Privacy and Security in Cloud Environment: A Survey of Recent Development

Amal Alsubaih, Alaaeldin Hafez, Khaled Alghathbar
Information System Department, King Saud University, Riyadh, Saudi Arabia

Keywords:

Cryptography
Authorization
Authentication
Cloud Environment
Privacy
Security

ABSTRACT

Recently, security and privacy issues have become barriers to the adoption of cloud computing. The main risks associated with cloud environments are the loss of control, lack of trust, and multi-tenancy. They raise serious concerns over privacy and security because the data and resources are located out of the users' boundary; the access control policies are enforced by a semi-trusted cloud; and the resources are shared between tenants. Because of the present issues of privacy and security, many studies have introduced solutions to protect the data stored in a cloud and prevent privacy violations. In this paper, we analyze the recent work on security and privacy in a cloud environment based on the related outlines provided by the Cloud Security Alliance (CSA); and provide an overview of the main mechanisms that have been used. This overview shows the existing research trends and presents open issues that offer a foundation for additional research in the cloud field.

1. INTRODUCTION

Cloud computing provides several benefits to the user, such as flexible and scalable on-demand services at a reduced cost [1]. While there are many benefits to adopting cloud computing, there are also some challenges and risks that accompany its adoption. Some of the biggest challenges facing cloud computing are privacy and security issues. Storing, processing, and sharing sensitive data like personal, financial, and medical data in an untrusted cloud could lead to privacy violations, primarily the disclosure of sensitive data by cloud service providers or external attackers. Moreover, the loss of control raises serious concerns over privacy and security because the data owner is unaware of the location of his/her data and the operations applied on his/her data in the cloud. Furthermore, unauthorized access to the stored data as a result of the weakness of the access control mechanism represents a serious threat to data confidentiality [2].

Currently, numerous cloud service providers have privacy and security problems that need to be addressed [3]. Borgmann et al. [4] studied several cloud storage service providers (CloudMe, Wuala, CrashPlan, Dropbox, Mozy, TeamDrive, and Ubuntu One), none of which were able to sufficiently meet all of the security requirements. Several vulnerabilities were found, including weak authentication, shared files being exposed using search engines, and data being stored without encryption or using only cloud side encryption, which did not prevent the disclosure of the sensitive data by the cloud. Similarly, Amazon S3 provides only cloud side encryption for the stored data, which does not protect the data confidentiality from the cloud provider [5]. Hadoop is a powerful framework that supports the processing and analysis of large data sets in a cloud computing environment using the MapReduce programming model [6]. Nevertheless, it has several problems related to security and privacy, one of which is a lack of data encryption when transferring and storing the data [7,8]. On the other hand, the research on data privacy in cloud computing is evolving over time. Various solutions have been proposed to preserve security and privacy in a cloud environment. Those solutions are usually based on concepts like cryptography, authorization, and authentication to prevent privacy violations and the disclosure of sensitive data by cloud service providers or external attackers.

The purpose of this paper is to enhance the security and privacy research in a cloud environment by providing an overview of the recent studies on this field. We aim to analyze the mechanisms used by the reviewed solutions to protect sensitive data outsourced to the cloud. This overview will show the existing research trends and technologies and will present open issues that offer a foundation for additional research in the cloud field. The remainder of this chapter is organized as follows: section 2 presents the research

method, and section 3 reviews some recent approaches proposed for preserving security and privacy in a cloud environment, along with the mechanisms for these approaches. Finally, section 4 concludes the paper.

2. RESEARCH METHOD

The existing privacy preserving solutions in a cloud environment are analyzed based on the related outlines provided by the Cloud Security Alliance (CSA). CSA is an organization that has specifically targeted the security and privacy issues in a cloud environment. It defines standards, outlines best practices, and provides guidelines to increase the security and enhance the privacy within a cloud environment. The CSA guidelines cover three aspects, the cloud architecture, governance, and operation, and address fourteen critical domains related to security in a cloud environment [9,10]. We focus on two core domains, the encryption and key management domain and the identity and access management domain. We choose these domains because they are essential to protect sensitive data within the cloud from being disclosed or leaked. From the requirements for these domains, we identify several questions and base our analysis on these to ensure that the current solutions for preserving security and privacy in a cloud environment satisfy the CSA requirements. Table 1 shows the related questions for each domain. We review twenty-three of the recent security and privacy preserving approaches for a cloud environment and map them into two main categories: the encryption and key management domain and the identity and access management domain. Then, we answer the related research questions for each approach and explore the main mechanisms that they depend on in order to protect the sensitive data stored within the cloud from security and privacy violations.

Table 1. Cloud security and privacy requirement by CSA and the related research questions.

CSA Domain	Requirement by CSA	Related Questions
Encryption and Key Management	1- Encryption: Policies and procedures shall be established and mechanisms implemented for encrypting sensitive data in storage and data in transmission. 2- Key Management: Policies and procedures shall be established and mechanisms implemented for effective key management to support encryption of data in storage and in transmission.	1. What is the encryption method used in the solution? 2. What is the key management followed by the approach? 3. What are the stages of the data life cycle covered by the protection?
Identity and Access Management	1- User Access Restriction / Authorization: Normal and privileged user access to sensitive data shall be restricted and approved by management prior to access being granted. 2- User Access Revocation: Revocation or modification of user access to the information assets and data shall be implemented upon any change.	4. What is the authorization mechanism used to grant the access? 5. Does the approach implement the user access revocation and modification efficiently? 6. Does the approach release sensitive information about data or policy?

3. RESULTS AND ANALYSIS

This section reviews twenty-three security and privacy preserving approaches for a cloud environment and maps them into two main categories: the encryption and key management domain and the identity and access management domain. Then, it explores the main mechanisms that the approaches depend on in order to protect the sensitive data stored within the cloud from privacy violations. Further, the review is broadened to include numerous issues in accordance with the CSA procedures. Table 2 shows a summary of the analysis results for the recent privacy and security developments in a cloud environment. Each column in Table 2 corresponds to one of the questions in Table 1.

3.1. Encryption and Key Management Approaches Analysis

Encryption is a technique that secures and converts data into a form that an unauthorized user cannot read. Recently, many researchers have suggested encrypting the data before outsourcing them into the cloud to ensure data privacy. Most of the reviewed studies have followed this suggestion and used encryption at the client side to protect the confidentiality of the data stored within the cloud. Nevertheless, some works [15,16,17,18,23] have not used encryption because they either targeted a private cloud or considered a public cloud to be a trusted party. Several works introduced a cloud service for data protection using encryption. The work in [12] presented Privacy as a Service (PasS) for data stored and processed in a cloud depending on the use of a cryptographic co-processor that offered a secure and trusted domain for storing and processing the private data. In [13], a cloud service was introduced called Confidentiality as a Service (CaaS), which was dedicated to protecting the confidentiality of the data stored in the cloud using two commutative layers of encryption. The first layer was added locally to the data at the data owner side. The data were then sent to

CaaS to add the second layer of encryption. In [32], a solution was presented that involved the selection of a suitable encryption algorithm for each set of data to satisfy the privacy demands of users with minimum system overhead. However, it did not protect the data from the cloud.

This discussion on encryption and key management domain focuses on certain areas (i.e. to answer the study questions Q1, Q2, Q3), including the type of encryption used by the approach, the key management that was followed by the approach, and the data life cycle that was covered. With regard to the type of encryption, they used several techniques, including symmetric key cryptography, asymmetric key cryptography, ciphertext policy attribute-based encryption, key policy attribute-based encryption, proxy re-encryption, and xml encryption .The majority of the reviewed approaches used symmetric key cryptography to encrypt the data and then used another type of cryptography to encrypt the key. In [14,20,21,22,25], the symmetric key was encrypted by attribute-based encryption, while in [12,19], the symmetric key was encrypted by asymmetric key cryptography. Moreover, the approaches in [20,21,22] used an attribute-based encryption and a proxy re-encryption. However, they combined them differently. The first encryption was used to enforce fine-grained access control, and the second was used to prevent a user whose permissions are revoked from accessing the data in the future.

The key generation processes in the reviewed studies were performed by a third trusted party [12,24,25,29] or by the data owner [14,19,20,21,22,26]. The key deriving method rarely appeared in the reviewed studies; only four methods were presented: the Extirpation-based Key Derivation Algorithm (EKDA) in [33], HMAC-based Key Derivation function in [13], key distribution and management algorithm in [27], and Attribute-Based Hierarchical Key Updating (AB-HKU) in [31]. Moreover, some studies did not indicate how to manage the keys, while other studies relied on pre-defined attribute-based encryption methods to generate the key.

Regarding the CSA requirement to protect sensitive data at rest and data in transmission, most of the reviewed works satisfied this requirement and protected the data privacy during those stages of the data life cycle using cryptography. On the other hand, some works did not protect the privacy of the data at rest or/and at transmission, as in [15,16,17,18,23].

3.2. Identity and Access Management Approaches Analysis

Identity and access management (IAM) allows the authorized users to access the right resources at specific times. Under IAM, we focus on authorization, which is the act of giving a user the needed permissions to access data. With regard to authorization in a cloud environment, some research solutions have considered the cloud service providers to be fully trusted, such as in [15,16,18], while other solutions have considered them to be semi trusted, honest-but-curious services, such as in [19,20,21,22]. The different considerations had an impact on the mechanisms introduced by the researchers.The former solutions protect the data from unauthorized users, whereas the latter protect the data from unauthorized users and the cloud servers.The discussion about authorization focuses on certain areas (i.e. to answer the study questions Q4,Q5,Q6), including the authorization mechanism used to grant access to the data, whether the approach implements user access revocation and permission modification efficiently , and whether the approach releases sensitive information about the data or policy.

Generally, in the reviewed solutions, authorization was performed by using one of these methods: using a trusted party as in [11,12,17,26,23,24,29], or cryptographic enforced access control as in [14,20,21,22,28,29,31]. The work in [25] is an example of using a trusted party; it introduced a service in a trusted cloud to manage the privacy preserving access control to other cloud services that contain the data. This service was called Permission as a Service (PaaS). PaaS permits the data owner to assign and update permissions for all his/her data located at different clouds in a single place. When the user requests access to specific data, the cloud redirects him to the permission trusted cloud in order to validate the request in the permissions list and then generate the decryption key for the user. Moreover, the basic idea behind cryptographic enforced access control is using attribute-based cryptography as an enforcement mechanism for authorization; they combine the access control policies with the encryption to protect the data.

With regard to the authorization mechanism used to grant access, different kinds of mechanisms have been used in recent developments for cloud including: attributes-based access control (ABAC) in [14,15,20,21,22,25,28,29,31], role-based access control (RBAC) in [16,17,18], a capability list in [30], access control list (ACLs) in [13,24,26,27,33], and guaranty-based access control in (GBAC) [23].

Our findings on user access revocation and permission modification are as follows: eleven solutions did not indicate how to implement this operation [11,12,14,15,16,17,18,24,28,29,32], the solutions in [22] and [21] presented efficient user revocation processes. However the former introduced another issue when the user re-joined with different access privileges, while the later revoked the user's rights automatically after a predetermined time. Thus, it is not suitable for an environment where it should be possible to revoke a user's rights at any time. Moreover, the rest of the solutions had a large computational overhead for either the

data owner or/and the cloud for data re-encryption and/or key re-generating and distributing after revoking a user's rights or changing his\her permissions.

Although the reviewed approaches that rely on encryption provide data confidentiality and fine-grained access control in a cloud, they rarely provide sufficient protection for the access control policy because they leak information about the data and user attributes; only five of the reviewed approaches were capable of providing privacy protection for the access control policy [23, 26, 27, 31, 33].

Table 2. Summery of the analysis results on the recent privacy and security development in cloud environment. [R= protect data at rest, T= protect data at transmissions, y= yes , n= no , N/A= access control policy are not developed].

Ref	Solution Description	Encryption Method	Key Management	R	T	Authorization Mechanism	User or Permission Revocation	Releasing Sensitive Information about Policy to Cloud
[11]	Privacy manager that provides data obfuscation data And preferences	Obfuscation data	Not indicated	y	y	Not indicated	Not indicated	N/A
[12]	privacy as a Service (PasS) using cryptographic co-processor	Symmetric Encryption for data Asymmetric Encryption for the key	Trusted Party Generates Key	y	y	Not indicated	Not indicated	N/A
[13]	Confidentiality as a Service (CaaS) using two commutative layers of encryption	Symmetric Encryption	HMAC-based Key Derivation function *(HKDF*) in cloud side to generate key for CaaS service only	y	y	ACLs	Remove user from the list and require data re-encryption and key re-distribution	Yes
[14]	Client-side architecture for a cryptographic storage service in the cloud.	Attribute-based Encryption Symmetric Encryption	No details By data owner	y	y	ABAC	Not indicated	Yes
[15]	Privacy preserving authorisation through sticky privacy policies	No Encryption	Not indicated	n	n	ABAC	Not indicated	Yes
[16]	Multi-tenancy authorization using role-based access control and path-based object hierarchy	No Encryption	Not indicated	n	n	RBAC	Not indicated	Yes
[17]	Combining Role-based Access Control with Attributes	No Encryption	Not indicated	n	n	ARBAC	Not indicated	Yes
[18]	It provides four privacy and security methods	No Encryption	Not indicated	n	n	RBAC	Not indicated	Yes
[19]	Plutus: Files group symmetric keys with lazy re-encryption	Symmetric Encryption	No details By data owner	y	y	Files group	Data re-encryption and key re-generation and re-distribution	N/A
[20]	It achieves secure, scalable, and fine-grained data access control using attribute-based encryption proxy re-encryption	Key Policy Attribute-based Encryption Proxy Re-Encryption	Generate key methods in attribute-based encryption and proxy re-encryption	y	y	ABAC	Data re-encryption on user revocation, key re-generation on revoking user's access privileges	Yes
[21]	Generic approach for attribute-based encryption	Attribute-based Encryption Proxy Re-Encryption	Generate key methods in attribute-based encryption and proxy re-encryption	y	y	ABAC	Removing the user from the list, Problem in re-join with different access privilege	Yes

[22]	It combines time with attribute-based encryption and proxy re-encryption	Cipher Policy Attribute-based Encryption Proxy Re-Encryption	Generate key methods in attribute-based encryption and proxy re-encryption	y	y	ABAC	After predefined time the user revoked Key re-issuing on permissions update	Yes
[23]	Persona approach that relies on identity-based cryptography And guarantee-based access control (GBAC)	No Encryption	Not indicated	n	n	Guarantee-based access control GBAC	Re-issuing key for unrevoked users	No
[24]	It relies on trusted platform module, eXtensible access control markup language (XACML)	XML-Encryption	Trusted Party	y	y	ACLs	Not indicated	Yes
[25]	Permission as a Service (PaaS) dedicated for authorization	Attribute based Encryption	Trusted Party Generates Key	y	y	ABAC	Costly key generating at each access a new key is issued	Yes
[26]	Dynamic cryptographic accumulators	Not indicated	Not indicated	y	y	ACLs	Updating the witness to all authorized users when permissions changed	No
[27]	It is two layer access control approach	Asymmetric Encryption and Selective Encryption	key distribution and management algorithms	y	y	ACLs	Re-encryption the data on user revocation	No
[28]	Using Attribute-Based Encryption, Attribute Based Signature depending on trusted party	Attribute-Based Encryption	Trusted Party Generates Key	y	y	ABAC	Not indicated	Yes
[29]	Using Attribute-Based Encryption, Attribute Based Signature depending on trusted party	Attribute-Based Encryption	Trusted Party Generates Key	y	y	ABAC	Not indicated	Yes
[30]	Each user associated with a set of authorized data file to represent the access control rather than ACLs	Asymmetric Encryption	Not indicated	y	y	Capability List	Data re-encryption and key re-generation and re-distribution	Yes
[31]	K2C: it provides anonymous access for the end user through using signatures of KP-ABE	Key-Policy Attribute-based Encryption	Attribute Based Hierarchical Key Updating AB-HKU	y	y	ABAC	Data re-encryption and key updated on users and permissions revocation	No
[32]	EPPS: chooses suitable encryptions algorithm for each data that satisfies the user privacy demands with the minimum system overhead	Symmetric Encryption	Not indicated	y	y	Not indicated	Not indicated	N/A
[33]	Privacy preserving framework for cloud storage using encryption.	Asymmetric, Symmetric Encryptions	Extirpation-based Key Derivation Algorithm (EKDA)	y	y	ACls	Data re-encryption and key updated on users and permissions revocation	No

3.3. Results and Recommendations:

In this survey, we reviewed the recent developments in security and privacy protection in a cloud environment, the summery of the study is shown on table 3. Although considerable progress has been made in protecting the data stored in a cloud, many important issues remain. We identified several issues and recommendations to form future research directions, including:

- Some of the reviewed authorization approaches protected the data outsourced into the cloud against external attackers only; however, the public cloud is semi-trusted and attacks can come from both inside and outside the cloud.
- Privacy aware authorization that provides privacy protection for both the data and access control policy is needed.
- Flexible authorization approaches that are capable of resolving access conflicts and following restriction rules in a privacy aware mode were missing from the reviewed solutions.
- Although the approaches provide data confidentiality and fine-grained access control, they lack efficient user revocation methods.
- More work should be done in key management.
- Role-based Access Control (RBAC) is flexible access control mechanism suitable for cloud environment [34], however only few works introduce it in their solutions.

Table 3. Summery of study results

Q	Area	Study Answer
1	The encryption method used in the solutions	Symmetric key cryptography, Asymmetric key cryptography, Ciphertext policy attribute-based encryption, Key policy attribute-based encryption, Proxy re-encryption, Xml encryption.
2	The key management used in the solutions	Extirpation-based Key Derivation Algorithm (EKDA), HMAC-based Key Derivation function, key distribution and management algorithm, Attribute-Based Hierarchical Key Updating (AB-HKU), Pre-defined attribute-based encryption methods to generate the key, Rely on trusted party.
3	Protection the stages of the data life cycle	Most of the reviewed works protect the privacy of the data at rest and at transmission except works in [15,16,17,18,23].
4	The authorization mechanism used in the solutions	Attributes-based Access Control (ABAC), Role-based Access Control (RBAC), Capability list, Access Control List (ACLs), Guaranty-based access control in (GBAC).
5	Implementing the user access revocation and modification efficiently	Most of the reviewed works did not implement the user access revocation and modification efficiently. They have computational overhead for either the data owner or/and the cloud.
6	Releasing sensitive information about data or policy	Most of the reviewed works release sensitive information about policy except works in **[**23, 26, 27, 31, 33].

4. CONCLUSION

The research on data privacy in cloud computing is evolving over time. Various solutions have been proposed to preserve security and privacy in a cloud environment. These solutions are usually based on concepts like cryptography and authorization. We identified and analyzed the main mechanisms used to reduce the privacy and security risks in these areas. Although considerable progress has been made in protecting the data stored in a cloud, many important issues remain, such as flexible and privacy aware authorization with efficient user revocation.

REFERENCES

[1] Q. Zhang, *et al.*, "Cloud computing: state-of-the-art and research challenges." *Journal of Internet Services and Applications* 1.1 (2010): 7-18

[2] S. Pearson, "Taking account of privacy when designing cloud computing services." *Software Engineering Challenges of Cloud Computing*, 2009. CLOUD'09. ICSE Workshop on. IEEE, 2009.

[3] V. Delgado, ".Exploring the limits of cloud computing ".Master's Thesis 2010 ,.

[4] M. Borgmann, *et al.* , "On the Security of Cloud Storage Services ".SIT Technical Reports , 2012.

[5] http//:aws.amazon.com/s3/#protecting [Accessed on: March 2013].

[6] http://hadoop.apache.org/ [accessed on : May 2013]

[7] C. Basescu, *et al.,* "Managing data access on clouds: A generic framework for enforcing security policies." *Advanced Information Networking and Applications (AINA),* IEEE, 2011.
[8] Intel Group, "Fast, Low-Overhead Encryption for Apache Hadoop." white paper,2013. [accessed on : May 2013] https://hadoop.intel.com/pdfs/IntelEncryptionforHadoopSolutionBrief.pdf
[9] Cloud Security Alliance ,"Security Guidance for Critical Areas of Focus in Cloud Computing V3 2011 ," https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf [Accessed on: January 2013].
[10] Cloud Security Alliance, "Domain 12 Guidance for Identity & Access Management", 2010 https://cloudsecurityalliance.org/guidance/csaguide-dom12-v2.10.pdf .[accessed on: March 2013.]
[11] S. Pearson *et al.,* "A privacy manager for cloud computing." *Cloud Computing. Springer Berlin Heidelberg*, 2009.
[12] W. Itani, *et al.,*" Privacy as a service: Privacy-aware data storage and processing in cloud computing architectures."*Dependable, Autonomic and Secure Computing. 8th IEEE International Conference on.* IEEE, 2009.
[13] S. Fahl, Sascha, *et al.* "Confidentiality as a Service--Usable Security for the Cloud." Trust, *Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference* on. IEEE, 2012.
[14] S. Kamara *et al.,*" Cryptographic cloud storage ".*Financial Cryptography and Data Security*, 2010.
[15] D. Chadwick, *et al.,* "A privacy preserving authorisation system for the cloud." *Journal of Computer and System Sciences* 78.5 2012, 1359-1373.
[16] J. Calero, *et al.,* "Toward a multi-tenancy authorization system for cloud services*." Security & Privacy*, IEEE 8.6 2010, 48-55.
[17] E. Mon, *et al.,* "The privacy-aware access control system using attribute-and role-based access control in private cloud." *Broadband Network and Multimedia Technology (IC-BNMT*), 2011 4th IEEE International Conference on. IEEE, 2011.
[18] Z. Wang, "Security and privacy issues within the Cloud Computing."*Computational and Information Sciences (ICCIS), 2011 International Conference on*. IEEE, 2011.
[19] M. Kallahalla, *et al*., "Plutus: Scalable secure file sharing on untrusted storage." *Proceedings of the 2nd USENIX Conference on File and Storage Technologies*. 2003.
[20] S. Yu, et al., "Achieving secure, scalable, and fine-grained data access control in cloud computing." *INFOCOM, Proceedings IEEE*. IEEE, 2010.
[21] Y. Yang, *et al.,* "A generic scheme for secure data sharing in cloud." *Parallel Processing Workshops (ICPPW), 2011 40th International Conference on. IEEE*, 2011.
[22] Q. Liu, *et al.,* "Time-Based Proxy Re-encryption Scheme for Secure Data Sharing in a Cloud Environment." *Information Sciences*, 2012.
[23] M. Hussain, "The design and applications of a privacy-preserving identity and trust-management system.", 2010, PHD thesis.
[24] U. Greveler, *et al.,* "A Privacy Preserving System for Cloud Computing." *Computer and Information Technology (CIT),* 2011 IEEE 11th International Conference on. IEEE, 2011.
[25] V. Echeverría, *et al.,* "Permission management system: permission as a service in cloud computing." *Computer Software and Applications Conference Workshops*, 2010 IEEE 34th Annual. IEEE, 2010.
[26] D. Slamanig, "Dynamic Accumulator based Discretionary Access Control for Outsourced Storage with Unlinkable Access." *Financial Cryptography and Data Security*. Springer Berlin Heidelberg, 2012.
[27] M. Raykova, *et al.,* "Privacy enhanced access control for outsourced data sharing." *Financial cryptography and data security*. Springer Berlin Heidelberg, 2012. 223-238.
[28] F. Zhao, *et al.,* "Realizing fine-grained and flexible access control to outsourced data with attribute-based cryptosystems." *Information Security Practice and Experience.* Springer Berlin Heidelberg, 2011. 83-97.
[29] S. Ruj, Sushmita, *et al.,* "Privacy preserving access control with authentication for securing data in clouds." *Cluster, Cloud and Grid Computing (CCGrid)*, 2012 12th IEEE/ACM International Symposium on. IEEE, 2012.
[30] C. Hota, *et al.,* "Capability-based Cryptographic Data Access Control in Cloud Computing." *Int. J. Advanced Networking and Applications* 3, no. 3 2011.
[31] S. Zarandioon, *et al.,* "K2C: Cryptographic cloud storage with lazy revocation and anonymous access."*Security and Privacy in Communication Networks. Springer Berlin Heidelberg*, 2012. 59-76.
[32] I. Chuang, et al., "An effective privacy protection scheme for cloud computing." *Advanced Communication Technology (ICACT*), 2011 13th International Conference on. IEEE, 2011.
[33] H. RuWei, *et al.,* "Study of privacy-preserving framework for cloud storage." *Computer Science and Information Systems* 8, no. 3 , 2011, : 801-819.
[34] H. Takabi, *et al.,* "Security and privacy challenges in cloud computing environments." *Security & Privacy*, IEEE 8, no. 6 , 2010,: 24-31.