International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307

A Unified Model for Cloud Data Confidentiality

[1]A.Vithya Vijayalakshmi, [2]N. Veeraragavan, [3]Dr. L. Arockiam
[1]Ph.D. Scholar, [2]Ph.D. Scholar, [3]Associate Professor,
Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli, Tamilnadu, India

ABSTRACT

Cloud computing is a developing technology which gets more attention from both the industries and academia. The cloud storage is one of the main benefit in cloud computing, which is particularly attractive for the users who needs unpredictable storage for their enterprises and so on. Minimum storage and processing cost is an obligatory requirement of all organization and industries, while analysis of data and information is mandatory in all organization. Although there is a reduction in cloud storage cost, customers has to face more technical and security problems such as data integrity, confidentiality and availability. If there is no confidentiality, then, there is no guarantee for the data on cloud. Many researches have been proposed number of techniques for data security in the cloud. However, there are still many issues in cloud data storage. The optimum solution to ensure the confidentiality in cloud storage is to encrypt the data whereas encryption alone fail to give high security to the data in the cloud storage. To give maximum protection to cloud storage, this paper proposes a unified combined model of both encryption and obfuscation. Encryption is the process of converting original text and Obfuscation is a process of encrypting numerical type of data, the researchers proved that by combining these two data protection techniques, the data will be more protective on cloud storage.

Keywords :- Cloud Storage, Database Management, Data Confidentiality, Encryption, Obfuscation

I. INTRODUCTION

Cloud computing is a recent trend in IT that moves computing and data away from desktop and portable PCs into large data centers. It refers to applications delivered as services over the Internet [1]. The definition of cloud computing provided by National Institute of Standards and Technology (NIST) says that: Cloud computing is a model for enabling 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 [2]. In cloud computing, there are basically three service models i.e. Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [3]. The main task of cloud computing is to make use of distributed resources in a better manner. Cloud computing deals with virtualization, scalability, interoperability, quality of service and the delivery models of the cloud, namely private, public and hybrid [2].

This paper is formatted in the following way: section II gives an overall idea of cloud data storage, section III discusses related work of this paper work, section IV discusses the issues in cloud data storage, section V describes proposed framework and its working steps, section VI describes the benefits of the proposed model and section VII describes the conclusion and future aspects related to this paper work.

II. CLOUD DATA STORAGE

Cloud data storage is growing very fast now-a-days because it can be available from various sources and the clients does not have to worry about maintaining, because the service provider will take all the responsibility [4]. Cloud Service Providers maintains computing resources and data automatically via software [5]. Security is very important issue in cloud computing and data security is one of the top most challenge. Whenever the client outsource their data to the cloud there may be data loss or leakage. Even though cloud service provider is honest but there may be internal threat. Any malicious user can harm

sensitive data because service providers can access to that data [4]. The four key components of data security in cloud computing are data availability, data integrity, data confidentiality, and data authentication [6]. Data availability refers to use the data in time whenever it's needed and also refers to the availability of cloud service provider on-demand. Data Integrity checks whether the information stored in the cloud gets modified or not. Data Confidentiality refers that only the authorized users are accessing the data. Data Authentication refers to the process of verifying whether the incoming user is authorized or not [14]. This paper focuses on data confidentiality in cloud storage. It is one of the major issues in data security. It allows only the authorized person can view the data. Whenever the confidentiality is less in the cloud storage, the data is not secure any more.

A. Cloud Database Management

A Cloud database management system (CDBMS) is a distributed database where the computing are delivered as a service. The main goal is to share the resources, software and information between multiply devices over a network which is mostly the internet. An example of this is Software as a Service, or SaaS, which is an application that is delivered through the web portals to customers. Cloud applications generally connects to a database that is being run on the cloud and have varying degrees of efficiency. Some are manually configured, some are pre-configured and some are native. Native cloud databases are traditionally better equipped and more stable, because that those that are modified to adapt the cloud [7].

Despite the benefits offered by cloud-based DBMS, many people still have problems. This is most likely due to the various security issues that have yet to be deal with. These security issues mainly focuses that cloud DBMS are hard to monitor since they often move across multiple hardware stacks and / or servers. Security becomes a serious issue with cloud DBMS when there's multiple virtual machines that might be able to access a database without being noticed or setting off any alerts [7]. In this situation a malicious person could potentially access the sensitive data or cause serious damage to the integral structure of a database, putting the entire system in danger.

B. Data Confidentiality

Data security covers the confidentiality and integrity of data, typically during transmission and storage. Data confidentiality is an important need for keeping information secret, maintaining personnel records and sensitive data. NIST defined confidentiality as "Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information" [8]. Data confidentiality is needed when the message sent contains sensitive material that should not be read by others and therefore must not be sent in an apprehensible format. To make comprehensible format, the users encrypt the data with traditional encryption technology before uploading the data. A key component to prevent information confidentiality would be encryption. Encryption ensures that only the right people (people who knows the key) can read the information. Encryption is ubiquitous in today's environment and can be found in almost every major protocol.

This paper proposes a unified model for cloud data confidentiality by using encryption and obfuscation technique [14]. Normally, data confidentiality is protected by various encryption technique, but it is not sufficient for the data protection in the cloud storage. In this model, both security algorithms and obfuscation are combined and used for giving secured cloud storage. Here files are encrypted with AES algorithm. Encryption is unified with obfuscation technique to protect the data from both insider attackers and outsider attackers. In the proposed technique, the data owner has the fill control over the data, he should encrypt and obfuscate the data before sending it to the cloud storage.

III. RELATED WORKS

Atiq ur Rehman et al. [13] introduced a framework model to ensure the data confidentiality in the cloud database. The data in the cloud are stored with a combination of encryption and obfuscation technique using metadata repository. According to the type of data in the table, only the sensitive character based columns are encrypted using data encryption standard algorithm and sensitive numerical based columns are obfuscated based on the metadata repository. The non-sensitive data columns are left unencrypted. Encryption

and obfuscation will be done only on the client side. The metadata repository available on the client side will have the privacy related information. Metadata repository is designed on client side and it will store encryption / decryption keys and obfuscation / de-obfuscation information of required data. This model concentrated on to query over encrypted plus obfuscated data.

Monikandan et al. [14] proposed a technique for efficient data confidentiality in cloud storage, they use both hybrid symmetric encryption algorithm and obfuscation techniques. This technique, first find out the type of data, based on the type of the data encryption or obfuscation is applied. If the data are alphabets or alphanumeric encryption is applied on the data, if the data are digits, then obfuscation is applied. Encryption technique uses the symmetric encryption algorithm, which includes both substitution and transposition to convert the original text into cipher text. Obfuscation is a technique applied through mathematical functions.

Varun Maheshwari et al. [15] introduced a texture scheme for representing the characters. To enhance the privacy, the textures are divided and noise is added to each of the portions such that all portions of a texture needs to be gathered to recover the original data. To overcome the data confidentiality, the main task of this scheme is to convert the data into images and then the images are split into pieces, each part of the image are send to different independent clouds. The system mainly consists of cloud service providers and users. This technique is applied to the data before uploading it into the cloud storage. When retrieving the data, the data are retrieved by matching between two images, normalized cross-correlation (NCC) is used to make a match. Aggregating the images, is one of the data obfuscation technique. Here, each part of the data stores on different clouds and also the data itself is highly noisy, this gives high security to the data. Malicious attackers cannot get the information from that data even if they finds the other parts because the data is highly noisy.

Yongkang Fu et al. [16] proposed a security scheme for the confidentiality of the data, recovery of lost and error data. The scheme is based on symmetric encryption technology and erasure codes. To ensure the data confidentiality, the user encrypt the data file names by using symmetric encryption techniques. Boot password technique is used to generate the encrypted key. It consists of two parts, one is memory password (input by the user) and the other is encrypt file name. Both added together and produced a boot password. Different boot passwords can be produced for different files, with the same memory password the user could generate different encrypted keys.

IV. ISSUES IN CLOUD DATA STORAGE

A. Trust

Trust is one of the important issue in cloud computing. Whenever the data are outsourced, the trust is based on the integrity of the data. Even though cloud provides major security, there may be malicious insiders who make use of the cloud data.

B. Ownership

The data owner has no complete control over the data, once the data has been submitted to the cloud. They unable to protect the data from the outsider attackers and the insiders. Many cloud providers address this issue with well-skilled user-sided agreements [10].

C. Security

Security is an important issue in cloud storage. When the data are outsourced, there will be less security, because there is no privacy and protection. Once the data has been sent to the third party, there is less protection to the data from the attackers. There are two type of attacker, one is insider attack and the other is outsider attack. Insider attack can't not be identified easily. So there is less security in storing the data in cloud storage.

D. Privacy

As all data can be accessed from any location, there is no privacy in cloud data storage. There may be data loss, data damage or data modified. On the other hand, users may crack hidden information when they accessing cloud computing services [10].

E. Data Loss and Leakage

Enterprises are more concerned about data loss and leakage. The data attack is much greater in the cloud. There are many ways data may be damaged in the cloud including data deletion, data modified, changing an encryption key and unauthorized access by insiders

or other cloud users. Encryption of sensitive data reduces the exposure of data loss and leakage [11].

V. PROPOSED FRAMEWORK

Security of the data in cloud storage is a critical act to both users and providers. This paper discusses the existing data confidentiality techniques. Normally, Confidentiality is ensured by encryption algorithm. Figure.1 represents a unified model for cloud data confidentiality using encryption and obfuscation technique. In this model, both encryption and obfuscation techniques are combined to secure the data stores in the cloud database. Based on the type of the data, this techniques can be applied. All the data must be encrypted and obfuscated before it is sent to the cloud database. Here, Advanced Standard Encryption algorithm is used for encrypting character based data, then it is converted to ASCII values and obfuscation technique is applied to the transformed encrypted values. The numerical values are obfuscated. Once the data is applied by proposed confidentiality technique at the client side, then the data is submitted to the cloud storage. Here, the data owner has the complete control over the data.

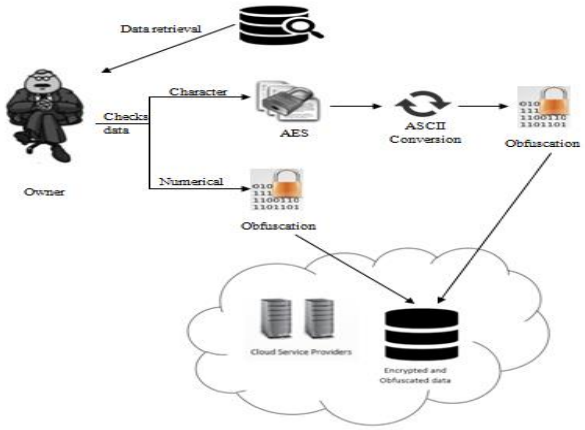

Figure 1 : A Unified Model for Cloud Data Confidentiality

A. TECHNIQUES USED IN THIS MODEL

- *Advanced Standard Encryption:*

It is an iterated symmetric block cipher, which means that works by repeating the same defined steps multiple times. It is a secret key encryption algorithm, operates on a fixed number of bytes. This key is expanded into individual sub keys, a sub keys for each operation round. This process is called key expansion. AES uses a key (cipher key) whose length is 128, 192 or 256 bits and 10, 12, or 14 round keys, respectively.

- *Obfuscation:*

Obfuscation is a form of data masking where data is scrambled to prevent unauthorized access to sensitive materials. This form of encryption results in unclear or confusing data. Obfuscation techniques are: Name obfuscation, Data obfuscation, Code flow obfuscation, Incremental obfuscation, Intermediate code optimization, Debug information obfuscation, Watermarking, Source code obfuscation [9].

- *Data Obfuscation (DO)*

Aimed at obscuring data. Types of data obfuscation are: Storage, Aggregation, Ordering and Encoding. Encoding uses mathematical functions to obfuscate the data. This model uses sqrt(), round() and user created equation y=x+x for obfuscation and for de-obfuscation pow() and user created equation x = y/2.

- *ASCII Conversion*

ASCII standards has been adopted by several American computer manufactures as their computer's internal code. The character based encrypted data are converted into ASCII values.

B. ALGORITHM

Step 1: Retrieval of data from repository
Step 2: Checks data type of the data
Step 2.1: if character
Step 2.1.1: perform AES encryption
Step 2.1.2: perform ASCII conversion
Step 2.1.3: perform of obfuscation
Step 2.2: if numerical
Step 2.2.1: perform of obfuscation
Step 3: Storage of obfuscation resultant data in cloud

This algorithm describes that the data owner will checks the data type of the data which has to be stored in the cloud database.

For encrypting: if the data type of the data is alphabet or alpha-numerical then it performs AES encryption then the encrypted text performs ASCII conversion at last obfuscation is applied. Here, sqrt( ), round( ), y=x+x are used to obfuscate the character based data. If the data type is numerical then obfuscation is applied, cbrt( ) is used to obfuscate the numerical values.

For decrypting: character based encrypted data are de-obfuscated using x = y / 2, pow( ), floor( ).

C. ALGORITHM WORKFLOW

The workflow of the proposed technique is given in Figure 2 followed by pseudo code.

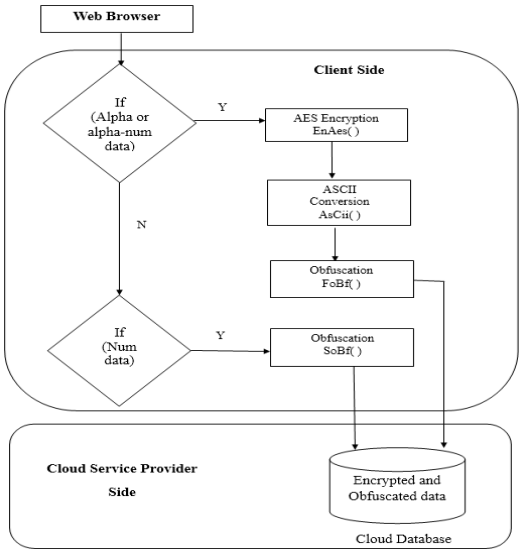

Figure 2: Workflow of the Unified Model

D. PSEUDO CODE FOR THE PROPOSED WORK

Pseudo code explains the working of the proposed model. Input will be the data from the table in database. Output will be the obfuscated data.

	ENCRYPT/OBFUSCATE	DECRYPT/DE-OBFUSCATE
input: $data_{ij}$ output: storage of data $mobfdata_{ij}$, $nobfdata_{ij}$ in cloud storage begin extract data, $data_{ij}$ from repository check $data_{ij}$ for datatype if datatype of $data_{ij}$ is character begin perform encryption function EnAes( ) on data $data_{ij}$; perform ASCII conversion AsCii( ) on data $endata_{ij}$; perform obfuscation function FoBf( ) on data $asdata_{ij}$; return $mobfdata_{ij}$ end if datatype of $data_{ij}$ is numerical begin perform obfuscation function SoBf( ) on data $data_{ij}$; return $nobfdata_{ij}$ end end	AsCii( ) begin $asdata_{ij}$ = ascii($data_{ij}$) return $asdata_{ij}$ end ---------- FoBf( ) begin s = sqrt($asdata_{ij}$) r = roundOf(s) $mobfdata_{ij}$ = r + r return $mobfdata_{ij}$ end ---------- SoBf( ) begin $nobfdata_{ij}$ = cbrt($data_{ij}$) return $nobfdata_{ij}$ end	AsCii( ) begin $data_{ij}$ = ascii($asdata_{ij}$) return $data_{ij}$ end ---------- FoBf( ) begin r = ($mobfdata_{ij}$/2) s = pow(r,2) $asdata_{ij}$ = floor(s) return $asdata_{ij}$ end ---------- SoBf( ) begin $data_{ij}$ = pow($nobfdata_{ij}$) return $data_{ij}$ end

Figure 3: Pseudo code for the model

VI. BENEFITS OF PROPOSED MODEL

Cloud computing data storage is growing now-a-days. The basic requirements of cloud are to meet the various needs of the organization and also to provide the secure storage. For the secure storage, encryption techniques are used. Only the encryption alone will not protect the data from the malicious attackers. When the data are stored in the cloud machine and the cryptographic keys are stored in the cloud servers it becomes untrusted data. Cloud service provider is not so trust worthy. Because they have the full control over the data once it is outsourced to the cloud, sometimes they may leak or damage the sensitive data. None of the cloud model provides the complete control to the data owner.

This proposed confidentiality technique will solve the above discussed problems. It also gives many benefit to store the data in the cloud in a secure manner.

- Here, the data owner has the full control over the data. This confidentiality technique is applied to the data only at the client side, before sending the data into the cloud.
- This model uses Advanced Standard Encryption algorithm to encrypt the data. It is one of the best symmetric encryption algorithm. Cloud storage contains a large set of databases and for such a large database asymmetric-key algorithms performance is slower when compared to symmetric-key algorithms.
- Obfuscation technique are used to encrypt the numerical values. When the character based data are converted into values and obfuscate then there will be more security to the data.
- This two data prevention techniques are combined together to secure the cloud data.

Hence, this proposed confidentiality technique provides speed in encrypting and decrypting. Privacy in data storage and also maximum security protection in cloud database.

VII. CONCLUSION

Cloud computing helps to save enterprises time and money, but trusting the cloud storage is difficult. The organization needs unpredictable data storage to store all their real asset, so they go for the cloud data storage. But the main issue in storing the data is security. Data security in the cloud will be ensured by the confidentiality of the data. This paper proposes a unified model to give maximum protection to the data. Encryption and obfuscation are the two different data protection techniques used in this model. To achieve high security, this model combines two data protection techniques encryption and obfuscation. Encryption alone will not give high security because secret keys

are shared among cloud service providers, there may also malicious insiders to attack the data. Obfuscation is a new technique to encrypt the data. The user data are encrypted and obfuscated at the client side before they are uploaded into the cloud storage. Unification of both the techniques gives major security to the user's data in the cloud storage. This model gives maximum protection, but the encryption algorithm used in this model is a commonly used algorithm. So, in future we intend to work with our own encryption algorithm to ensure more security on the data over the cloud.

VIII. ACKNOWLEDGEMENTS

This research was supported by a grant from University Grants Commission (UGC) awarded to A. Vithya Vijayalakshmi under Junior Research Fellowship.

IX. REFERENCES

[1]. Pradnyesh Bhisikar and Prof. Amit Sahu, "Security in Data Storage and Transmission in Cloud Computing", International Journal of Advanced Research in Computer Science and Software Engineering, 2013, Vol.3, Iss. 3, pp. 410-415.

[2]. Yashpalsinh Jadeja and Kirit Modi, "Cloud Computing - Concepts, Architecture and Challenges", International Conference on Computing, Electronics and Electrical Technologies, 2012, pp. 877-880.

[3]. Sarika Gupta, Sangita Rani Satapathy, Piyush Mehta and Anupam Tripathy, "A Secure and Searchable Data Storage in Cloud Computing", IEEE, 2012, pp. 106-109.

[4]. Fawaz S. Al-Anzi, Ayed A. Salman, Noby K. Jacob, and Jyoti Soni "Towards Robust, Scalable and Secure Network Storage in Cloud Computing", IEEE, 2014, pp. 51-55.

[5]. Manpreet Kaur and Rajbir Singh, "Implementing Encryption Algorithms to Enhance Data Security of Cloud in Cloud Computing", International Journal of Computer Applications, 2013, ISSN: 0975 - 8887, Vol. 70, pp. 16-21.

[6]. R. Anitha, P.Pradeepan, P.Yogesh, and SaswatiMukherjee, "Data Storage Security in Cloud using Metadata", International Conference on Machine Learning and Computer Science, 2013, pp. 26-30.

[7]. Hyun-Suk Yu, Yvette E. Gelogo and Kyung Jung Kim, "Securing Data Storage in Cloud Computing", Journal of Security Engineering, 2012, pp. 251-260.

[8]. Chaitanya Dwivedula and Anusha Choday, "Research on preserving User Confidentiality in Cloud Computing - Design of a Confidentiality Framework", International Journal of Engineering Research and Applications, 2013, ISSN: 2248-9622, Vol. 3, Iss. 2, pp. 35-52.

[9]. Popa Marius, "Techniques of Program Code Obfuscation for Secure Software", Journal of Mobile, Embedded and Distributed Systems, 2011, ISSN 2067 - 4074 vol. 3, pp. 205-219.

[10]. Dr. L. Arockiam and S. Monikandan, "Data Security and Privacy in Cloud Storage using Hybrid Symmetric Encryption Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, 2013, ISSN: 2319-5940, Vol. 2, Iss. 8, pp. 3064-3070.

[11]. Uegebe Ikechukwu Valentine and Omenka Ugochukwu Enyinna, "Building Trust and Confidentiality in Cloud computing Distributed Data Storage", West African Journal of Industrial & Academic Research, 2013, Vol. 6, PP.78-83.

[12]. Kawser Wazed Nafi, Tonny Shekha Kar, Sayed Anisul Hoque and Dr. M. M. A Hashem, "A Newer User Authentication, File encryption and Distributed Server Based Cloud Computing security architecture", International Journal of Advanced Computer Science and Applications, 2012, Vol. 3, pp.181-186.

[13]. Atiq ur Rehman and M.Hussain, "Efficient Cloud Data Confidentiality for DaaS", International Journal of Advanced Science and Technology, 2011, Vol. 35, pp. 1-10.

[14]. S. Monikandan and Dr. L. Arockiam, "Efficient Cloud Storage Confidentiality to Ensure Data Security", International Conference on Computer Communication and Informatics, 2014.

[15]. Varun Maheshwari, Arash Nourian and Muthucumaru Maheswaran, "Character-based Search with Data Confidentiality in the Clouds", IEEE International Conference on Cloud Computing Technology and Science, 2012, pp.895-899.

[16]. Yongkang Fu and Bin Sun, "A scheme of data confidentiality and fault-tolerance in cloud storage", IEEE CCIS, 2012, pp.228-233.