RESEARCH ARTICLE OPEN ACCESS

Securing Data in Cloud using Disinformation Data in Fog Computing

Urmi Priyadarshani Das [1], VijayaKumar R [2], Dr B Rajalakshmi Ravishankar [3]
M.Tech [1], Assistant Professor [2], HOD [3]
Department of Computer Science and Engineering
New Horizon College Engineering, Bangalore
India

ABSTRACT

Increase in usage of computer and internet increases data. Storage of data in the conventional system is impossible now a days. Insufficient storage space enforces people to move to cloud computing. Cloud computing has flexibility, scalability, multiporatability, elasticity.Eventhough efforts are taken in the security of the cloud but still it contains loopholes which is restricting people to use cloud. Fog computing is the extension of the cloud computing to the edge of the network. Fog computing extends the storage, networking and computing facility of cloud computing. Security in data access is increased day by day though data is stored remotely. Volume of the data also increases day by day. Securing data by using various behaviour parameters and decoy data technology maintain confidentiality of data. This paper proposes an approach of using fog computing for securing the data with efficient algorithms and behaviour analysis on large data.

Keywords:-Cloud Security

I. INTRODUCTION

In this digital world storage and security of data is plays a measure role. Cloud computing overcome the storage issue of the digitization .cloud computing spreads its computing services through World Wide Web.

All large to small scale industries use cloud to store data. Now a days cloud is used widely to store data.[1] Cloud storage increases usability, accessability.It saves bandwidth by sending link instead of mailing file. Annual cost of the businesses and organizations is being reduced by cloud storage. Cost of setting up on premise infrastructure is also reduced by storing files into cloud and cloud service provider charge depending on the usage of the storage. Redundant data storage helps in disaster recovery of the stored data. Data is maintained and managed remotely in cloud storage. Back up of data is also stored remotely. [1]Files are stored online and can be accessed from any location via internet. Along with all the advantages of cloud storage it has some disadvantages. Security and accessibility of the data is one among the disadvantages of the cloud.

Data Stored in remote so internet is required to connect and access to the data, failure in connection causes failure in data access. Data theft is the measure issue in cloud security. [3]Personal information from 500 million accounts was being stolen claimed by yahoo. The internet company attributed the breach as a “State Sponsored” attack. Another data theft occurs when passport scan of first lady Obama was posted online. [3]In Equifax Company, hackers accessed the company’s system and people’s names, Social Security numbers, birth dates, addresses and in some instances, driver’s license numbers. California- households based data analytics firm had leaked online sensitive personal information of 123 Million American . [3]Accenture Cloud Platform Customer Sensitive Data got affected when an Unsecured Server Hosted on Amazon’s S3 storage service which is holding 137 gigabytes of Customer data has completely left unsecured. Uber data Breach, Hackers stole 57 million [3] Uber users around the world and 600,000 drivers names including their license numbers. To overcome such kind of data hacking in cloud storage, Fog computing is introduced by CISCO.
Fog computing introduced new characteristics of security and privacy using the decoy data for the detected malicious user.

A.FOG COMPUTING AND SECURITY

Fog computing was first initiated by CISCO which extend the cloud computing to the edge of the network. Virtualized platform of Computing, networking, storage facility is provided in between data centres and End Users by FOG Computing. FOG nodes requires trust model to establish reliability and security in fog network. Authentication in fog networking is the first and fore most requirement for security.

Authentication by the end user is required to ensure uninterrupted service to the registered users. Centralized control is difficult in scattered FOG nodes in large area. Intruder can enter into the fog network through poorly secure fog node. Most important privacy in location and data leakage depends on the location of the fog node. More communication between three layers increase security issue.[5]Most of the data theft happens by the malicious user insider to the cloud service The malicious insider to a cloud can easily attack to the user data. .Unauthorized access cannot be detected by end user science the attack came from the cloud service provider. To monitor the amount or the duration of the user data access User behaviour profiling is useful. Cloud computing uses many methods to secure data like encryption, access control misconfigured services, but fog computing uses behavioural analysis to find the intruder .If user is detected as intruder then according to disinformation attack decoy data is provided to the user .[5]

The proposed system detect intrusion by using behaviour analysis, On detection of intruders decoy data is provided to the user.

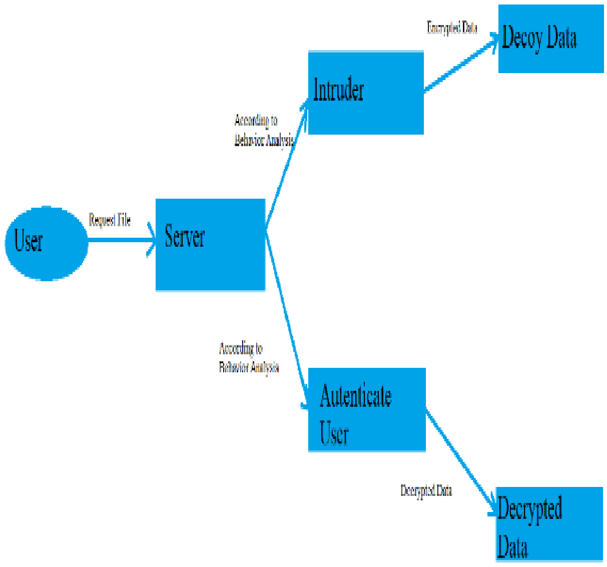


Figure.1 [Figure shows File request by user and output received by user.]

II.RELATED WORK

System to monitor data from the users to store in database and provides data security from malicious intruders detecting using behaviour analysis of the characteristics of the users data, protecting original data by providing encrypted data.[7]

Privacy and security in fog computing is measure concern in order to save the data and processing. Utilization of hardware resources, scalability, reduced cost, easy deployment are some areas of advantage of the cloud system. Hosting multiple users and data storage to the cloud environment has significant risk. Security concern around the traffic though node such as routers. Malicious software may be installed by hackers on the node of the fog computing. If temporary data need to be stored in this node then also privacy issue along with security challenges need to be addressed. Vulnerability can also increases security and privacy issues in fog computing. To overcome from the issue decoy data concept and disinformation attack against malicious insiders to the cloud computing.[6]

Malicious Insider to the cloud: Malicious insiders attack is one of the server attack to the cloud computing which has data theft attack by malicious insider to the cloud provider. Lack of authentication from the cloud service provider causes data theft. User data can be get accessed easily by malicious insiders to the cloud, though the attack came from the cloud service provider end users don’t detect the unauthorized access.[5]

Sophisticated attack cannot be fully protected if there are bugs in the code or faulty implementation of encryption and access control, mis configured service. Monitoring of the amount and duration of the data access is controlled by user behaviour profiling. Prediction of the malicious attack can be possible using user behaviour profiling by detecting abnormal behaviour of the end user [5].Stolfo et al.[4] proposed a new level of security for the cloud based on the user behaviour and profiling, if the abnormal behaviour is detected, then the decoy information is delivered to the true uses to obtain the response by many ways, e.g., security challenges. Otherwise, the decoy delivers a massive amount of garbage data to the attackers, thereafter, reducing the stolen information of the users. At the same time following issues arise as:

- Where to place the decoy in fog networks?
- How to design on-demand decoy information to further reduce the amount of stolen data.

Decoy Data: The header section of the decoy document contain Hash Message Authentication Code (HMAC) hidden inside it. HMAC act as a unique key to a file for each user. Verification of the decoy document is happening by computing a HMAC based on the content of the document. When user is requesting for a file then Comparison of the HMAC in the document will happen if two HMACs match the document is determine as a decoy document and one alert is being issued.[4]

Behaviour analysis: The Behaviour analysis of the user data to detect intruder can be done using two algorithm Naïve Bayes and SVM algorithm. Naive Bayes theorem is a probabilistic multiclass classification algorithm which aims the conditional probability distribution of feature such as login time, session time and upload count of the file [12].

SVM algorithm based on binary machine learning classification, which classifies into two classes taking all items as input. Finding maximum margin hyper plane is the main task. [12]

According to [9] Naïve Bayes and SVM have same measures of AUC but Naïve Bayes is simpler to implement than SVM as it calculates probability using simple mathematical probability concept.[7]

Shifting to Fog: In order to get better Quality of service, low latency, location awareness, real-time applications heterogeneity, wireless access the proposed system adopts FOG computing concept.[5],[7]

A.FEATURES TO SECURE DATA

Security of the data in cloud can be increased by detecting intruders and providing decoy data through a disinformation attack. Security in cloud services achieved by the feature

User Profile Mapping:-It is the type of mapping of the existing pattern to the current user behaviour pattern to find intruder. The behaviour of the user is continuously checked to access if any abnormality present .The behaviour of the person who access through the cloud is detected base on the characteristics:[8] Login Time, Session Time, Upload Count, Download Count, How many files are read and how often. Naive Bayes algorithm [7] compare is used to compare existing user data set to the new user dataset.

Decoy Data Technology:-When user upload any file that is being encrypted using AES encryption algorithm and stored in database in blob. If the user is detected as a unauthorized user the encrypted data is provided to the user .The encrypted data cannot be understandable by the unauthorized user.

If the user detected as authorized user then the encrypted file is being decrypted using key and the decrypted data given to the user.

HADOOP: Large data set in a distributed environment is analysed by HADOOP, which is a JAVA based programing framework. Files will be stored in HDFS and output being produced using mapreduce function on the dataset. In this proposed system mapreduce uses Naïve Bayes algorithm to analyse the data. The users behaviour is stored in My SQL server and detection of the abnormal behaviour is done Naïve Bayes algorithm in mapreduce function. If unauthorized user is detected decoy data is provided to the user. [7]

III. IMPLEMENTATION

A.MODULES

The implementation of project is divided into three module for easy understanding and development purpose. Control part of the project is done by admin module, cloud operations is being done by Cloud server module and the operation performed form the user is controlled by Client module .

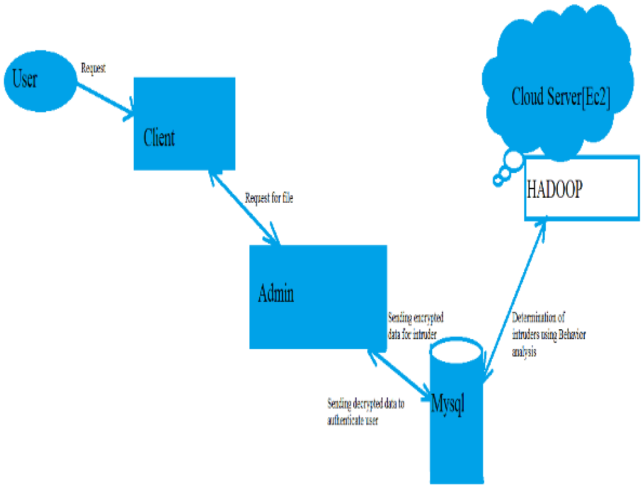

User
Request
Client
Cloud Server[Ec2]
HADOOP
Request for file
Admin
Sending encrypted data for intruder
Determination of intruders using Behavior analysis
Sending decrypted data to authenticate user
Mysql

Figure.2 [Flow between the modules]

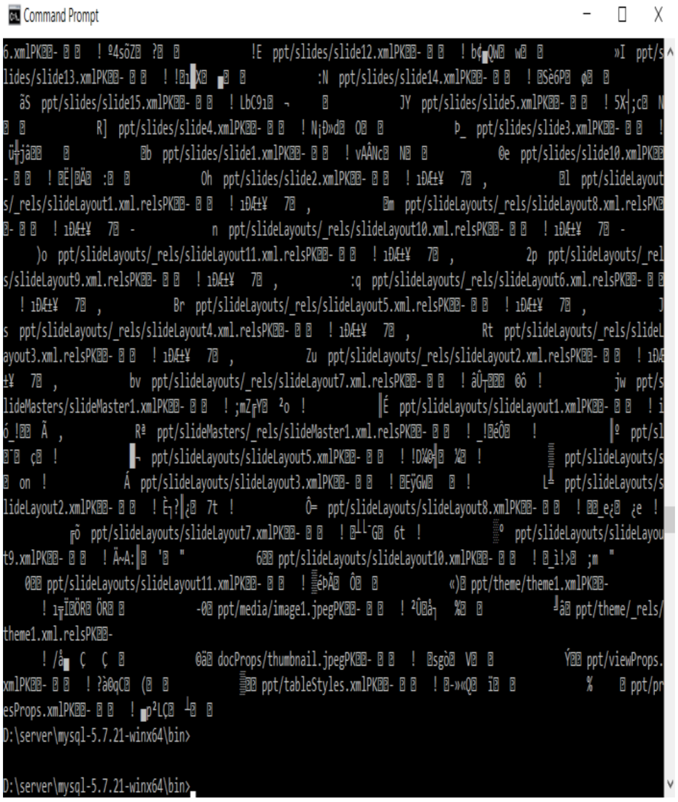
ADMIN: Admin module is responsible to enter user data and also decoy data into the database.

The work flow for user data and decoy data contains following steps

1. User account is created using login screen

2. Password is being encrypted and stored in the database .Encryption is done using salt value. For authenticated user password is being decrypted allow user to login into account.

3. Every time someone want to enter into the account the above two steps will repeat.



```
6.xmlPK▯▯- ▯ ▯  ! º4sõZ▯  ?▯  ▯            !E  ppt/slides/slide12.xmlPK▯▯- ▯ ▯  ! b¢▯QW▯  w▯  ▯          »I  ppt/s
lides/slide13.xmlPK▯▯- ▯ ▯  ! !▯1▯X▯  ▯▯  ▯          :N  ppt/slides/slide14.xmlPK▯▯- ▯ ▯  ! ▯Sè6P▯  ø▯  ▯
  ãS  ppt/slides/slide15.xmlPK▯▯- ▯ ▯  ! LbC9ı▯  ¬      ▯           JY  ppt/slides/slide5.xmlPK▯▯- ▯ ▯  ! 5X┤;c▯  N
▯  ▯          R]  ppt/slides/slide4.xmlPK▯▯- ▯ ▯  ! N¡Ð»d▯  O▯  ▯           Þ_  ppt/slides/slide3.xmlPK▯▯- ▯ ▯  !
 ü╫jâ▯▯    ▯          ▯b  ppt/slides/slide1.xmlPK▯▯- ▯ ▯  ! vÂÂNc▯  N▯  ▯          0e  ppt/slides/slide10.xmlPK▯▯
- ▯ ▯  ! ▯Ë|▯Ä▯  :▯  ▯          Oh  ppt/slides/slide2.xmlPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          ▯l  ppt/slideLayout
s/_rels/slideLayout1.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          ▯m  ppt/slideLayouts/_rels/slideLayout8.xml.relsPK▯
▯- ▯ ▯  ! ıÐÆ±¥   7▯  -          n  ppt/slideLayouts/_rels/slideLayout10.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  -
    )o  ppt/slideLayouts/_rels/slideLayout11.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          2p  ppt/slideLayouts/_rel
s/slideLayout9.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          :q  ppt/slideLayouts/_rels/slideLayout6.xml.relsPK▯▯- ▯ ▯
  ! ıÐÆ±¥   7▯  ,          Br  ppt/slideLayouts/_rels/slideLayout5.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          J
s  ppt/slideLayouts/_rels/slideLayout4.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          Rt  ppt/slideLayouts/_rels/slideL
ayout3.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ±¥   7▯  ,          Zu  ppt/slideLayouts/_rels/slideLayout2.xml.relsPK▯▯- ▯ ▯  ! ıÐÆ
±¥   7▯  ,          bv  ppt/slideLayouts/_rels/slideLayout7.xml.relsPK▯▯- ▯ ▯  ! âÛ┬▯▯▯  0ô  !          jw  ppt/s
lideMasters/slideMaster1.xmlPK▯▯- ▯ ▯  ! ;mZ╓Y▯  ²o  !          ║É  ppt/slideLayouts/slideLayout1.xmlPK▯▯- ▯ ▯  ! i
ó_!▯▯  Ã  ,          Rª  ppt/slideMasters/_rels/slideMaster1.xml.relsPK▯▯- ▯ ▯  ! _!▯éÔ▯   !          ║º  ppt/sl
▯¯▯  ç▯  !          ▐┤  ppt/slideLayouts/slideLayout5.xmlPK▯▯- ▯ ▯  ! !D¼0╢▯  ¼▯  !          ▒  ppt/slideLayouts/s
▯  on  !          Á  ppt/slideLayouts/slideLayout3.xmlPK▯▯- ▯ ▯  ! ▯EÿGW▯   ▯  !          L╧  ppt/slideLayouts/s
lideLayout2.xmlPK▯▯- ▯ ▯  ! È┐?║¿▯  7t  !          Ô=  ppt/slideLayouts/slideLayout8.xmlPK▯▯- ▯ ▯  ! ▯▯_e¿▯  ¿e  !
          ╓õ  ppt/slideLayouts/slideLayout7.xmlPK▯▯- ▯ ▯  ! ▯┴└¯G▯  6t  !          ░º  ppt/slideLayouts/slideLayou
t9.xmlPK▯▯- ▯ ▯  ! Ä~A:║▯  '▯  "          6▯▯ ppt/slideLayouts/slideLayout10.xmlPK▯▯- ▯ ▯  ! ▯_ì!›▯  ;m  "
   0▯▯ ppt/slideLayouts/slideLayout11.xmlPK▯▯- ▯ ▯  ! ▒éÞÃ▯  Ô▯  ▯          «)▯ ppt/theme/theme1.xmlPK▯▯-
    ! ı╤Ï▯ÖR▯ ÖR▯ ▯          -0▯ ppt/media/image1.jpegPK▯▯- ▯ ▯  ! ²Û▯å┐   %▯  ▯          ╛â▯ ppt/theme/_rels/
theme1.xml.relsPK▯▯-
    ! /å▄  Ç   Ç  ▯          0ä▯ docProps/thumbnail.jpegPK▯▯- ▯ ▯  !  ▯sgò▯  V▯  ▯          Ý▯▯ ppt/viewProps.
xmlPK▯▯- ▯ ▯  ! ?à0qC▯  (▯  ▯          ▒▯▯ ppt/tableStyles.xmlPK▯▯- ▯ ▯  ! ▯-»«Q▯  ï▯  ▯          %   ▯ ppt/pr
esProps.xmlPK▯▯- ▯ ▯  ! ▄p²LÇ▯  ┴▯  ▯
D:\server\mysql-5.7.21-winx64\bin>

D:\server\mysql-5.7.21-winx64\bin>
```

Figure.3[Encrypted File using AES]

Workflow for file encryption

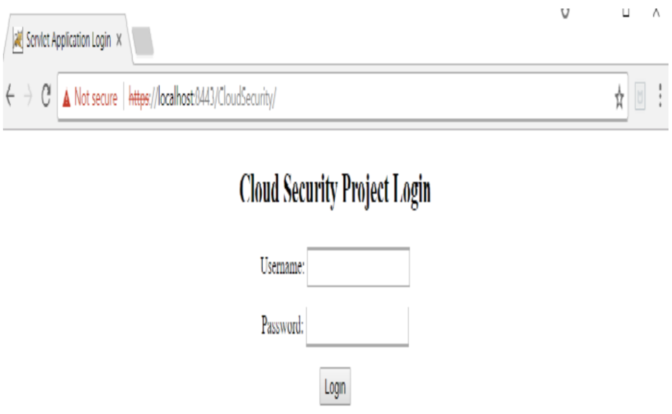
1. When user uploads file, file is being encrypted using AES algorithm and stored to the MySQL database.

2. on request for the file data decrypted and send to the user for authenticated user but for unauthenticated user file in form of encrypted format will display which can not be in human understandable form .The file cannot be decrypted without static key which only authenticate user has.

In User behaviour mapping parameters like upload rate, down load rate, backlisted count will be considered to predict valid or invalid user.

CLIENT: To access cloud from the desktop of the client webapp is required. That webapp is present under client module. Login page, File upload page client, Page through which file data can be checked.

When user gives request to login into the system, the password of the user is being checked with the encrypted password stored in the database. If password matched with the decrypted password. Then user allowed to login into the system to upload or access to the file stored in database which is stored in encrypted format in the database. If user is an authenticate user



Servlet Application Login
Not secure
https://localhost:8443/CloudSecurity/
Cloud Security Project Login
Username:
Password:
Login

Figure.4[Login screen of the user for secure login to the server]

CLOUD SERVER: The server which is responsible for user profile data, and user storage data. Responsible to control user behaviour through the data generated by the behaviour analysis.

AWS Ec2 is used for the cloud server model implementation.

HADOOP is installed to analyse the user behaviour data.

IV.CONCLUSION

Security and privacy of the cloud computing is increased by using disinformation attack of Fog computing Thus, data in cloud is secured by using decoy data in place of original data. It uses SHA1 algorithm to encrypt password, AES algorithm to encrypt and decrypt file uploaded by the user, and Naïve Bayes theorem in HADOOP to detect intruder. In this system instead of two files one encrypted file is stored into the server which saves space and access faster. When user detected as an intruder then encrypted file will be given to the user which cannot be decrypted without key. For authenticated users file being decrypted using static key value of AES algorithm and original file produced to client. Space for storage of file and access time is reduced. This system improves the cloud security by using secure Fog computing.

ACKNOWLEDGEMENT

I would like to thank Mr VijayKumar.R assistant professor, and Dr. Rjalakshmi Ravishankar Head of the Department Computer science and Engineering in New Horizon college of

Engineering for all her support and guidance. The Facility provided helped to complete good quality project work. Your guidance on technology helps me to develop a strong technical system.Aslo I would like to thank facility provided by the institution to complete my project.

REFERENCE

[1]https://www.levelcloud.net/why-levelcloud/cloud-education-center/advantages-and-disadvantages-of-cloud-computing/

[2]http://bigdata-madesimple.com/5-advantages-and-disadvantages-of-cloud-storage/

[3] https://gbhackers.com/top-10-biggest-data-breaches-2017/

[4]S.J.Stolfo,salem,andA.D.Keromytis,‘‘FogComputing:MitigatingInsiderDataTheftAttacksintheCloud,’’inProc.IEEE Symp.Security Privacy Workshops (SPW), May 2012, pp. 125–128.

[5]M. Mukherjee et al. Rakesh Matam, Lei Shu, Lenadros Maglara, Mohamed Amine Ferrag, Nikumani Choudhury AND Vikas Kumar,” Security and Privacy in Fog Computing: Challenges” date of current version October 12, 2017 in IEEE Access

[6]Blesson Varghese, Nan Wang, Dimitrios S. Nikolopoulos, Raj Kumar Buyya,” Feasibility of Fog Computing” arXiv:1701.05451v1 [cs.DC] 19 Jan 2017

[7] Shreya Waghmare, Shruti Ahire, Himali Fegade, Pratiksha Darekar,” Securing Cloud using Fog Computing with Hadoop Framework”, International Journal of Science, Engineering and Technology An Open Access Journal, Shreya Waghmare et al. 2017, Volume 5 Issue 3

[8] Fog Computing: preventing Insider Data Theft Attacks in Cloud Using User Behavior Profiling and Decoy Information Technology

[9] Comparing Naive Bayes, Decision Trees, and SVM with AUC and Accuracy, Jin Huang, Jingjing Lu, Charles X. Ling

[10] FogComputing: Securing the cloud and preventing insider attacks in the cloud. Aatish B. Shah1, Jai Kannan2, Deep Utkal Shah3 Prof. S.B.Ware4, Prof. R.S.Badodekar5 2016

[11] Younghee Park, Salvatore J. Stolfo, Software Decoys for Insider Threat, ACM 2012.

[12] Wael Etaiwi*, Mariam Biltawi and Ghazi Naymat,” Evaluation of classification algorithms for banking customer’s behaviour under Apache Spark Data Processing System ”, Wael Etaiwi et al. / Procedia Computer Science 113 (2017) 559–564