A novel approach for data mining on encrypted data in cloud computing environment

1st Trushita Patel, 2nd Mitesh Thakor
[1] M. Tech Student,,[2]Assistant Professor,
[1] Computer Engineering
[1] Department of Computer Engineering, MEC, Basna. India.
[1]trushpatel.ce@gmail.com, [2]mitesh.thakore@gmail.com

Abstract - Cloud Computing has become a main source for data processing, data storage and distribution .the storage of data is simple and free to use. Data mining is becoming an increasingly important technology for the information society.. The integration of data mining techniques into normal day-to-day activities has become common place. Traditional *Data Storage* systems are not able to handle *Large amount of data* and also analyzing the *Large amount of data* becomes a challenge and thus it cannot be handled by traditional analytic tools. *Cloud Computing* can resolve the problem of handling, storage and analyzing *Large amount of data* as it distributes the big data within the cloudlets. Data Privacy and Data security is one of the major issues while storing the *Large amount of data* in a Cloud environment because sensitive information is centralized into the cloud, so must be required this information must be encrypted and uploaded to cloud for the data privacy and efficient data utilization.

The proposed scheme, guarantees top-n multi keyword retrieval over encrypted cloud data using DES with high Security and practical efficiency using Apriory Algorithm, where in the majority of computing work is done on the server

Index Terms - Cloud Computing, Data Mining, Association rules, Top-n multi keyword.

I. INTRODUCTION

Cloud Computing and Data Mining are the two emerging trends of today in the world of information technology an dcomputing environment. “Cloud computing refers to the web-based computing, providing users or devices with shared pool of resources, information or software on demand and pay per-use basis”. It allows end user and small companies to make use of various computational and distributed resources like storage, software and processing capabilities provided by other companies such as Amazon or Microsoft.
Cloud Services provided by the clouds are broadly divided into three categories: Infrastructure-as-a- Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

Infrastructure as a service (IaaS): In the IaaS model computers are offered as physical or as virtual machines, and other resources.

Platform as a service (PaaS): In the PaaS model, cloud providers offers a computing platform including operating system, programming language execution environment, database, and web server. Without buying and managing hardware and software on a cloud platform.

Software as a service (SaaS): In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients.

“Data mining is the process of extracting useful patterns or knowledge from large databases**”**. Data mining tasks include Classification)Data mining tasks include Classification, Association Rules, Clustering,(, Association Rule)It helps to find out the interesting relationship between the products(, Clustering)is the technique of grouping of several objects unto groups of similar attributes in order to simplify large, complex sets...etc (
Data mining process……

- **Selection:** select data from various resources where operation to be performed.
- **Preprocessing:** also known as data cleaning in which remove the unwanted data.
- **Transformation:** transform /consolidate into a new format for processing.
- **Data mining:** identify the desire result.
- **Interpretation / evaluation:** interpret the result/query to give meaningful report/information.

II. LITERATURE SURVEY

Author at[1] proposed a model where they focus on file searching based on multiple keywords. They argue that normaly all the current keyword based search where file may contain the keyword or not . so, the authors provide scalable systerm with minimize information lekage.their model prevent overload by working at user side for ranking files,where consume less bandwidth. there perform analysis show efficiency of their proposed solution.

In this paper [2] Author work on effectively collobrative outsourced data mining process with multi owner for provide security in cloud computing . So, the author focusing on utilized trusted computing technology to detect malicious cloud assumption. They analyses different- different technology for security and efficiency of data in cloud .that the author prove the correctness of framework with KNN,K-Means and SVN in outsourced collaborative environment .so finally they provide analysis for security and efficiency of framework.

Author proposed for the same [3] but In this paper author focuses on Homomorphic Encryption for secure data mining in cloud .they argue that normally all current issue like traditional system can't be able for storing and large amount of data .so the author provide easy solution in cloud with secure data mining technique. Authors propose a secure K-Means data mining approach on data may be distributed among no of host security and privacy of the data. This paper able for represent the pallier Homomorphic encryption system and perform analysis of K-Means result in cloud computing.

This Paper Author [4] work on secure semantic search using Query Keyword. base on keyword co-occurrence probability and semantic relationship library. They proposed scheme for not only match exact files but work on Query keyword. so, author produce architecture using query keyword technique for cloud. (A) Store SRL in private cloud (B) Retrieve index on public cloud We also derive the one to many OPE technique to protect related score and compute total Score.

Author [5] focuses on remotely stored encrypted data store in cloud computing .so, author argue that normally all the traditional searchable encryption .Other Produced method for searchable encrypted data using keyword is contain or not in file using Boolean search method. But this author proposed framework for secure rank keyword search over large amount of encrypted data file in cloud. so, author also define crypto primitive OPSE and one to more order preserving mapping for retrieve efficient data from cloud.

COMPARISON OF VARIOUS RESEARCH SCHEMES

The table below shows a short comparison about the various schemes proposed by a researcher by taking different parameters. The table gives the description about the basic technique used with the benefits that researcher gets the limitations found in schemes.

Criteria Group →	**Encryption/ Data Mining/Cloud Computing oriented measures**							Others
Individual Criteria → Providers ↓	**Cloud Computing**	**Encrypted Data**	**Data Mining**	**Data Security**	**Keyword based**	**Secure SQL Query**	**Secure Searching**	**Less Overhead**
[1]	√	√	√	√	√	X	√	X
[2]	√	X	√	√	X	X	X	X
[3]	√	√	√	√	X	√	X	X
[4]	√	√	√	√	X	X	X	X
[5]	√	√	X	√	√	X	√	√
[6]	√	√	X	√	√	X	√	X

Table 1. Comparison study

III. CONCLUSIONS

In this paper we study how datamining is used on encryptred data in cloud computing.so we proposed framework to make syatem highly scalable, costly less, secure and highly perfomance orientad. In our proposed scheme, ,bination approch of Data miningcom Encryptin and Cloud computing for reduce overhead on client side. method like DES As a case of study classical, ,for Encryption ngApriory for Data mini. so proposed solution is secure,scalable, less costly,highly performance oriented compared to the other technique of data mining in cloud computing.

IV. REFERENCES

[1] D.Pratiba1, Dr.G.Shobha and Vijaya Lakshmi.P.S ,"Efficient Data Retrieval From Cloud Storage Using Data Mining Technique" in International Journal on Cybernetics & Informatics (IJCI) Vol. 4, No. 2 , April 2015 , pp. 271-279.

[2] Yila Huangy, Qiwei Luz, Yan Xiongz ,"Collaborative Outsourced Data Mining for Secure Cloud Computing" on JOURNAL OF NETWORKS, VOL. 9, NO. 10, OCTOBER 2014, pp. 2655- 2664.

[3] Deepti Mittal, Damandeep Kaur, Ashish Aggarwal ," Secure Data Mining in Cloud using Homomorphic Encryption".

[4] Xingming Sun, Yanling Zhu, Zhihua Xia and Lihong Chen ," Privacy- Preserving Keyword-based Semantic Search over Encrypted Cloud Data ", International Journal of Security and Its Applications Vol.8, No.3 (2014), pp.9-20.

[5] Ms. M. R. Girme, Prof. G.M. Bhandari ," Efficient Ranked Keyword Search For Achieving Effective Utilization Of Remotely Stored Encrypted Data In Cloud", in *IJAIEM* Volume 3, Issue 6, June 2014 ISSN 2319 – 4847 .pp. 105-113.
[6] AbdulRahman R. Alazmi & AbdulAziz R. Alazmi ," Data Mining And Visualization of Large Databases" , in International Journal of Computer Science and Security (IJCSS), Volume (6) : Issue (5) : 2012, pp.295-314.

[7] Smita, Priti Sharma "Use of Data Mining in Various Field: A Survey Paper", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 3, Ver. V (May-Jun. 2014), PP 18-21.
[8] CH.Sekhar, S Reshma Anjum" Cloud Data Mining based on Association Rule" CH.Sekhar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, pp 2091-2094.
[9] Fei Long" Research on algorithms of data mining under cloud computing environment" Journal of Chemical and Pharmaceutical Research, 2014, 6(7): pp 1152-1157.
[10] B. Kamala "A Study On Integrated Approach Of Data Mining And Cloud Mining" International Journal Of Advances In Computer Science And Cloud Computing, Issn: 2321-4058.
[11] Mr.A.Srinivas, M. Kalyan Srinivas, A.V.R.K.Harsha Vardhan Varma "A Study On Cloud Computing Data Mining" International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 5, July 2013.
[12] Zeba Qureshi, Jaya Bansal, Sanjay Bansal "A Survey on Association Rule Mining in Cloud Computing", International Journal of Emerging Technology and Advanced Engineering 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 4, April 2013.
[13] N. Janardhan ,T. Sree Pravallika , Sowjanya Gorantla," An Efficient Approach For Integrating Data Mining Into Cloud Computing", International Journal of Computer Trends and Technology (IJCTT) - volume4 Issue5–May 2013.
[14] Ruxandra-Ştefania PETRE," Data mining in Cloud Computing", Database Systems Journal vol. III, no. 3/2012 pp.67-71.