iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** XII **Month of publication:** December 2016

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Secure Outsourced Comparison Scheme for Partitioned Databases

Neethu Rosiline[1], T Sivakumar[2], Sreedarsa K. S[3]
[1]*PG Student, Maharaja Institute of Technology, Coimbatore*
[2]*M.E.(CSE), Professor, Maharaja Institute Of Technology, Coimbatore*
[3]*PG Student, Maharaja Institute of Technology, Coimbatore*

Abstract: In different applications that uses cloud, studies a range of data analysis techniques. The most popular data analysis techniques were association rule mining and frequent itemset mining. In this paper we focus on outsourced comparison scheme for partitioned databases. This paper mainly concern about data security. It prevents accessing of shared data among different users. We provide homomorphic encryption scheme and a comparison scheme for data privacy. Association rule mining can be used for effectively getting data from portioned databases. Resource consumption can be reduced by the usage of cloud servers for data and computing.
Index Terms- Association rule mining, frequent itemset mining, outsourced comparison scheme.

I. INTRODUCTION

In a centralized database raw data is mined centrally by classic frequent item set mining and association rule mining algorithms,[13,14].In this privacy is not considered. Due to an enlarged understanding of the significance of data privacy, a number of privacy-preserving mining solutions have been proposed in recent times. In this there are different data owners' aims to study association rule or frequent item sets from their joint data. However, the data owners are not willing to send their raw data to a central site due to privacy concerns. If each data owner has one or more rows (i.e. transactions) in the joint database, we say that the database is horizontally partitioned [1]. If each data owner has one or more columns in the joint database, the database is considered vertically partitioned [2]. This paper focuses on vertically partitioned databases. Vertically partitioned databases are helpful for market basket analysis [8] health care [9], web usage mining [10] and bioinformatics [11].

For example, different businesses, such as a fashion designer and a luxury watch designer, sell different products to the same community. These businesses collaborate to mine customer buying patterns from the joint database. A transaction of the database contains the products that a customer had bought from one or more of the participating businesses, and attributes such as the customer credit card number and date of purchase are used as TIDs. Therefore, each of the businesses (i.e. data owners) will own some transaction partitions in the joint database. However, these businesses may not wish to disclose such data, which include trade secrets (e.g. there may be other competing businesses sharing the same joint database) and customer privacy (e.g. due to regulations in existing privacy regime). Therefore, a privacy-preserving mining solution must be applied. Other use cases can also be found in areas such as automotive safety and national security.

Our solutions are designed for outsourced databases that allow multiple data owners to efficiently share their data securely without compromising on data privacy [3]. Our solutions leak less information about the raw data than most existing solutions. In comparison to the only known solution achieving a similar privacy level as our proposed solutions, the performance of our proposed solutions is 3 to 5 orders of magnitude higher.

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

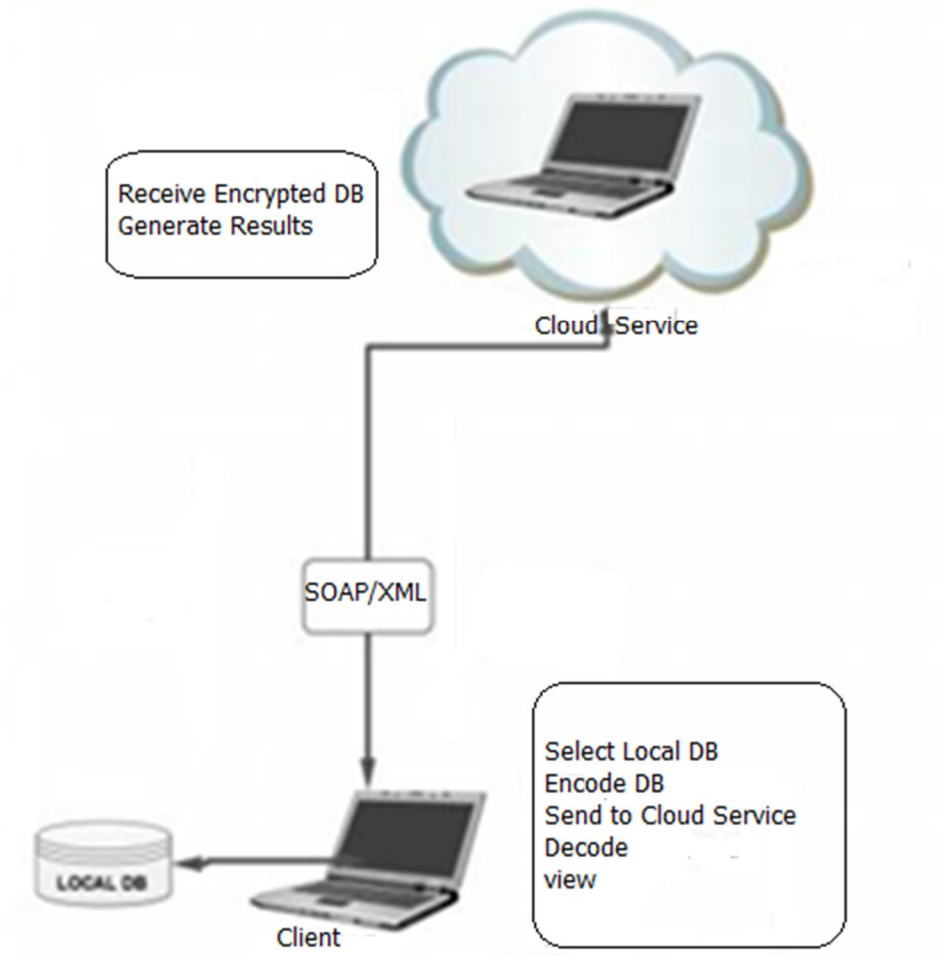


Figure1 Overall diagram

A. System Details

1) Existing System: All existing solutions, do not utilize a third-party server to compute the mining result (fig1). Some solutions use asymmetric homomorphic encryption to compute the supports of itemsets. While other uses a secure scalar product protocol, a set intersection cardinality protocol or a secret sharing scheme to perform these computations. Most of these solutions give exact support to all data owners, resulting leakage of information about data owners raw data. This can be avoided by using the frequent item set mining. The frequent item set mining [5] does not expose exact support to all data owners. But by using association rule we can't mine the result if the exact support cannot computed. Frequent itemsets play an essential role in many data mining tasks that try to find interesting patterns from databases, such as association rules, correlations, sequences, episodes, classifiers and clusters. The mining of association rules is one of the most popular problems of all these. The identification of sets of items, products, symptoms and characteristics, which often occur together in the given database, can be seen as one of the most basic tasks in data mining.

2) Proposed System: Proposed a solution based on k-anonymity frequency. To counter frequency analysis attack. The data owner inserts fictitious transactions in the encrypted database to conceal the item frequency. After inserting the fictitious transactions, any item in the encrypted database will share the same frequency with at least k−1 other items. The data owner sends the encrypted database of both the real and fictitious transactions to the cloud. Finally, the data owner decrypts the received item sets with the revised supports higher than the frequency threshold, and generates association rules based on found frequent itemsets [5]. Our solutions use their techniques to conceal the raw data from the cloud and mitigate frequency analysis attack that can be undertaken by the cloud. Privacy-preserving outsourced association rule mining solution[6] based on predicate encryption. This solution is resilient to chosen-plaintext attacks on encrypted items, but it is vulnerable to frequency analysis.

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

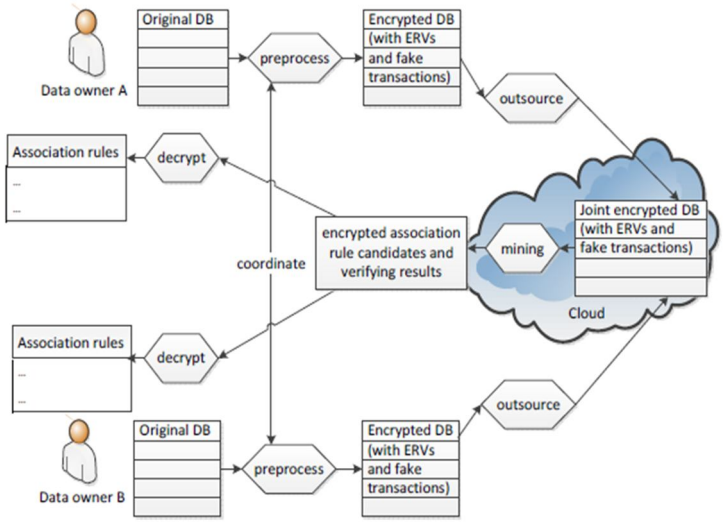


Figure 2 Privacy-preserving outsourced association rule mining.

a) *Proposed homomorphic encryption scheme:* Homomorphic encryption scheme allows one or more plaintext operations (e.g addition and multiplication) to be carried out on the cipher texts. If the addition operation is allowed, then the scheme is known as additive homomorphic encryption [7]. Existing homomorphic encryption schemes are generally asymmetric. In this paper, we propose a symmetric homomorphic encryption scheme (using only modular additions and multiplications), which is significantly more efficient than asymmetric schemes.

b) *Proposed Secure Outsourced Comparison Scheme:* The proposed secure comparison scheme is based on the symmetric homomorphic encryption scheme[7]. In our privacy-preserving data mining solutions, data owners require the cloud to compare supports/confidences with thresholds. However, both supports and confidences must be kept secret from the cloud and data owners, while the comparison results must be kept secret from the cloud.

c) *Privacy-preserving outsourced association rule mining:* Privacy-preserving outsourced association rule mining (fig 2) solution based on predicate encryption [6]. This solution is flexible to chosen-plaintext attacks on encrypted items, but it is vulnerable to frequency analysis attacks. Applying this solution to vertically partitioned databases will also result in the leakage of the exact supports to data owners.

II. SYSTEM ARCHITECTURE

The system architecture (fig 3) is consists of two or data owners and a cloud. Each data owner has a private database. The data owner encrypts their private data and outsourcing the cloud. The cloud mine association rules or frequent itemsets from the joint/merged database on data owner request [3, 5]. The cloud stores the databases received from different data owners, verifying the database, and the sending of the mining result to relevant data owners. [4]

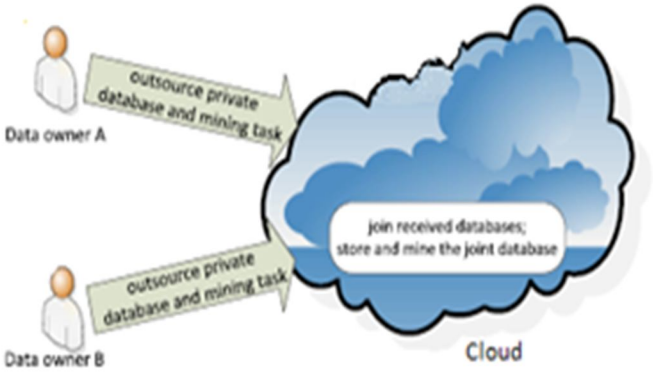


Figure 3 System architecture of outsourced data mining on joint data base.

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

There are five different varieties of modules present in the architecture. Authentication Module, Encryption one Module, Encryption and Cloud Upload Module, Mining and Verifying Result Module and Decryption Module. The data owners can work along with the details of modules mentioned below,

A. *Authentication Module*
   1) Login validation of owner.
   2) Registration.
   3) File upload.
   4) User Activation.

B. *Encryption One Module*
   1) Encryption one by User A.

C. *Encryption and Cloud Upload Module*
   1) Encryption two by User B.
   2) Merging with first encrypt.
   3) Cloud Upload.

D. *Mining and Verifying Result Module*
   1) Mining data from cloud.
   2) Verifying result.

E. *Decryption Module*
   1) Decrypt.
   2) Design Goals

The goals of our secure outsourced comparison scheme for partitioned databases are as follows:

a) *Privacy*. Data owners get small information regarding databases of other data owners. Each raw transaction details of users were concealed. And mining result should be verified by the cloud.

b) *Efficiency*. The data owners have different keys for encryption, encrypted files were merged and uploaded and with both data owner keys only we can decrypt the data. This process may be time consuming but an efficient one.

III. RESULTS

The results of outsourced comparison scheme for partitioned databases are given below:

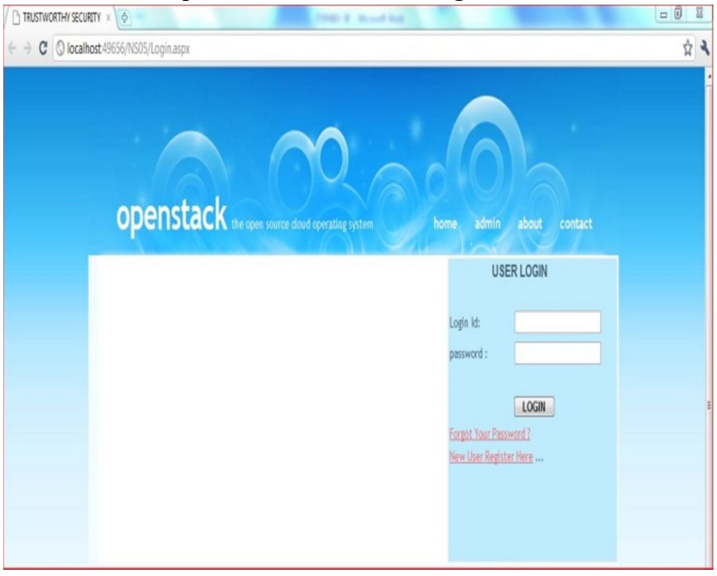


Figure 4: Login page

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

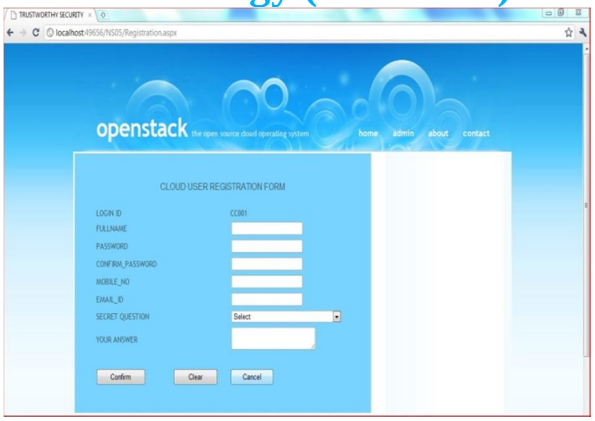


Figure 5: Registration page

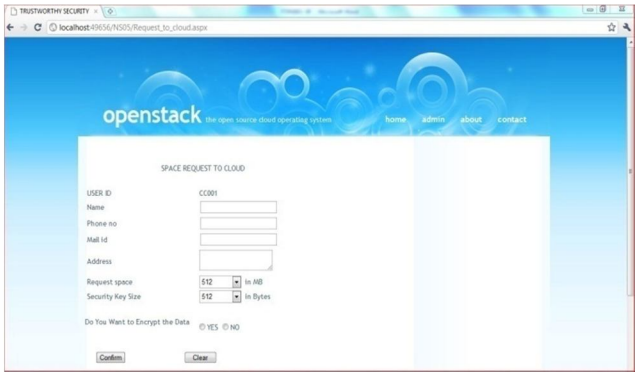


Figure 6: Contact Cloud and payment page

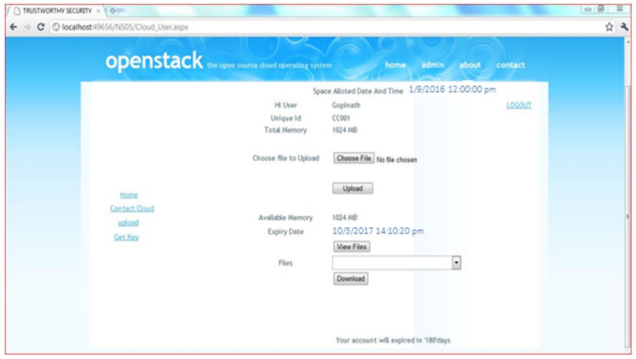


Figure 7: User profile page

IV. FUTURE WORK

An efficient homomorphic encryption scheme and a secure outsourced comparison scheme were presented in this paper. Both schemes have possible usage in other secure computation applications, such as secure data aggregation [15]. Signifying the service of the proposed homomorphic encryption scheme and outsourced comparison scheme in other settings will be the focus of future research.

In the first pass, the algorithm counts occurrence of items (attribute-value pairs) in the dataset, and stores them to 'header table'. In the second pass, it builds the FP-tree structure by inserting instances. Items in each instance have to be sorted by descending order of their frequency in the dataset, so that the tree can be processed quickly. Items in each instance that do not meet minimum coverage threshold are discarded.

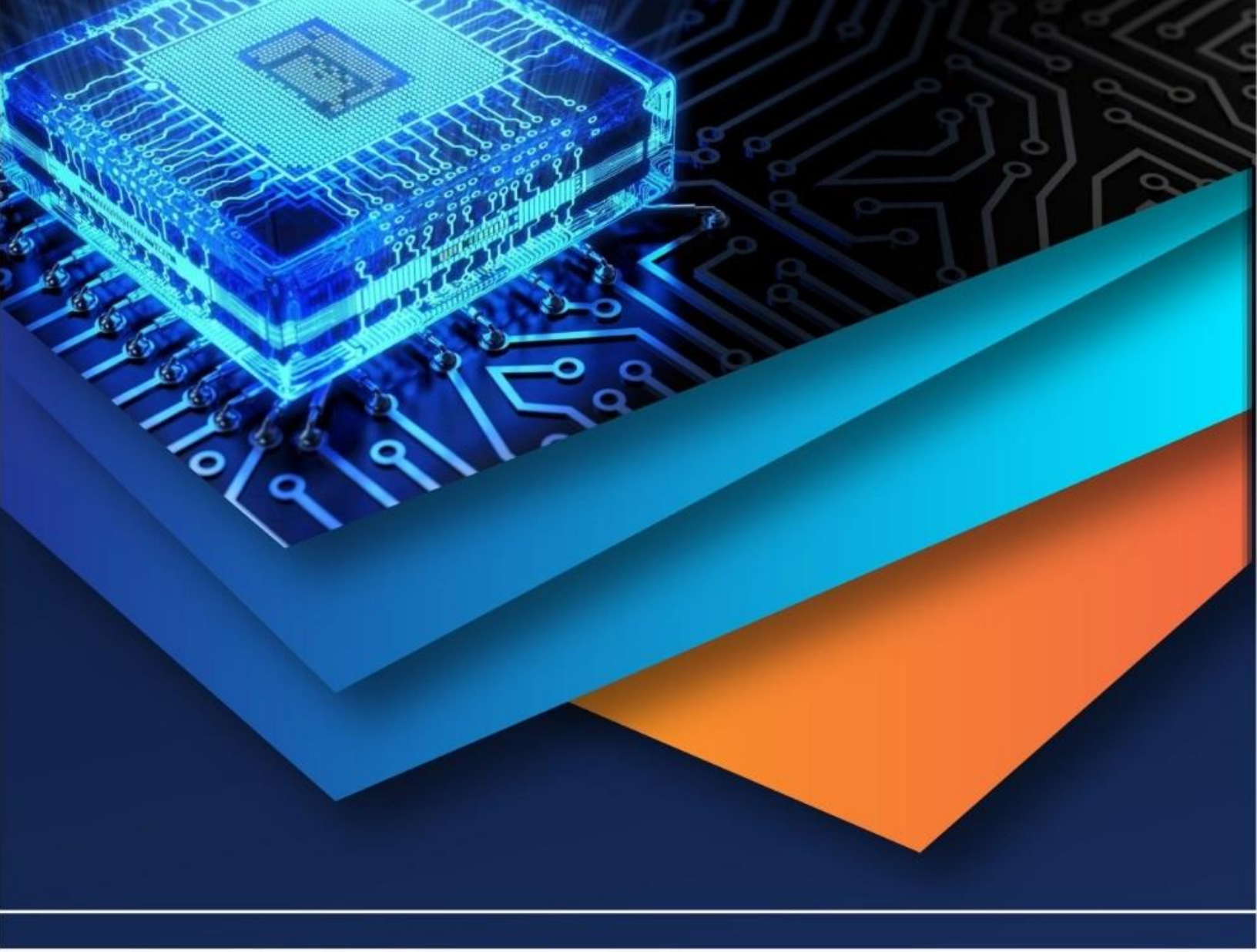
V. CONCLUSION

The concepts of secure outsourced comparison scheme on partitioned databases provide one more data owners to outsource the data

securely. One data owner will encrypt database using a key and another data owner encrypt the database using another key. Both the results were merged and verified by the cloud can be uploaded. The uploaded data can be decrypted by different data owners with their own key.

REFERENCES

[1] M. Kantarcioglu and C. Clifton, "Privacy-preserving distributed mining of association rules on horizontally partitioned data," IEEE Trans. Knowl. Data Eng., vol. 16, no. 9, pp. 1026–1037, Sep. 2004.
[2] B. Rozenberg and E. Gudes, "Association rules mining in vertically partitioned databases," Data Knowl. Eng., vol. 59, no. 2, pp. 378–396,2006.International Journal of Computer Engineering and Applications, Volume 8, Issue 1, oct 2014.
[3] F. Giannotti, L. V. S. Lakshmanan, A. Monreale, D. Pedreschi and H. Wang, "Privacy-preserving mining of association rules from outsourced transaction databases", IEEE Syst. J., vol. 7, no. 3, pp. 385-395, 2013.
[4] B. Dong, R. Liu and H. Wang, "Result integrity verification of outsourced frequent itemset mining",Proc. 27th Annu. IFIP WG Conf. Data Appl. Secur. Privacy (DBSec), pp. 258-265,
[5] R. Liu and H. Wang, "Result integrity verification of outsourced privacy preserving frequent itemset mining", Proc. SIAM Int. Conf. Data Mining, pp. 244-252,
[6] v F. Giannotti, L. V. S. Lakshmanan, A. Monreale, D. Pedreschi, and W. Wang, "Privacy-preserving data mining from outsourced databases," in Proc. CPDP, 2011, pp. 411–426.
[7] Iram Ahmad and Archana Khandekar "Homomorphic Encryption Method Applied to Cloud Computing" International Journal of Information & Computation Technology. ISSN 0974-2239 Volume 4, Number 15 (2014), pp. 1519-1530 © International Research Publications House.
[8] T. Brijs, G. Swinnen, K. Vanhoof, and G. Wets, "Using association rules for product assortment decisions: A case study," in Proc. SIGKDD, 1999, pp. 254–260.
[9] S. E. Brossette, A. P. Sprague, J. M. Hardin, K. B. Waites, W. T. Jones, and S. A. Moser, "Association rules and data mining in hospital infection control and public health surveillance," J. Amer. Med. Inform. Assoc., vol. 5, no. 4, pp. 373–381, 1998.
[10] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Effective personalization based on association rule discovery from Web usage data," in Proc. WIDM, 2001, pp. 9–15.
[11] C. Creighton and S. Hanash, "Mining gene expression databases for association rules," Bioinformatics, vol. 19, no. 1, pp. 79–86, 2003.
[12] X. Yin and J. Han, "CPAR: Classification based on predictive association rules," in Proc. SIAM SDM, 2003, pp. 1–5.
[13] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in Proc. VLDB, 1994, pp. 1–13.
[14] M. J. Zaki, "Scalable algorithms for association mining," IEEE Trans.Knowl. Data Eng., vol. 12, no. 3, pp. 372–390, May/Jun. 2000.
[15] N.Karthick and X.Agnes Kalrani "A Survey on Data Aggregation in Big Data and Cloud Computing" International Journal of Computer Trends and Technology (IJCTT) – volume 17 number 1 – Nov 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY