



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 3, March 2017

A Fair Solution for Outsourced Data Mining rule on Secure Data in Cloud Computing

Maitri Bhavsar¹, Jayesh Mevada²

M. Tech Student, Department of Computer Engineering, MEC, Basna, India¹

Assistant Professor, Department of Computer Engineering, MEC, Basna, India²

ABSTRACT: Now a day's user of computer system need everything on hand without location dependency with least cost and efficiently. Here the important term Cloud computing comes with lots of benefits. The user store lots of information on Cloud server and able to access from anywhere, anytime. With this important advantage this technology also having some issues like, Security of user information, efficiently fetch information from the various large data and many more. So, to increase the use of Cloud and to deal with the various issue related to cloud environment as mentioned above we are here try to proposed a efficient secure searching mechanism which user can have all the details handy without getting any type of burden. The scheme will provides efficient searching, Group based retrieval, efficient mining technique and more which not only increase trust of user on Cloud as well reduce the time taken for related to searching of important information.

KEYWORDS: cloud computing, datamining, association rule, encryption, improvised apriori algorithm.

1. INTRODUCTION

Cloud computing is the delivery of computing services over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications. Cloud computing is receiving a great deal of attention, both in publications and among users, from individuals at home to the U.S. government. Yet it is not always clearly defined. Cloud computing is a subscription-based service where you can obtain networked storage space and computer resources. One way to think of cloud computing is to consider your experience with email. Your email client, if it is Yahoo!, Gmail, Hotmail, and so on, takes care of housing all of the hardware and software necessary to support your personal email account. When you want to access your email you open your web browser, go to the email client, and log in. The most important part of the equation is having internet access. Your email is not housed on your physical computer; you access it through an internet connection, and you can access it anywhere. If you are on a trip, at work, or down the street getting coffee, you can check your email as long as you have access to the internet. Your email is different than software installed on your computer, such as a word processing program. An email client is similar to how cloud computing works. Except instead of accessing just your email, you can choose what information you have access to within the cloud.

II. BACKGROUND THEORY

We are in an age often referred to as the information age. In this information age, because we believe that information leads to power and success, and thanks to sophisticated technologies such as computers, satellites, etc., we have been collecting tremendous amounts of information. Initially, with the advent of computers and means for mass digital storage, we started collecting and storing all sorts of data, counting on the power of computers to help sort through this amalgam of information. Unfortunately, these massive collections of data stored on disparate structures very rapidly



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

became overwhelming. This initial chaos has led to the creation of structured databases and database management systems (DBMS)..

A. Why Data mining on cloud?

Data mining in cloud computing is the process of extracting structured information from unstructured or semi-structured web data sources. Efficient analysis. It generate faster and proper output according to user demand. User can store and retrieve meaningful information from cloud at anytime from anywhere.

III. RELATED WORK

Authors of [1] Proposed technique executes an encryption scheme, called AES that is based on 1–1 substitution ciphers for items and adding fake transactions to make each cipher item share the same frequency as $\geq k-1$ others. It makes use of a compact synopsis of the fake transactions from which the true support of mined patterns from the server can be efficiently recovered. This technique also proposed a strategy for incremental maintenance of the synopsis against updates consisting of appends and dropping of old transaction batches. Unlike previous works, proposed technique formally proved that our method is robust against an adversarial attack based on the original items and their exact support.

Authors of [2] an overview of the necessity and utility of data mining in cloud computing. Cloud computing offers benefits for organizations and individuals. There are also privacy and security concerns. If you are considering a cloud service, you should think about how your personal information, and that of your customers, can best be protected. Actually we are discussing the cloud computing data mining for the advance use of security in data loss purpose. While the data we are storing in cloud is being separated in different servers for a security but the hackers using the cheap and raw cloud computing for the misuse of the software.

Authors of [4] In this paper, the memory space is drastically reduced when large number of transactions are performed from the data warehouses and repositories and an improvised Apriori is proposed by reducing the time consumed in transactions scanning for candidate itemsets and also by reducing the number of transactions to be scanned. Whenever the k of k -itemset increases, the gap between our improved Apriori and the original Apriori increases from view of time consumed, and whenever the value of minimum support increases, the gap between our improved Apriori and the original Apriori decreases from view of time consumed. The time consumed to generate candidate support count in our improved Apriori is less than the time consumed in the original Apriori; our improved Apriori reduces the time consuming by 67.87%. Hence, this approach is far more efficient than the original apriori algorithm as it uses the approach of parallel algorithm and clustering method by which the memory space is reduced and it can be successfully used in the real time applications especially in the library as it can save a lot of time by giving all the information about those books which are frequently read.

Authors of [6] Cloud Computing provides storage of data in a server by protecting data by using data mining concept. Actually, we are discussing the cloud computing data mining for the advance use of security in data loss purpose. In Cloud computing, the data is being shifted from one server to another server in a peer to peer transaction. Data mining technologies provided through Cloud Computing is an absolutely necessary characteristic for today's businesses to make proactive, knowledge driven decisions as it helps them have future trends and behaviors predicted.

COMPARISON OF VARIOUS RESEARCH SCHEMES

We have made exhaustive survey on various recent research papers relating to Association Rule in cloud environment. Table shows the comparison of cloud computing approaches.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 3, March 2017

Ref - no.	Cloud Computing	Associati-on rule	Encryption& Decryption	Data -mining - technique	Privacy-preservi-ng outsourcing	Authenti cati-on
[1]	NO	YES	NO	YES	YES	YES
[2]	NO	YES	YES	NO	YES	NO
[3]	YES	YES	YES	NO	YES	NO
[4]	YES	NO	YES	NO	NO	YES
[5]	YES	NO	YES	YES	NO	YES
[6]	YES	YES	NO	YES	NO	YES

Table 1. Comparison study

IV. OUR CONTRIBUTION

Various researchers have worked to achieve better shown in figure 4.

4.1 Proposed Work

The proposed system flowchart based on **Uploading Data on Cloud** and **Retrieve Data from Cloud base on Query** in Fig.

(a) Uploading Data on Cloud

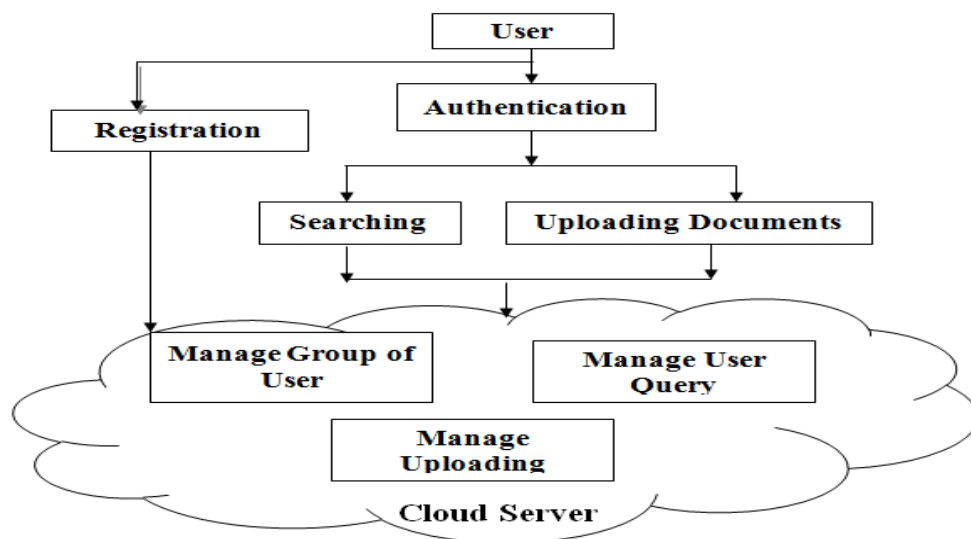


Fig. 2. Uploading data in cloud process.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 3, March 2017

In Fig 2 shown we propose a process for Upload data on Cloud. So. First any user try to login if login successfully then this user eligible for upload data. But here before data is store in Database it must be convert in encrypted form After perform encryption on uploading file, then it store in database and modify database. At last server send reply to user for user conformation.

(b) Retrieve Data from Cloud base on Query

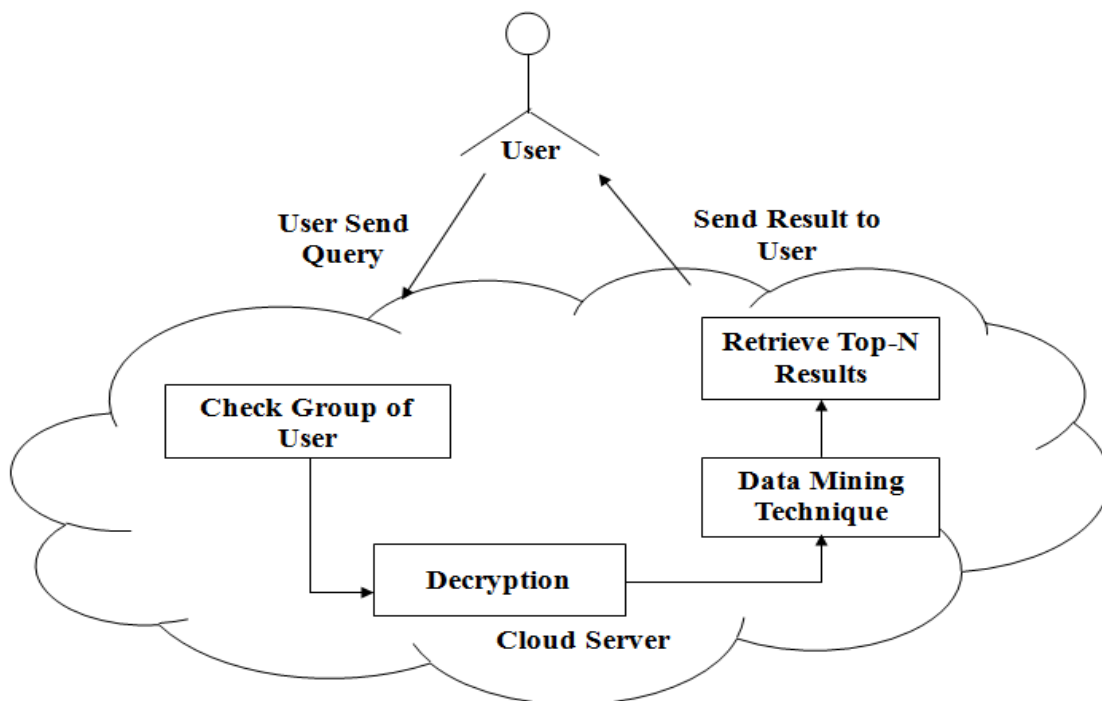


Fig. 3. Data retrieving and mining process.

In this process if user can retrieve any data from cloud then it must be performed logging process. If user can't log, then do not retrieve data from cloud. After perform logging process then write send query with specific keyword to server. Server received user query and keyword and find data in the server. After finding large amount of data then apply data mining technique on this data. At last server performing data mining process on data and send result to user.

V. RESULTS

[1] Data Set Uploading Time

Table 5.1 shows the simulation results for uploading time. Fig. 5.9 shows the Uploading file time chart, in existing system has more time for uploading a data. But by using the proposed system, we can reduce the uploading time.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

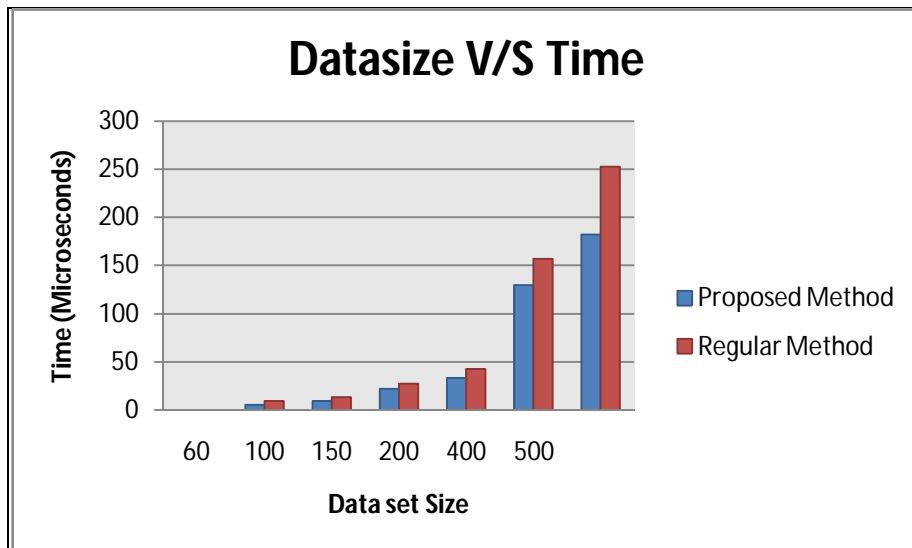
Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Table : 5.1 Data Uploading Time

Dataset size (file)		
Data size	Proposed Method	Regular Method
60	5.173849	9.542035
100	9.281112	13.40456
150	21.968036	27.551625
200	33.361004	42.393826
400	129.911101	156.902231
500	182.576441	252.693603

Data uploading File Chart



[2] Support Time

Table 5.2 shows the results for Support time. Fig. 5.10 shows the no. of support time chart, in existing system has support as compared to proposed system.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

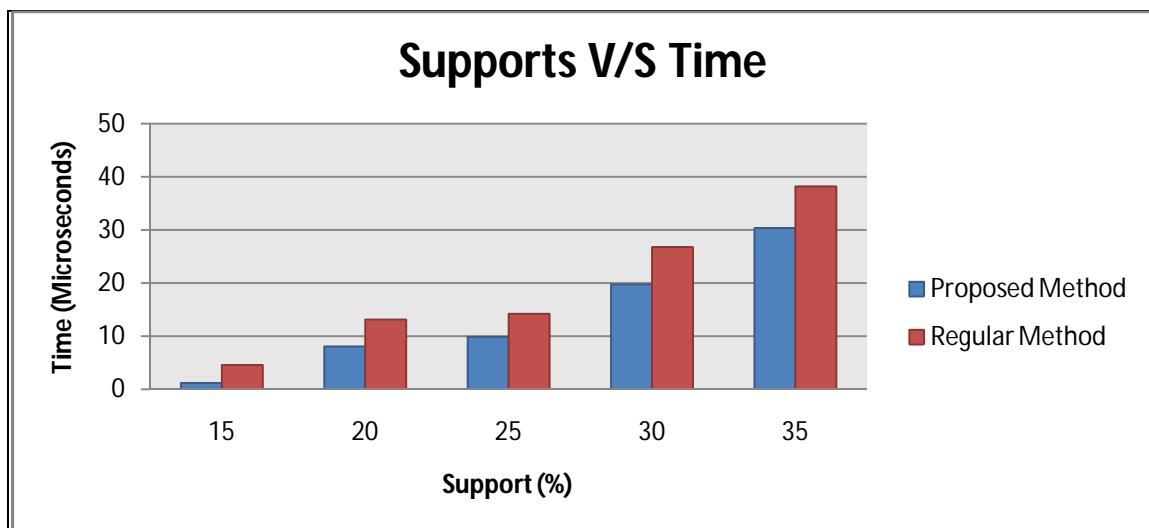
Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

Table : 5.2 Support Time

Support time		
Supports	Proposed Method	Regular Method
15	1.192577	4.530819
20	8.040051	13.126228
25	9.937169	14.201934
30	19.762233	26.700172
35	30.418617	38.236826

Support Time Chart



[3] Time Consumption

Table 5.3 shows the simulation results for time consumption. Fig. 5.11 shows the time consumption chart which input different data file and get the time consumption. But by using the proposed system, we can reduce the time consumption as compared to existing system.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

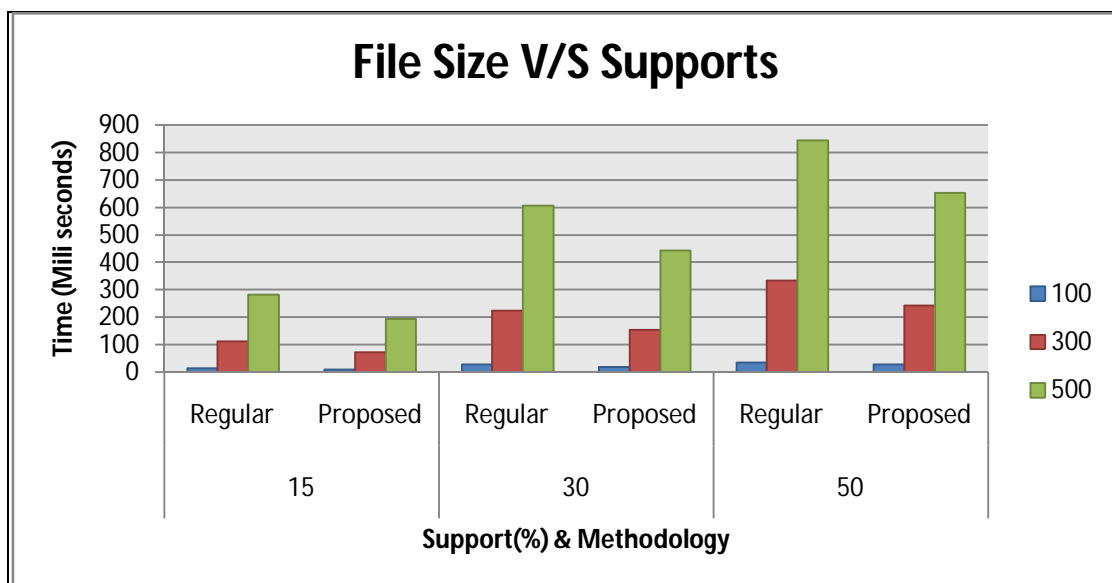
Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Table : 5.3 Time Consumption

Time Consumption						
Support	15		30		50	
File Size	Regular	Proposed	Regular	Proposed	Regular	Proposed
100	13.038242	9.735852	28.850318	19.642627	35.550624	28.89295
300	111.220191	72.88903	224.636711	153.868229	332.656355	243.222226
500	283.007615	194.009277	605.554299	442.559695	844.541366	653.940269

Time Consumption Chart



VI. CONCLUSION

Here we used Method of Data Mining for group of user which will take less searching time and fast retrieval of data. Proposed scheme used for Secure and Mine large data set(Scalable).System provides higher performance with low cost on cloud environment.To retrieve data set in a human-understandable language Provides lightness(less overhead) and fast access technology Provide Multi keyword based searching technique. To achieving high storage, efficient Data Retrieval and computing services at a low cost level. We used one of the existing encryption/decryption method for faster execution of proposed scheme.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 3, March 2017

REFERENCES

- [1] Swapnil C. Salunke , Prof. Rphshani Raut(Ade) , “Encrypted association rule mining for outsourced data mining” volume 5, issue 11, November 2015.
- [2] Miss Rohini A. Dhote , Dr. S. Deshpande , “Data mining with cloud Computing - An Overview” . International journal of advanced research in computer engg.& technology (IJARCET) VOLUME 5 ISSUE1, January 2016.
- [3] Fang Liu, Wee Keong Ng, Wei Zhang, “Encrypted association rule mining for outsourced data mining”. 2015 IEEE 29th International conference on advanced Information Networking & Application.
- [4] Sudhakar Singh, Pankaj Singh, Rakhi Garg, P.K. Mishra, “Mining Association Rules In Various Computing Environments: A Survey” , ISSN9973-4562 volume11 , Number 8(2016) pp 5629-5640.
- [5] Akshita Bhandari , Ashutosh Gupta, Debasis Das, “Improvised Apriori algorithm using frequent pattern tree for real time application in data mining” ELSEVIER, (ICICT 2014) .
- [6] CH.Sekhar, S Reshma Arjum , “ Cloud Data Mining Based on Association Rule” (IJCSIT) International Journal of Computer Science and Information Technology Vol.5 (2) .2014 2091-2094 .
- [7] Fenghua Liu, “Research On Data mining algorithm based on business cloud platform for mobile internet”, Vol.9, No. 2 (2016), pp.343-352.
- [8] A. Beloglazov and R. Buyya “Energy efficient resource management in virtualized cloud data centers,” In Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing. IEEE Computer Society, 2010, pp. 826-831.
- [9] A. Hameed, A. Khoshkbarfoushha, R. Ranjan, P. P. Jayaraman, J. Kolodziej, P. Balaji, and A. Zomaya, “A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems,” Springer, 2014, pp. 1-24.
- [10] Y. Chen, D. Gmach, M. Arlitt, M. Marwah, and A. Gandhi “Minimizing data center SLA violations and power consumption via hybrid resource provisioning,” In: Second international green computing conference (IGCC 2011), IEEE 2011, pp. 1–8.
- [11] C. Goldman, M. Reid, R. Levy, and A. Silverstein “Coordination of energy efficiency and demand response,” Environmental Energy Technologies Division, Berkeley National Laboratory, 2010, pp. 1-74.
- [12] S. Srikantaiah, A. Kansal, and F. Zhao “Energy aware consolidation for cloud computing,” In: Conference on power aware computer and systems, 2008.
- [13] A. Verma, P. Ahuja, and A. Neogi “pMapper: power and migration cost aware application placement in virtualized systems,” In: 9th ACM/IFIP/USENIX international conference on middleware (Middleware’08), 2008, pp. 243–264.
- [14] Y. C. Lee and A. Y. Zomaya, “Energy efficient utilization of resources in cloud computing systems,” Journal of Supercomputing, 60(2), Springer 2012, pp. 268-280.
- [15] A. Beloglazov, J. Abawajy, and R. Buyya “Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing,” Future generation computer systems, 28(5), ScienceDirect 2012, pp. 755-768.
- [16] H. M. Lee, Y. S. Jeong, and H. J. Jang, “Performance analysis based resource allocation for green cloud computing,” The Journal of Supercomputing, 69(3), Springer 2014, pp. 1013-1026.
- [17] A. Horri, M. S. Mozafari, and G. Dastghaibifard, “Novel resource allocation algorithms to performance and energy efficiency in cloud computing,” The Journal of Supercomputing, 69(3), Springer 2014, pp. 1445-1461.
- [18] “Greencloud,” Available at: <https://greencloud.gforge.uni.lu/>
- [19] “Xen,” Available at: <https://en.wikipedia.org/wiki/Xen>
- [20] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. D. Rose, and R. Buyya, “CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms,” Software: Practice and Experience, Wiley Press, New York, USA, 2010.
- [21] “Eclipse,” Available at: [http://en.wikipedia.org/wiki/Eclipse_\(software\)](http://en.wikipedia.org/wiki/Eclipse_(software)).