International Journal of Research in Advent Technology, Vol.4, No.3, March 2016
E-ISSN: 2321-9637
Available online at www.ijrat.org

Data Classification Based Security Framework in Cloud Environment

Apeksha Patel[1], Dishant Soni[2]
Computer Engineering Department,Final Year Student,SPCE,Visnagar, India[1]
Computer Engineering Department,Assistant Professor,SPCE,Visnagar, India[2]
patelapexa1@gmail.com[1], drsoni.ce@spcevng.ac.in[2]

Abstract— Cloud Computing gives many security challenges because of the increased use of Cloud services day by day. The Cloud platform are offered with many options and billing facility. In the race of achieving more security, user are always compromising performance and cost. The classifications technique of the data determines the extent to which the data needs to be secured. To secure the data of Cloud Computing classification based frameworks are found useful to minimize the perfomance overhead and cost. In this paper a study has been made in the direction of classification based data security in Cloud and compared using the parameters data integrity, confidentiality, frequent access, speed, time and cost. The proposed scheme will offer a proper data protection mechanism based on the user classification of data. So based on the scheme user can be provided various way to protect his data on the Cloud platform, Which guarantees the optimized IO access and also balance the performance and security parameters.
Keywords—Data classification,k-nn classifier,VDCI technique

1. INTRODUCTION

Cloud computing is a large scale distributed computing environment where software, platform and infrastructure services are available on demand to the users. The concept of Cloud computing implies providing everything i.e. hardware or software resources to customers on the virtual platform of Cloud. Users can use this services on rent so they do not have to purchase their own infrastructure and software for short period of time so they can save their time and money using the Cloud services. All of this services and computing resources are available through a simple Internet connection using a standard browser or other tools. However, there are still many problems exist in Cloud computing today, a recent survey shows that data security and privacy risks have become the primary concern for people to shift to cloud computing. The need of security in cloud computing has been increased due to the widespread use of Cloud computing in different area.

Cloud computing needs secure environment to fulfill the user's requirement of on-demand scalable and inexpensive services. Security of the sensitive data is a major issue in Cloud environment. data theft, data loss, and inaccurate data are the serious issue which require more secrecy and more control in cloud computing. Monitoring of data access by service provider is necessary to assure who accessed the data, for which purpose and when the data is being accessed. Verification of user data is also needed which consists that service provider must ensure about user or consumer data which must be accessed only by authorised users. To protect the data we can use third party auditor who manage and audit the user data and can periodically check data integrity. so it gives data correctness and accuracy of consumer or user data. using classifier approach we can give enhancements in perfomance of security of sensitive data in cloud computing.

The main Cloud service providers are Google, Microsoft, Salesforce.com, Vmforce.com and Amazon etc. The Cloud computing system depends on the layers SaaS, PaaS, and Iaas for information transportation. [1]

2. RELATED WORK

In 2015, Darwazeh, Nour S., Raad S. Al-Qassas, and Fahd AlDosari[2] analyzed all security issues and than consider the problem that encrypting all data without consideration of its confidentiality increase the processing time and complexity. so they proposed a framework which is used to encrypt the data based on confidential level. Using the proposed framework users can encrypt their own data using a key which only user knows, not knowing by service provider. A cloud storage model classified into three levels, basic, confidential, and highly confidential according to confidential degree. An efficient confidentiality-based cloud storage framework enhances the processing time and assures confidentiality and integrity through data classification and applying TLS, AES and SHA based on the type of classified

data. Shaikh, Rizwana, and M. Sasikumar[3] provide a data classification method to secure the data. A proposed data classification method has three type of characteristics access control, content and storage. Based on this three characteristics all the data stored on cloud have been classified. Based on type of content and access control parameter security levels are provided as per requirement of confidentiality of user. Also they provide a regular backup and recovery plan for data.

In Existing work Zardari, Munwar Ali, Low Tang Jung, and Nordin Zakaria[4] analyzed that security approach must be decided after classification of data based on security need which data need low level security, which data need higher security and which does not need. They proposed a K-NN classifier for data confidentiality in cloud. They have used K-NN classifier to classify the data in sensitive and non-sensitive data. They used RSA algorithm to encrypt sensitive data for protecting sensitive data from unauthorized users. Public data does not need any encryption so VM store public data directly without encryption. So this type of classifier approach based on classification save the processing time and memory resources. Yogesh V. Bhapkar, Rakesh S. Gaikwad and Milind R. Hegade[5] have examined about the problem of ensuring the integrity of data storage in cloud. So they proposed a system to ensuring the correctness of user data in Cloud computing system. A proposed system uses Third party auditor (TPA) to check the integrity periodically of user data stored on cloud and providing increased security level system.

Islam, Md Rafiqul and Mansura Habiba[6] proposed a system architecture for three tier security framework provides a security strategy as per user's demand. Different Security levels are provided based on classification of data. They have analyzed performance with respect to overhead for different security services such as confidentiality, integrity, and authenticity and they showed that using this framework performance of system is increased. In 2013, Zardari, Munwar Ali, Low Tang Jung and Muhamed Nording B. Zakaria[7] learned about the security concerns of the user and malicious activities of cloud. So they provided Hybrid Multi-cloud data security (HMCDS) model in which data is classified into three classes most sensitive, sensitive and non-sensitive data. They provided high level of confidentiality using implemented HMCDS model and low level of confidentiality using data classification technique. Multi cloud and clusters are used in this model. Moghaddam, Faraz Fatemi, Moslem Yezdanpanah and Touraj Khodadadi, Mohammad Ahmadi, and Mohammad Eslami[8] found that Applying same security strategy in all data decrease the level of efficiency and reliability. So they proposed a VDCI technique is based on three security parameters confidentiality, integrity and availability. Using the proposed technique the values of the three security parameters are determined automatically regarding stored specification in the history of data. They designed the algorithms of 3-level security and 5-level security for assigning process.

From the study and analysis of the research papers we have analyzed some parameters such as confidentiality, integrity, Frequent access, Cost, time and speed and compare them as shown in table 1. Also we have found the security mechanisms which are used to secure the data in each paper.

TABLE.1 SECURITY BASED ON DATA CLASSIFICATION RELATED WORK

References	**Security mechanisms**	**confidentiality**	**integrity**	**Frequent access**	**cost**	**time**	**speed**
[2]	SHA512+ AES12	1	1	4	2	3	2
[3]	Encryption technique	4	1	1	4	4	4
[4]	RSA algorithm	3	4	4	2	2	1
[5]	TPA+ AES algo.+ hashing	3	1	4	2	2	4
[6]	DES+IDEA +3DES	1	1	4	4	2	1
[7]	Encryption algorithm	1	4	4	2	4	4
[8]	VDCI based technique	3	4	4	4	4	4

1. High, 2. Less , 3. Improved 4. None

3. PROBLEM DOMAIN

Users are very much concerned about security of their data on Cloud platform. To make the data secure they

have always been offered Encryption solution to protect data on Cloud storage. Users are forced to use the scheme irrespective of the sensitivity of his data. Encryption and Integrity check on the large data which creates IO-Cost in Cloud computing. Encryption schemes always creates performance trade-off against security. Applying same security strategy in all data takes more processing time and memory resources. It also decrease the efficiency and reliability of the system.

There are many domain of research in Cloud computing , but we decided to explore data storage security . As it is dealing with user's data which is most important entity in Cloud Data Storage."

4. RESEARCH GAPS

The classified approach will easily decide the security needs of the data. After the data classification, we can easily select an appropriate security mechanism for data according to the need of data.

we can plan to enhance the framework by considering other aspects. This includes automatic data classification and the use of different cryptographic algorithms such as asymmetric public key, RSA, and Elliptic curve cryptography that could provide higher degree of confidentiality and security [2]. The efficiency of the proposed classification scheme is analyzed with the sample dataset collected. The proposed classification can be implemented as a working module i.e. prototype and simulation of the classification technique can be evaluated[3]. Providing security services according to different levels of security enhances the performance of the system without taking additional time overhead [6].

5. PROPOSED WORK

We aim to develop A Mechanism which works on the user's classification of data. So based on the proposed scheme user can be provided various way to protect his data on the Cloud platform Data classification can help in choosing best encryption mechanism to maximize the performance. We will try to employ suitable data classification technique so that data categorization can be atomize. Higher degree of confidentiality can be achieved with suitable combination of encryption techniques. The proposed model will also take care about integrity of the data. Proposed framework will help to choose the appropriate algorithm and key size according to the classification of data.
The Encryption and classification flow diagram gives the process flow of our proposed work as following :

Encryption and classification flow diagram

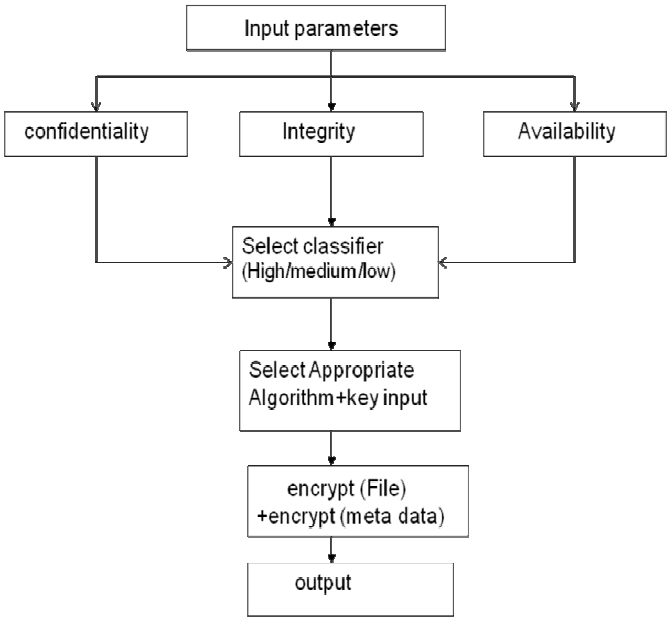

Fig. 1 flow diagram of proposed system

In the proposed system user will give the input of three parameters confidentiality, integrity and availability. Then the classifier will choose the classification high, medium or low using the three parameters. According the classification the system will select the appropriate algorithm and key size. Then we will have encrypted file and encrypted meta data and thus it gives the final output.

6. CONCLUSION AND FUTURE WORK

From the study and analysis presented in this papers we found that, The Cloud has still security and privacy issues in the main focus due to the cost and performance of the chosen model. We have studied several classification techniques and frameworks to protect the data stored on Cloud storage. classifier approach achieves better security of sensitive data in Cloud computing. Different standard encryption algorithms such as DES, AES, SHA, RSA are used to protect the sensitive data. Using the hybrid model user can save their cost of security and increase the confidentiality of the sensitive and most sensitive data. Using our proposed scheme we can easily select an appropriate security strategy according to the security need of the data. The proposed scheme will provide balance between processing time and security of the data.
It will also consider IO access requirement to minimize the resource billing.

REFERENCES

[1] Mahalle, Sheetal, and Ranjeet Jaiswal. "Cloud Computing Security: A Survey."International

Journal of Computer Applications 115, no. 6 (2015).

[2] Darwazeh, Nour S., Raad S. Al-Qassas, and Fahd AlDosari. "A Secure Cloud Computing Model based on Data Classification." Procedia Computer Science 52 (2015):,pp. 1153-1158.

[3] Shaikh, Rizwana, and M. Sasikumar. "Data Classification for Achieving Security in Cloud Computing." Procedia Computer Science 45 (2015): pp. 493-498.

[4] Zardari, Munwar Ali, Low Tang Jung, and Nordin Zakaria. "K-NN classifier for data confidentiality in cloud computing." In Computer and Information Sciences (ICCOINS), 2014 International Conference on, IEEE, 2014, pp. 1-6.

[5] Yogesh V. Bhapkar, Rakesh S. Gaikwad, Milind R. Hegade" Providing Security And Privacy To Cloud Data Storage" International Journal of Computer Science and Information Technologies(IJCSIT), Vol. 6 (2) , 2015, pp. 969-971 .

[6] Islam, Md Rafiqul, and Mansura Habiba. "Agent Based Framework for providing Security to data storage in Cloud." In Computer and Information Technology (ICCIT), 2012 15th International Conference on, IEEE, 2012, pp. 446-451..

[7] Zardari, Munwar Ali, Low Tang Jung, and Muhamed Nording B. Zakaria. "Hybrid Multi-cloud Data Security (HMCDS) Model and Data Classification." In Advanced Computer Science Applications and Technologies (ACSAT), 2013 International Conference on, IEEE, 2013, pp. 166-171..

[8] Moghaddam, Faraz Fatemi, Moslem Yezdanpanah, Touraj Khodadadi, Mohammad Ahmadi, and Mohammad Eslami. "VDCI: Variable data classification index to ensure data protection in cloud computing environments." In Systems, Process and Control (ICSPC), 2014 IEEE Conference on, IEEE, 2014, pp. 53-57.

[9] Saxena, Tunisha, and Vaishali Chourey. "A survey paper on cloud security issues and challenges." In IT in Business, Industry and Government (CSIBIG), 2014 Conference on, pp. 1-5. IEEE, 2014, pp. 53-57.

[10] Zhang, Qi, Lu Cheng, and Raouf Boutaba. "Cloud computing: state-of-the-art and research challenges." Journal of internet services and applications 1, no. 1 (2010): IEEE, pp. 7-18.