Scientific Journal Impact Factor (SJIF) : 4.364

International Journal of Modern Trends in Engineering and Research

IJMTER

ISSN (ONLINE) : 2349 - 9745

ISSN (PRINT) : 2393 - 8161

PRIVACY - PRESERVING OUTSOURCED ASSOCIATION RULE MINING USING APRIORI ALGORITHM

Pooja Somwanshi[1], Rohini Sonawane[2], Twinkle Patil[3], Diksha Siksure[4] and Professor K.C.Kulkarni[5]

[1,2,3,4,5] *Computer Department, R.H.Sapat College Of Engineering, Nashik*

Abstract— Association rule mining and frequent itemset mining are two popular and widely studied data analysis techniques for a range of applications. We focus on privacy preserving mining on vertically partitioned databases. In such a scenario, data owners wish to learn the association rules or frequent itemsets from a collective dataset, and disclose as little information about their (sensitive) raw data as possible to other data owners and third parties. To ensure data privacy, we design an efficient homomorphic encryption scheme and a secure comparison scheme. We then propose a cloud-aided frequent itemset mining solution, which is used to build an association rule mining solution. Our solutions are designed for outsourced databases that allow multiple data owners to efficiently share their data securely without compromising on data privacy. Our solutions leak less information about the raw data than most existing solutions.

Keywords—Association rule mining, frequent itemset mining, privacy-preserving data mining

I. INTRODUCTION

Association rule mining and frequent itemset mining are two popular and widely studied data analysis techniques. These two techniques have been employed in applications such as market basket analysis [1], health care [2], web usage mining [3], bioinformatics [4] and prediction [5]. A transaction database is a set of transactions, and each transaction is a set of data items with a unique TID (Transaction ID).This topic focus on privacy-preserving mining on vertically partitioned databases. To ensure data privacy the system will design an efficient homomorphic encryption scheme. Then the system propose a cloud-aided frequent item set mining solution, which is used to build an association rule mining solution. System design solutions for outsourced databases that allow multiple data owners to efficiently share their data securely without compromising on data privacy.

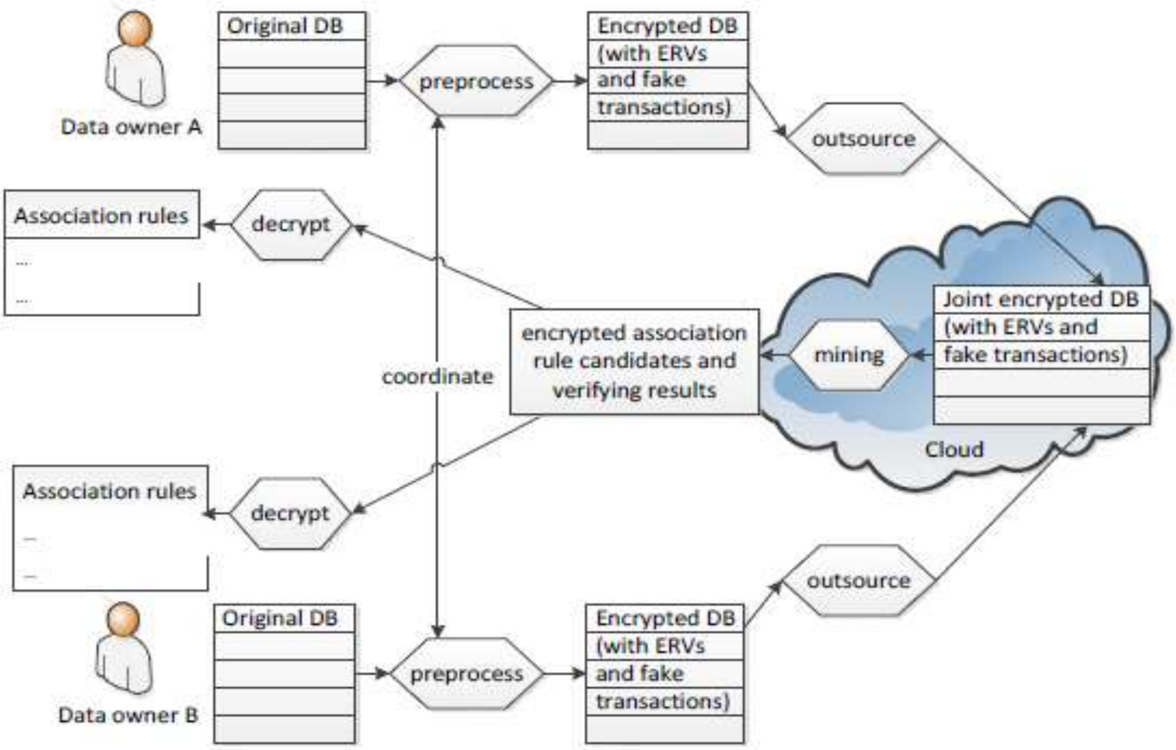
II. EXISTING SYSTEM

Classic frequent item set mining and association rule mining algorithms, such as Apriori, Eclat and FP-growth, were designed for a centralized database setting where the raw data is stored in the central site for mining. Privacy concerns were not considered in this setting. Due to an increased understanding of the importance of data privacy a number of privacy-preserving mining solutions have been proposed in recent times. In their settings, there are multiple data owners wishing to learn association rules 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. If each data owner has one or more columns in the joint database, the database is considered vertically partitioned.

III. PROPOSED SYSTEM

This paper focuses on vertically partitioned databases, such databases are useful for market basket analysis. 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. In this paper, we propose a cloud-aided privacy-preserving frequent itemset mining solution for vertically partitioned databases, which is then used to build a privacy-preserving association rule mining solution. Both solutions are designed for applications where data owners have a high level of privacy requirement. The solutions are also suitable for data owners looking to outsource data storage – i.e. data owners can outsource their encrypted data and mining task to a semitrusted (i.e. curious but honest) cloud in a privacy-preserving manner.

IV. SYSTEM ARCHITECTURE


Data owner A
Original DB
preprocess
Encrypted DB (with ERVs and fake transactions)
outsource
Association rules
...
...
decrypt
coordinate
encrypted association rule candidates and verifying results
mining
Joint encrypted DB (with ERVs and fake transactions)
Cloud
Association rules
...
...
decrypt
outsource
Data owner B
Original DB
preprocess
Encrypted DB (with ERVs and fake transactions)

V. MATHEMATICAL MODEL

Let S is system for implementation

S={U,I,O,D,P}

Where

U=set of users

Ui={u1,u2,u3…….un}

I=set of inputs

Ii={i1,i2,i3………in}

for ex login details, registration details etc

O=set of outputs

Oi ={o1,o2,03…….on}

For ex. login /denied, registration successful/ unsuccessful

D=set of devices

Di={d1,d2,d3……dn}

P=set of processing

Pi={p1,p2,p3,…….pn}
For ex authentication, update information

VI. FUTURE SCOPE

Apart from the work towards this system, future work mainly comprises of the following
Objective: In future we can perform functions like Privacy preserving mining on vertically partitioned database more effectively.

VII. CONCLUSION

In this paper, system proposed a privacy-preserving outsourced frequent itemset mining solution for vertically partitioned databases. This allows the data owners to outsource mining task on their joint data in a privacy-preserving manner. Based on this solution, we built a privacy-preserving outsourced association rule mining solution for vertically partitioned databases. This solutions protect data owner’s raw data from other data owners and the cloud. Our solutions also ensure the privacy of the mining results from the cloud. Compared with most existing solutions, our solutions leak less information about the data owners’ raw data. Evaluation has also demonstrated that this solutions are very efficient; therefore, this solutions are suitable to be used by data owners wishing to outsource their databases to the cloud but require a high level of privacy without compromising on performance.

REFERENCES

[1] T. Brijs, G. Swinnen, K. Vanhoof, and G. Wets, “Using association rules for product assortment decisions: A case study,” in SIGKDD 1999.

[2] 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,” Journal of the American medical informatics association, vol. 5, no. 4, pp. 373–381, 1998.

[3] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, “Effective personalization based on association rule discovery from web usage data,” in WIDM 2001.

[4] C. Creighton and S. Hanash, “Mining gene expression databases for association rules,” Bioinformatics, vol. 19, no. 1, pp. 79–86, 2003.

[5] X. Yin and J. Han, “Cpar: classification based on predictive association rules.” in SIAM SDM 2003.