built through dynamic virtualization of hardware, software, and datasets. Cloud security infrastructure and the trust reputation management play a vital role to upgrade the cloud services [55]. The Internet access security, server access

security, program access security, and database security are the main security issues in the cloud.

6. Conclusion

Cloud computing is a promising and emerging technology for the next generation of IT applications. The barrier and hurdles toward the rapid growth of cloud computing are data security and privacy issues. Reducing data storage and processing cost is a mandatory requirement of any organization, while analysis of data and information is always the most important tasks in all the organizations for decision making. So no organizations will transfer their data or information to the cloud until the trust is built between the cloud service providers and consumers. A number of techniques have been proposed by researchers for data protection and to attain highest level of data security in the cloud. However, there are still many gaps to be filled by making these techniques more effective. More work is required in the area of cloud computing to make it acceptable by the cloud service consumers. This paper surveyed different techniques about data security and privacy, focusing on the data storage and use in the cloud, for data protection in the cloud computing environments to build trust between cloud service providers and consumers.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research work was supported by International Science and Technology Cooperation Program of China under Grant no. 2014DFA11350 and the National Natural Science Foundation of China (Grant nos. 61371185, 61171014, and 61202436).

References

[1] N. Leavitt, "Is cloud computing really ready for prime time?" *Computer*, vol. 42, no. 1, pp. 15–25, 2009.

[2] P. Mell and T. Grance, "The nist definition of cloud computing," *National Institute of Standards and Technology*, vol. 53, no. 6, article 50, 2009.

[3] F. Berman, G. Fox, and A. J. G. Hey, *Grid Computing: Making the Global Infrastructure a Reality, Volume 2*, John Wiley and sons, 2003.

[4] M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," *IACR Cryptology EPrint Archive*, vol. 186, 2008.

[5] Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 843–859, 2013.

[6] N. Kshetri, "Privacy and security issues in cloud computing: the role of institutions and institutional evolution," *Telecommunications Policy*, vol. 37, no. 4-5, pp. 372–386, 2013.

[7] R. Latif, H. Abbas, S. Assar, and Q. Ali, "Cloud computing risk assessment: a systematic literature review," in *Future Information Technology*, pp. 285–295, Springer, Berlin, Germany, 2014.

[8] A. Avižienis, J. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11–33, 2004.

[9] Z. Mahmood, "Data location and security issues in cloud computing," in *Proceedings of the 2nd International Conference on Emerging Intelligent Data and Web Technologies (EIDWT '11)*, pp. 49–54, IEEE, September 2011.

[10] D. Sun, G. Chang, L. Sun, and X. Wang, "Surveying and analyzing security, privacy and trust issues in cloud computing environments," in *Proceedings of the International Conference on Advanced in Control Engineering and Information Science (CEIS '11)*, pp. 2852–2856, chn, August 2011.

[11] A. Pandey, R. M. Tugnayat, and A. K. Tiwari, "Data Security Framework for Cloud Computing Networks," *International Journal of Computer Engineering & Technology*, vol. 4, no. 1, pp. 178–181, 2013.

[12] D. A. Klein, "Data security for digital data storage," U.S. Patent Application 14/022,095, 2013.

[13] M. Y. A. Younis and K. Kifayat, "Secure cloud computing for critical infrastructure: a survey," Tech. Rep., Liverpool John Moores University, Liverpool, UK, 2013.

[14] S. Kardaş, S. Çelik, M. A. Bingöl, and A. Levi, "A new security and privacy framework for RFID in cloud computing," in *Proceedings of the 5th IEEE International Conference on Cloud Computing Technology and Science (CloudCom '13)*, Bristol , UK, 2013.

[15] A. Behl, "Emerging security challenges in cloud computing: an insight to cloud security challenges and their mitigation," in *Proceedings of the World Congress on Information and Communication Technologies (WICT '11)*, pp. 217–222, IEEE, December 2011.

[16] D. Chen and H. Zhao, "Data security and privacy protection issues in cloud computing," in *Proceeding of the International Conference on Computer Science and Electronics Engineering (ICCSEE '12)*, vol. 1, pp. 647–651, Hangzhou, China, March 2012.

[17] K. D. Bowers, A. Juels, and A. Oprea, "Proofs of retrievability: theory and implementation," in *Proceedings of the ACM Workshop on Cloud Computing Security (CCSW '09)*, pp. 43–53, November 2009.

[18] K. D. Bowers, A. Juels, and A. Oprea, "HAIL: a high-availability and integrity layer for cloud storage," in *Proceedings of the 16th ACM conference on Computer and Communications Security*, pp. 187–198, ACM, Chicago, Ill, USA, November 2009.

[19] J. Schiffman, T. Moyer, H. Vijayakumar, T. Jaeger, and P. McDaniel, "Seeding clouds with trust anchors," in *Proceedings of the ACM workshop on Cloud computing security workshop (CCSW '10)*, pp. 43–46, ACM, October 2010.

[20] D. H. Rakesh, R. R. Bhavsar, and A. S. Thorve, "Data security over cloud," *International Journal of Computer Applications*, no. 5, pp. 11–14, 2012.

[21] R. L. Rivest, L. Adleman, and M. L. Dertouzos, "On data banks and privacy homomorphisms," *Foundations of Secure Computation*, vol. 4, no. 11, pp. 169–180, 1978.

[22] C. Gentry, *A fully homomorphic encryption scheme [Ph.D. thesis]*, Stanford University, 2009.

[23] D. Boneh, "The decision Diffie-Hellman problem," in *Algorithmic Number Theory*, vol. 1423, pp. 48–63, Springer, 1998.

[24] A. Kaur and M. Bhardwaj, "Hybrid encryption for cloud database security," *Journal of Engineering Science Technology*, vol. 2, pp. 737–741, 2012.

[25] R. Arora, A. Parashar, and C. C. I. Transforming, "Secure user data in cloud computing using encryption algorithms," *International Journal of Engineering Research and Applications*, vol. 3, no. 4, pp. 1922–1926, 2013.

[26] D. Manivannan and R. Sujarani, "Light weight and secure database encryption using tsfs algorithm," in *Proceedings of the International Conference on Computing Communication and Networking Technologies (ICCCNT '10)*, pp. 1–7, IEEE, 2010.

[27] F. Pagano and D. Pagano, "Using in-memory encrypted databases on the cloud," in *Proceedings of the 1st IEEE International Workshop on Securing Services on the Cloud (IWSSC '11)*, pp. 30–37, September 2011.

[28] K. Huang and R. Tso, "A commutative encryption scheme based on ElGamal encryption," in *Proceedings of the 3rd International Conference on Information Security and Intelligent Control (ISIC '12)*, pp. 156–159, IEEE, August 2012.

[29] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222–233, 2014.

[30] M. A. AlZain, B. Soh, and E. Pardede, "Mcdb: using multi-clouds to ensure security in cloud computing," in *Proceedings of the IEEE 9th International Conference on Dependable, Autonomic and Secure Computing (DASC '11)*, pp. 784–791, 2011.

[31] C. P. Ram and G. Sreenivaasan, "Security as a service (sass): securing user data by coprocessor and distributing the data," in *Proceedings of the 2nd International Conference on Trendz in Information Sciences and Computing, (TISC '10)*, pp. 152–155, IEEE, December 2010.

[32] M. Asad Arfeen, K. Pawlikowski, and A. Willig, "A framework for resource allocation strategies in cloud computing environment," in *Proceedings of the 35th Annual IEEE International Computer Software and Applications Conference Workshops (COMPSACW '11)*, pp. 261–266, July 2011.

[33] A. Rao, "Centralized database security in cloud," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 1, pp. 544–549, 2012.

[34] E. M. Mohamed, H. S. Abdelkader, and S. El-Etriby, "Enhanced data security model for cloud computing," in *Proceedings of the 8th International Conference on Informatics and Systems (INFOS '12)*, pp. CC-12–CC-17, IEEE, 2012.

[35] S. Biedermann and S. Katzenbeisser, "POSTER: event-based isolation of critical data in the cloud," in *Proceedings of the ACM SIGSAC Conference on Computer & Communications Security*, pp. 1383–1386, ACM, 2013.

[36] C. Delettre, K. Boudaoud, and M. Riveill, "Cloud computing, security and data concealment," in *Proceedings of the 16th IEEE Symposium on Computers and Communications (ISCC '11)*, pp. 424–431, Kerkyra, Greece, July 2011.

[37] Y. Tang, P. P. C. Lee, J. C. S. Lui, and R. Perlman, "Fade: secure overlay cloud storage with file assured deletion," in *Security and Privacy in Communication Networks*, pp. 380–397, Springer, New York, NY, USA, 2010.

[38] K. Benson, R. Dowsley, and H. Shacham, "Do you know where your cloud files are?" in *Proceedings of the 3rd ACM workshop on Cloud computing security workshop*, pp. 73–82, ACM, October 2011.

[39] P. Mahajan, S. Setty, S. Lee et al., "Depot: cloud storage with minimal trust," *ACM Transactions on Computer Systems*, vol. 29, no. 4, article 12, 2011.

[40] A. J. Feldman, W. P. Zeller, M. J. Freedman, and E. W. Felten, "SPORC: group collaboration using untrusted cloud resources," in *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation (OSDI '10)*, vol. 10, pp. 337–350, 2010.

[41] E. Pinheiro, W.-D. Weber, and L. A. Barroso, "Failure trends in a large disk drive population," in *Proceedings of the 5th USENIX conference on File and Storage Technologies (FAST '07)*, vol. 7, pp. 17–23.

[42] T. Tsai, N. Theera-Ampornpunt, and S. Bagchi, "A study of soft error consequences in hard disk drives," in *Proceeding of the 42nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN '12)*, pp. 1–8, Boston, Mass,USA, June 2012.

[43] J. Krumm, "A survey of computational location privacy," *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 391–399, 2009.

[44] S. Pearson and A. Benameur, "Privacy, security and trust issues arising from cloud computing," in *Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom '10)*, pp. 693–702, IEEE, December 2010.

[45] E. Stefanov, M. van Dijk, E. Shi et al., "Path oram: an extremely simple oblivious ram protocol," in *Proceedings of the ACM SIGSAC Conference on Computer & Communications Security*, pp. 299–310, ACM, 2013.

[46] S. Paquette, P. T. Jaeger, and S. C. Wilson, "Identifying the security risks associated with governmental use of cloud computing," *Government Information Quarterly*, vol. 27, no. 3, pp. 245–253, 2010.

[47] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 1–11, 2011.

[48] C. Cachin and M. Schunter, "A cloud you can trust," *IEEE Spectrum*, vol. 48, no. 12, pp. 28–51, 2011.

[49] J. Idziorek, M. Tannian, and D. Jacobson, "Attribution of Fraudulent Resource Consumption in the cloud," in *Proceedings of the IEEE 5th International Conference on Cloud Computing (CLOUD '12)*, pp. 99–106, June 2012.

[50] Z. Shen, L. Li, F. Yan, and X. Wu, "Cloud computing system based on trusted computing platform," in *Proceedings of the International Conference on Intelligent Computation Technology and Automation (ICICTA '10)*, vol. 1, pp. 942–945, IEEE, May 2010.

[51] R. Neisse, D. Holling, and A. Pretschner, "Implementing trust in cloud infrastructures," in *Proceedings of the 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid '11)*, pp. 524–533, IEEE Computer Society, May 2011.

[52] R. Yeluri, E. Castro-Leon, R. R. Harmon, and J. Greene, "Building trust and compliance in the cloud for services," in *Proceedings of the Annual SRII Global Conference (SRII '12)*, pp. 379–390, San Jose, Calif, USA, July 2012.

[53] R. Ranchal, B. Bhargava, L. B. Othmane et al., "Protection of identity information in cloud computing without trusted third party," in *Proceedings of the 29th IEEE Symposium on Reliable Distributed Systems (SRDS '10)*, pp. 368–372, November 2010.

[54] A. Squicciarini, S. Sundareswaran, and D. Lin, "Preventing information leakage from indexing in the cloud," in *Proceedings of the 3rd IEEE International Conference on Cloud Computing (CLOUD '10)*, pp. 188–195, July 2010.

[55] K. Hwang and D. Li, "Trusted cloud computing with secure resources and data coloring," *IEEE Internet Computing*, vol. 14, no. 5, pp. 14–22, 2010.