IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

VOLUME : 2 ISSUE : 2 Print / Issue Publication Date: 27-Jan-2017

ISSN : 2455-2143

Indexed In

Google SCHOLAR OPEN ACCESS doi Crossref

WWW.IJEAST.COM

editor@ijeast.com

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 2, Issue 2, ISSN No. 2455-2143, Pages 141-146
Published Online December-January 2016 in IJEAST (http://www.ijeast.com)

THREE LEVEL AUTHENTICATION SYSTEM TO ENHANCE SECURITY ON CLOUD USING K-MEAN CLUSTERING TECHNIQUE

Harmanpreet Kaur Sidhu
Research Scholar
Department Of Computer Science Engineering
Chandigarh University, Gharaun, India

Er. Mamoon Rashid
Assistant Professor
Department Of Computer Science Engineering
Chandigarh University, Gharaun, India

Abstract- Cloud computing is a model that enables convenient and on demand network access to a shared pool of configurable computing resources where millions of users share an infrastructure. Cloud is a multi-tenant environment in which consumers consume Cloud Service provider's infrastructure which is shared among other consumers. Privacy and Security is significant obstacle that is preventing the extensive adoption of the public cloud in the Industry. In this paper, the authors preserved privacy in data storage by validating through IP based detection, Geographical location based security technique. After that privacy is enabled by using the k-mean clustering technique for validating the user accesses based on spending subscriptions when consumer want to use the organization services and constant key length encryption technique irrespective of users is used to secure data on cloud storage.

Keywords- **Computing, Multi-tenancy, IP addressing, K-mean, clustering.**

I. INTRODUCTION

Cloud has begun to grow as a hotspot in both industry and scholarly world. Cloud computing is a not new paradigm it is the further advancement of grid computing and distributed computing. Cloud Computing provides on demand provision of resources. Cloud computing provides economics benefits to the small and large enterprises by reducing the capital expenditure and other expenditures. Cloud storage is a new business architecture for delivering virtualized storage to customers on demand [2]. A cloud storage provider gives storage as a support to the consumers to store their sensitive and non-sensitive data to the cloud. Cloud storage models can be implemented as public storage model, private storage model and hybrid storage model. A storage provider provide additional services to the consumers like file sharing, document compatibility, backup of data, big data storage, scale up and scale down etc. In spite of the enormous business and specialized favourable circumstances of the cloud storage benefits, the data confidentiality is one of the major flaw in widespread acceptance of cloud. Cloud is a multi-tenant environment in which consumers consume Cloud Service provider's infrastructure which is shared among other consumers. But whenever a user shares information in cloud, the question of privacy and confidentiality comes in mind immediately. Privacy and security becomes an obstacle in adoption of cloud on the massive level. The major problem in cloud storage is that owner do not have full control over the data which is stored on the cloud. If any unauthentic user want to use the services of cloud of any user then there should be some security mechanism to provide privacy to the data. Therefore, consumer wants some security policies before storing data on cloud.

Clustering can be used in different ways according to the need of categorization. It is very frequently used method to save time for analyzing the large number of consumers in cloud and they are grouped together based on the spending subscriptions in this paper. Different algorithms are K-mean, fuzzy c-means, mountain, subtractive etc. [7].

This paper deals with the privacy concerns on the cloud storage which is due to the multi-tenant environment of cloud. The privacy is enhancing by applying the proposed algorithms such as IP based detection, Geographical location based security technique, k-mean clustering algorithm and by using encryption algorithm.

The rest of the paper is organized as follows: Section II describes the related work on privacy and security of data storage on cloud .In Section III, the problem is proposed by applying k-mean clustering technique and other security techniques for storing data on cloud. Section V discusses experimental setup and desired results are shown in Section VI. Section VII gives conclusion and future scope of the approach.

II. RELATED WORK

In recent years, several methods have been proposed to preserve privacy and provide efficient analysis to the shared data.

Witold Litwin [31] provided security assessment of cloud in 4 categories, namely customary security in a PC system, Availability of distributed computing applications, Third-party information protection, and Third-party information control. The traditional security covers issues such as

attacks on virtual machines, cloud provider weaknesses, authentication and authorization, and data stealing or leakage.

Ko et al. [30] proposed an execution model called HybrEx (Hybrid Execution) model. This model is outlined with a perspective to guarantee confidentiality and privacy. In this model basically more emphases is on the storage of sensitive and non-sensitive data. Public cloud was used to store the non-sensitive data and for computations and private cloud infrastructure is used to store sensitive data and computation of sensitive data is also done on the infrastructure of organization. The model supports application level partitioning i.e. application is divided into two partitions when an application works with both public and private data, the partition which involves with the public data made run on the public cloud. On the other hand partition with private data run on the private cloud. Basic purpose of the model is divide the data according to private and public cloud .this model allowed the integration of more resources to the private cloud so that confidentiality and security is maintained at the higher level. Although this approach provides security benefits but this model avoids the public clouds for sensitive data and computation.

Jianping Wu et al. [11], a proposed a Source Address Validation Architecture to give a straightforward system benefit that each packet received and sent have validated IP address. This architecture helped in minimizing the serious security and accounting problems.

Aniello Castiglione et al. [12], people and their devices move from one network to another and their IP addresses are changing as consequences. Author used Network and Port Address Translation logs to analyze the different piece of activities to discover the information. Recognized the recurring activity patterns helps to identify any offender's action by any device

Thorsten Ries et al. [14], proposed an geo-location approach based on network coordinate systems and even if cloud provider used other measures to hide the resource location and accuracy increased with modified approach to find the exact location.

III. PROBLEM FORMULATION

In accordance to the literature survey, it was found that, although current privacy models preserve the security on the cloud storage by using various encryption algorithms on client side and even by the cloud provider, still when any malicious user attacks the encryption algorithm and finds the decryption key then these models fails not in terms of security but also in the over usage of subscription which effects users in terms of financial loss. These confinements makes construction modelling quit ordinary as there is no such component to give protection on the data storage even any unauthenticated user have decryption key. Further it was investigated that there is no transaction limit for users to access the resources and clients can download and retrieve any amount of data from the cloud storage. This flaw leads to insecurity of the architecture because any number of resources can be accesses without any investigation at free from any user, if the account is hacked then it is a huge loss to organization and particular user**.** Moreover, cloud is a multi-tenant environment in which number of users using the same cloud storage platform to store their data. On the other hand, if any intruder got access to the authorized user then it will delete, download the private data from the authorized user's account. Furthermore, k-mean clustering algorithm only used to cluster the large amount of data and retrieve the information from them. Keeping each one of these blemishes of present models into thought, there is need of a compositional framework which will address every one of these issues and thus will come about as an all the more effective model.

IV. PROPOSED APPROACH

The proposed approach aims to enhancing the privacy on cloud storage using K-mean clustering algorithm, IP addressing based detection technique, Longitude and latitude technique and Cryptography technique for data uploading on cloud.

Administrator is solely responsible for managing users, uploading data in the cloud storage and management of keys which are used for encryption and decryption of the data to upload on the cloud and download from the cloud storage. Administrator has all the access over data and will upload, read and download the data anytime and anywhere. Firstly, when user is created by administrator, users are restricted to use the organizations resources on the subscription bases. Although it looks like inconvenience to users as they cannot access the organizations resources free and they have limited transaction limits. However, this feature adds security to the system because imposing transaction limit will reduce the loss even when unauthorized user will enters into the system. Normal users are allowed to use the data according to membership type and premium user access all the services of organization's system. Finally, the data is stored on the cloud storage.

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 2, Issue 2, ISSN No. 2455-2143, Pages 141-146
Published Online December-January 2016 in IJEAST (http://www.ijeast.com)

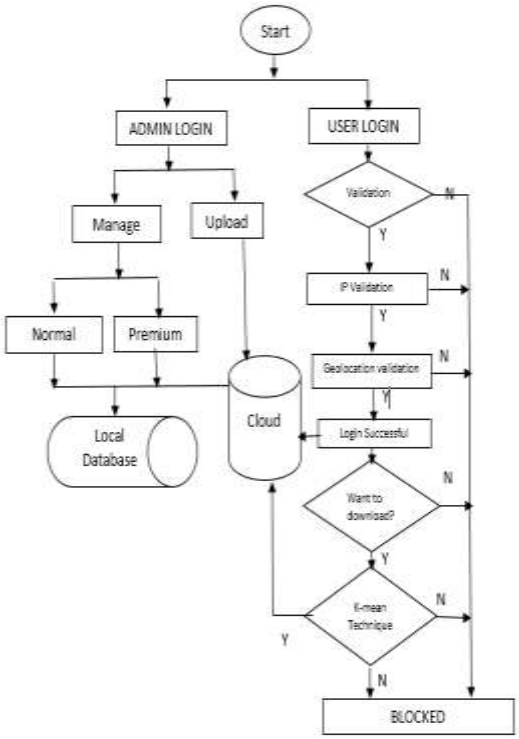

Fig. 1 Flow Chart of Proposed work.

4.1. **Privacy Perspective:**

1. Security is applied to the system when user tried to log in the system:
First of all when user login to the organizations system it will validate through the basic validation technique. After this another method is used to validate the genuine user is IP based detection, through this if user is using its account from the same IP address then it will be allowed to access its account. If IP address is different from previous one then location of user will be determined and calculates the distance between two locations. After calculating the distances it will be measured that is it possible to travel that distance in the two login time. If it is not possible then the user will be blocked. Figure 2 shows flow chart of login validation.

2. Privacy is applied through K-mean clustering Technique when user wants to download:
This technique is used to cluster the similar type of data in one cluster. In this approach k-mean is used to validating the user's access based on the spending subscriptions. When any user wants to download any file from the storage account then k-mean will find the cluster of user in which its subscription limit exists. If the requested type of transaction is not belongs to that cluster then it will block the user to protect the data because any intruder can misuse the sensitive data. Figure 3 shows K-mean clustering Technique.

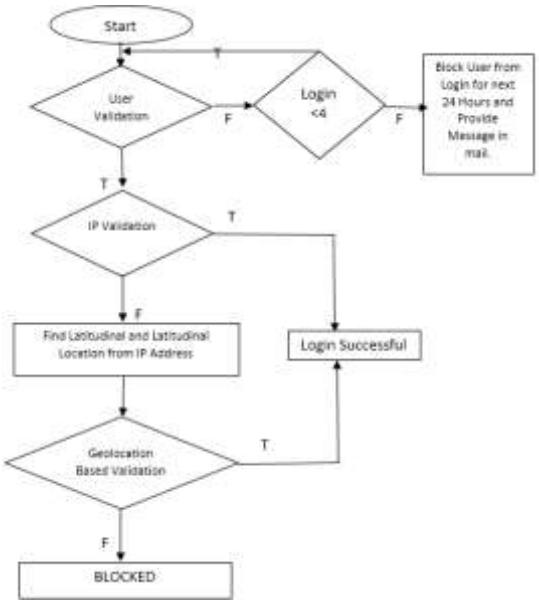

Fig. 2 Flow Chart of Login Validations

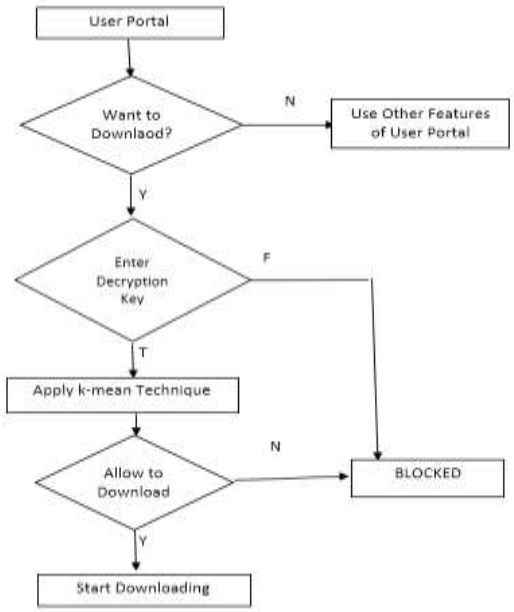

Fig. 3 Flow Chart of K-mean clustering Validation

Algorithm 1-LOG IN Validation

Step 1: Start
Step 2: Validate Username and Password.
Step 3: If recent IP address is equal to any of the previous IP address,
then
login successful

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 2, Issue 2, ISSN No. 2455-2143, Pages 141-146
Published Online December-January 2016 in IJEAST (http://www.ijeast.com)

else
end If
goto step 4
Step 4: Calculate distance between recent IP and Previous IP
Step 5: Calculate time form last login to recent login.
Step 6: If distance is covered in time
then
login successful
else
end If
login failed

Algorithm 2- K-mean Technique

Step 1: Select k objects from dataset D as initial cluster set.
Step 2: Calculate the distance d between each object d_i (1 <=i <=n) as Euclidean distance and assign data object d_i to the nearest cluster.
Step 3: Find distance d_n of new object to Cluster centers c_j.
Step 4: If distance of new object d_n<= Euclidean distance
then
allow to download and add object to cluster
else
block the user to download.
end If

V. IMPLEMENTATION

The system is implemented in MVC Framework with C# Scripting Language. Testing of web application is done in local environment and after that web application is deployed on the Microsoft Azure cloud platform and any user can use this application from anywhere and anytime. On Microsoft Azure cloud SQL database, storage account is used to store the data on the cloud. In the next step the data is uploaded on the cloud through our implemented Graphical User Interface. The storage account is created for uploading the data on Microsoft account for the same purpose. In the last, all the users are allowed to use the services and access the resources and data on the Microsoft Cloud which is stored by administrator through this implemented architecture.
On assumption is that the administrator successfully login into the system. Now admin is redirected to the main interface of this model and admin can now manage users and upload the data to the cloud storage. Administrator is solely responsible for managing users, uploading data in the cloud storage and management of keys which are used for encryption and decryption of the data to upload on the cloud and download from the cloud storage. Administrator has all the access over data and will upload, read and download the data anytime and anywhere. Firstly when user is created, users can access the data on the cloud according to membership. Although it looks like inconvenience to users as they cannot access the organizations resources free and they have limited transaction limits. However, this feature add security to the system because imposing transaction limit will reduce the loss when system is used by unauthorized user. Moreover, administrator can block any user according to the trust level of the user. User are allowed to access the resources according to subscription and premium user access all the services of organization's system. Finally, the data is stored on the cloud storage.
The system work as when any user want to login into the portal and enters username and password. Here first it validates the username and password and here if any unknown user tries to login into the system then after three attempts a mail is sent to the mail address which is registered by the administrator to inform about the blocking of account. This will add the more security to our system. After the successful login user can access the stored data on the cloud. Furthermore, the user validation IP address validation is applied if IP address of the user is matched with the previous login IP address then user make successful login. Otherwise Geographical location of the IP address is find with the help of service provider and from longitude and latitude the distance between the previous location and location of current login is calculated. If the calculated distance is covered with maximum travelling speed between two login times then the user will login successfully to the system otherwise user is blocked and message is mailed to the user. When user successfully crossed all the enabled security mechanisms then it will be able to access all the services provided by the cloud service provider.
Now when user successfully login into the system and want download the stored data on cloud then a k-mean security mechanism is implemented to avoid the case where resources can be accessed by the unauthentic user. If user want to download the stored data then k-mean technique is applied in which cluster is formed according to the spending subscription history. If the current request is not lie within the cluster then the user is not allow to download. In this way system the loss of resources of a particular user is diminished and system will work in their way to provide a security to the system even intruder or unauthenticated user has the decryption key.

VI. RESULTS

In this executed model, we have effectively included the features of imposing downloading limits for each user and changing access limits for clients to get the data taking into account their premium membership or normal membership. IP based security validation and geographical location based validations are imposed on the system so that unauthorized users are blocked before login into the system. Sometimes even after the stringent login validations unauthorized user may login into the system and tries to access the user private data then k-mean based validation approach is used to

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 2, Issue 2, ISSN No. 2455-2143, Pages 141-146
Published Online December-January 2016 in IJEAST (http://www.ijeast.com)

secure data from unauthorized user. For test results we have created different scenario in this system.

Scenario 1: First administrator make successful login into the system. Administrator is the only premium users and other users are normal user. For test purpose 5 users are added into the system by administrator and data limit is 5 MB, 6MB, 8MB, 10MB and 12MB respectively for the users. For testing purpose admin login into the system and tries to download the data for any size and it can download any number of data with any size without any decryption key. After that for testing purpose any user login into the system successfully and tries to download the data beyond its limit then data is not downloaded

Scenario 2: For testing an unauthorized user tries to login into the system with different location. Unauthorized user got the username and password by eavesdropping from the network. When unauthorized user successfully makes basic validation after that IP address matching is executed and IP address is different from previous one in this case then geographical location is found of the IP address. Current location of the user is not able to cover in the time between two logins then the user is blocked.

Scenario 3: When unauthorized user successfully login into the system and tries to download the data. Suppose an unauthorized user login into the user portal and tries to download the users data. System will not allow the user to download the data without decryption key this will add security into the system.

Scenario 4: when unauthorized user have decryption key for downloading. When user enters the decryption key to download then k-mean technique is executed in which cluster of that user is formed from the usage pattern of downloading and unauthorized user just want to fully used the data left for this purpose large size of data is downloaded by the unauthorized user. That request lies beyond the cluster's centroid then data is not downloaded even the user knows the decryption key.

Hybrid execution model is outlined with a perspective to guarantee confidentiality and privacy. In this model basically more emphases is on the storage of sensitive and non-sensitive data.The model supports application level partitioning i.e. application is divided into two partitions when an application works with both public and private data, the partition which involves with the public data made run on the public cloud. In this encryption done with the symmetric key but this model solely depend upon the encryption technique. If cryptographic technique used is failed then this model fails. In efficient privacy protection scheme is IP based security is used in which source address validation scheme is used and with this encryption algorithm is used. If encryption is failed no other security mechanism is applied to add security to the system. In the proposed model k-mean enhances the privacy on cloud storage even when an unauthorized user has access to the decryption key.

Table 1: Comparison between Approaches.

VII. CONCLUSION AND FUTURE WORK

To summarize, the work gives a model of a framework that can be used by the organization's data to protect and manage their stored data over un-trusted public clouds. As part of the work the possibilities of using IP address based validations and geographical based validations along with the constant key length key encryption scheme irrespective of users to encrypt files. The cryptographic algorithm will deal with individual files which are stored on cloud storage and decrypt the files by using different decryption keys. Moreover, K-Mean clustering technique validates the user accesses based on spending subscriptions. This model blocked the users depending upon the history of using its subscription. This model works in the swear circumstances when an unauthorized user have decryption key even then intruder unable to download the files from the user portal.

Table -1 Comparison of proposed model with others

Models / Features	HbrEx	EPPS	Proposed Model
IP Address Based Security	NO	YES	YES
Geolocation based security	NO	NO	YES
Encryption Algorithm	YES	YES	YES
Work when encryption fails	NO	NO	YES
K-mean clustering technique	NO	NO	YES

If encryption fails even then this model is active to provide the security to the system. The implemented framework can possibly be valuable in business circumstances where privacy and security of the client's data is main priority, as it catches reasonable practical access arrangement in an adaptable way and gives secure data storage in the cloud enforcing these privacy approaches.

In future, k-mean technique can be also used for deleting and reading the data from cloud. Moreover, we can use different length for encryption technique with respect to users which will make the system more reliable.

VIII. REFRENCES

[1] A Vouk, Mladen. "Cloud computing–issues, research and implementations", *CIT. Journal of Computing and Information Technology 16*, Vol. no. 4, PP. 235-246, 2008.

[2] Mell, Peter, Tim Grance. "Draft NIST working definition of cloud computing", *Referenced on June*, 2009.

[3] Hyun-Suk Yu, Yvette E. Gelogo, Kyung Jung Kim, "Securing Data Storage in Cloud Computing", *Security Engineering Research Institute (Journal of Security Engineering)*, Vol. No. 9, Issue No. 3, pp 251-260, 2012

[4] Tim Mather, Subra Kumaraswamy, and Shahed Latif, "Cloud Security and Privacy", *O'Reilly Media, Inc*., r 4, pp 61-71, 2009.

[5] W.K. Denial, "Challenges on privacy and reliability in cloud computing security", *Electronics and Electrical Engineering (IEEE), International Conference*,Vol.2, pp. 1181 - 1187, 2014.

[6] M. Zhou, R. Zhang, W. Xie, W. Qian, A. Zhou ,"Security and Privacy in Cloud Computing: A Survey," *Sixth International Conference on Semantics, Knowledge and Grids*, Vol.2, pp. 105 - 112 ,2014.

[7] K. Hammouda, "A Comparative Study of Data Clustering Techniques," *International conference*, 2012.

[8] S. Na, L. Xumin, "Research on k-means Clustering Algorithm," *Third International Symposium on Intelligent Information Technology and Security Informatics*, 2010.

[9] G.Logeswari, D.Sangeetha, V.Vaidehi," A Cost Effective Clustering based Anonymization Approach for Storing PHR's in Cloud", *International Conference on Recent Trends in Information Technology*, vol. 3 , pp. 234-235, 2014.

[10] Chang F, Dean J, Ghemawat S, "Bigtable: A distributed storage system for structured data", *ACM Transactions on Computer Systems (TOCS)*, 2013.

[11] Wu, Jianping, Gang Ren, and Xing Li. "Source address validation: Architecture and protocol design." *IEEE International Conference* on, pp. 276-283. IEEE, 2007.

[12] Castiglione aniello,"Device Tracking in Private Networks via NAPT Log Analysis", *Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing-IEEE*, 2007.

[13] Ries Thorsten ,"Verification of Data Location in cloud Networking", *Fourth IEEE conference on utility and cloud computing-IEEE*,2011.

[14] Jyoti Yadav and Monika Sharma, "A Review of K-mean Algorithm", *International Journal of Engineering Trends and Technology (IJETT)*, Volume 4 Issue 7, 2013.

[15] Chander Kant and Yogesh Sharma, "Enhanced Security Architecture for Cloud Data Security", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume3, Issue 5, 2013.

[16] Juntao Wang; Xiaolong Su, "An improved K-Means clustering algorithm", *3rd IEEE International Conference on Communication Software and Networks (ICCSN)*, pp. 44,46, 2011.

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

ABOUT IJEAST

International Journal of Engineering Applied Science and Technology (IJEAST) is a peer-reviewed, open access journal that publishes high-quality research papers in the field of Engineering, Applied Science and Technology.

IJEAST aims to provide a platform for researchers, academicians, and professionals to share their innovative ideas, research findings, and practical experiences with the global scientific community.

FOCUS AREAS

- Engineering
- Applied Science
- Technology
- Innovation & Development
- Interdisciplinary Studies

PEER REVIEWED
All submissions are rigorously peer reviewed to ensure quality.

OPEN ACCESS
Free and unrestricted access to research for all.

GLOBAL REACH
Connecting researchers and professionals worldwide.

TIMELY PUBLICATION
We ensure a swift and efficient publication process.

For more information, visit our website
www.ijeast.com

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

editor@ijeast.com
www.ijeast.com
India

2455-2143