South Asian Journal of Engineering and Technology Research Article

Open Access Full Text Article 123(2022) 158-162 DOI: 10.26524/sajet.2022.12.52

Improved privacy-preserving using elliptic curve diffie hellman and k-anonymity methodology

M. Karthick[1], P.Sathish Kumar[2]

[1]Assistant Professor, Nandha College of Technology, Erode.

[2]PG scholars, Nandha College of Technology, Erode.

Corresponding author.
Correspondence: M. Karthick
E-mail:magukarthik@nandhatech.org

Article info
Received 22 th August 2022
Received in revised form 12 September 2022 Accepted
3 October 2022

Keywords
cypher texts, homomorphic encryption, k-anonymity

Retrived from: https://sajet.in/index.php/journal/article/view/234

Abstract
Within a corporate privacy-preserving system, the issue of outsourcing the association rule mining task is taken into account. Therefore, privacy-preserving data mining is a research topic that is concerned with the security established from personally identifiable information when taken into account for data mining. We create a secure comparison technique and an effective homomorphic encryption scheme to guarantee data privacy. then suggests a frequent item set mining approach supported by the cloud that is used to create an association rule mining approach. The solutions are made for external databases that let several data owners communicate their information securely and productively without sacrificing data privacy. Compared to the majority of other solutions, the solutions leak less information about the raw data. The Rob Frugal encryption method is suggested as a solution to the security issue created by authorised users using external datasets.This system's proposed Rob algorithm incorporates fictitious patterns in cyphers for objects stored in the database. The false patterns contained in the outsourced data may increase the capacity overhead. We include weighted support in the original support of items to address this issue by lowering the number of false patterns and raising the security level for outsourced data with less complexity. To reduce storage requirements, the fictitious transaction table data is transformed into a matrix format. Outsourced data is more secure in the proposed work because process attacks based on items and item sets are not feasible.

ABSTRACT

Within a corporate privacy-preserving system, the issue of outsourcing the association rule mining task is taken into account. Therefore, privacy-preserving data mining is a research topic that is concerned with the security established from personally identifiable information when taken into account for data mining. We create a secure comparison technique and an effective homomorphic

encryption scheme to guarantee data privacy. then suggests a frequent item set mining approach supported by the cloud that is used to create an association rule mining approach. The solutions are made for external databases that let several data owners communicate their information securely and productively without sacrificing data privacy. Compared to the majority of other solutions, the solutions leak less information about the raw data. The Rob Frugal encryption method is suggested as a solution to the security issue created by authorised users using external datasets.

This system's proposed Rob algorithm incorporates fictitious patterns in cyphers for objects stored in the database. The false patterns contained in the outsourced data may increase the capacity overhead. We include weighted support in the original support of items to address this issue by lowering the number of false patterns and raising the security level for outsourced data with less complexity. To reduce storage requirements, the fictitious transaction table data is transformed into a matrix format. Outsourced data is more secure in the proposed work because process attacks based on items and item sets are not feasible.

INTRODUCTION

Association rules for identifying patterns between products in extensive transaction data captured by supermarket Point-Of-Sale (POS) systems. For instance, the rule revealed by a supermarket's sales data might suggest that if a consumer purchases potatoes and onions at the same time, they are likely to do the same with hamburger meat. Such data can serve as the foundation for judgments on marketing initiatives like, for example, promotional pricing or product placement. Association rules are used today in numerous application fields, including as Web usage mining, intrusion detection, and bioinformatics, in addition to the example from market basket analysis given above.

Homomorphic Encryption Scheme

The cypher texts can be subjected to one or more plaintext operations (such as addition and multiplication) using a homomorphic encryption technique. Additive homomorphic encryption is the term for the system if the addition operation is permitted. The method is referred to as multiplicative homomorphic encryption if the multiplication operation is permitted. The suggested approach suggests a safe outsourced comparison strategy in addition to an effective homomorphic encryption method. They provide an effective homomorphic encryption strategy to enable safe outsourced computation of supports and confidences, as well as a secure outsourced comparison scheme for comparing supports and confidences with thresholds, to prevent the exposure of supports and confidences.

Efficient Homomorphic Encryption Scheme

Most homomorphic encryption techniques are asymmetric. One encryption/decryption key using only modular additions and multiplications is proposed in this study, which is substantially more efficient than asymmetric (public and private keys) systems. The method is more effective than the homomorphic encryption techniques used in other association rule mining and frequent item set mining methods since it just calls for modular additions and multiplications. The suggested system is included in the scheme, which allows for numerous homomorphic additions and a constrained number of homomorphic multiplications.

The proposed techniques can possibly be used in a wide variety of secure compute applications, although they are created for the data mining solutions described in this research.

Preprocessing and mining are the two stages of the association rule mining system. Data owners and the cloud work together during the preprocessing stage to create an encrypted joint database at the cloud's end as well as certain auxiliary data for mining that protects privacy. The cloud mines association regulations for data owners in the mining stage while protecting privacy. Instead of mining association rule candidates during the mining stage, the cloud mines frequent item set candidates. In order to recover the actual frequent item sets, the data owners must first decrypt the encrypted verifying results and frequent item set candidates.

Privacy-Preserving Outsourced Mining

Before outsourcing, the database's data objects are often encrypted with a substitution cypher. Thus, the substitution cypher is not vulnerable to counter frequency analysis attacks. The data owner hides the item frequency in the encrypted database by inserting bogus transactions as a defence against frequency analysis attacks. Any item in the encrypted database will have a frequency with at least k 1 other items after the bogus transactions have been inserted. The owner of the data uploads to the cloud the encrypted database of both genuine and made-up transactions. The frequent item sets and their supports are returned to the data owner after being processed by a traditional frequent item set mining algorithm on the cloud.

The owner of the data updates these item sets by deducting them from the associated occurrence counts in the fictional transactions. The data owner then creates association rules based on the frequent item sets discovered after decrypting the received item sets with the revised supports higher than the frequency threshold. The solutions employ their methods to thwart frequency analysis attacks that the cloud might launch and hide the raw data from it. The current implementation is resistant against chosen-plaintext attacks on encrypted things, but it is prone to frequent analysis attacks The same supports provided to data owners will likewise leak if this technique is applied to databases that are vertically partitioned.

Homomorphic encryption is a form of cryptography that dispenses with the necessity to decrypt ciphertexts in order to evaluate arbitrary functions on encrypted data. This project uses the most advanced homomorphic encryption techniques. The approach is based on mining in vertically partitioned databases while protecting privacy. In this case, data owners are interested in learning the association rules or frequent item sets from a large data collection while disclosing as little as possible about their (sensitive) raw data to other data owners and outside parties. The privacy-preserving frequent item set mining solution for vertically partitioned databases is part of the proposed system, and it is used to create a privacy-preserving association rule mining solution.

In the cloud, you can obtain a better level of privacy because the majority of solutions currently in use call for sharing/exposing raw data or disclosing the precise supports to data owners. Due to these restrictions, sensitive data from the raw data is exposed. Privacy - Data owners should have as limited access to other data owners' databases as possible. To prevent information about the raw data from leaking, a data owner's raw transaction details should not be exposed, and the supports should be hidden. The same goes for exact confidence levels, which can be utilised to deduce details about the raw data. The suggested solutions ought to guard the cloud-based mining results as well. Performance is typically lowered as a result of privacy-preserving measures.

The k-anonymity

Both of these techniques modify individual pieces of data to prevent their precise identification. The k-anonymity method effectively reduces the granularity of data representation so

that a particular set of data maps onto at least k other records in the data. Techniques like generalization and suppression are required. The k-anonymity method is flawed in that sensitive values can be deduced for the altered data if there is uniformity of sensitive values within a group. In order to achieve anonymization, the l-diversity model was created to handle this problem by requiring intra group variety of sensitive values. The goal is to make it as difficult as possible for adversaries to precisely identify individual records using combinations of data properties. The dataset needs to be made anonymous as the initial step.

The ARX Anonymization Tool is used to perform the anonymization procedure. Using this technology, the attributes of age, the number of pregnant women, and class were chosen for anonymization. In order to achieve data anonymity, the anonymization procedure uses the generalisation technique, which involves generalising an attribute from the chosen dataset. The hiding failure after the data has been anonymized is the outcome that is gathered following the anonymization procedure.

CONCLUSION

A privacy-preserving outsourced frequent item set mining solution for vertically partitioned databases is included in the proposed system. Consequently, the data owners can outsource mining tasks on their shared data while protecting privacy. A privacy-preserving outsourced association rule partitioned database is created based on this solution. The solutions shield a data owner's raw data from cloud storage and other data owners. The technologies also guarantee the confidentiality of cloud mining findings.

The solutions leak less information about the data owner's raw data than the majority of existing solutions. Solutions are suited for use by data owners who want to outsource their databases to the cloud but demand a high level of privacy without compromising on performance because evaluation has also shown that they are very efficient. An effective homomorphic encryption method and a safe outsourced comparison scheme were given in this research to actualize the solutions. Beyond the data mining techniques discussed in this study, both schemes have potential applicability in other secure computation contexts, such as secure data aggregation. Future study will concentrate on demonstrating the applicability of the suggested homomorphic encryption scheme and outsourced comparison scheme in various contexts.

REFERENCES

1. Agrawal R and Srikant R, (1994), “Fast algorithms for mining associationrules,” in Proc. VLDB, pp. 1–13.
2. Amer. Med Inform J, Brossette S E, Jones W T, Hardin J M, Moser S A,Sprague A P, and Waites K B, (1998),“Association rules and data mining inhospital infection control and public health surveillance,” Assoc., vol. 5, no.4, pp. 373–381.
3. Brijs T, Swinnen G, Vanhoof K and Wets G, (1999), “Using association rules for product assortment decisions: A case study,” in Proc. SIGKDD,pp. 254–260.
4. Chang L, Matwin S and Zhan J, (2005), “Privacy-preserving collaborative association rule mining,” in Proc. DBSEC , pp. 153–165.
5. Clifton C and Kantarcioglu M, (2004), “Privacy-preserving distributed mining of association rules on horizontally partitioned data,” IEEE Trans. Knowl. Data Eng., vol. 16, no. 9, pp. 1026–1037.
6. Clifton C and Vaidya J, (2002), “Privacy preserving association rule mining in vertically partitioned data,” in Proc. SIGKDD, pp. 639–644.

7. Cramer R., Gennaro R and Schoenmakers B, (1997), "A secure and optimally efficient multi-authority election scheme," Eur. Trans. Telecommun., vol. 8, no. 5, pp. 481–490.
8. Creighton C and Hanash S, (2003), "Mining gene expression databases for association rules," Bioinformatics, vol. 19, no. 1, pp. 79–86.
9. Dai H,B. Luo T, Mobasher and Nakagawa M,(2001), "Effective personalization based on association rule discovery from Web usage data," in Proc. WIDM,, pp. 9–15.
10. Gudes E and Rozenberg B,(2006), "Association rules mining in vertically partitioned databases," Data Knowl. Eng., vol. 59, no. 2, pp. 378–396.
11. Han J and Yin X, (2003), "CPAR: Classification based on predictive association rules," in Proc. SIAM SDM, pp. 1–5.
12. Han J, Pei J and Yin Y, (2000), "Mining frequent patterns without candidate generation," in Proc. ACM SIGMOD, pp. 1–12.
13. V.S. Sureshkumar, A.Chandrasekar, "Fuzzy-GA Optimized Multi-Cloud Multi-Task Scheduler for Cloud Storage and Service Applications", International Journal of Scientific & Engineering Research, Volume 4, Issue3, March-2013.
14. V.S. Sureshkumar "Optimized Multicloud Multitask Scheduler for Cloud Storage and Service by Genetic Algorithm and Rank Selection Method", International Journal of Advanced Science Engineering and Technology, pp:2-7, Issue 4, volume 3,2014.
15. E.prabhakar, V.S.Suresh kumar, Dr.S.Nandagopal "Mining better advertisement tool for government schemes using machine learning",International Journal of psychosocial Rehabilitation,pp:1122-1135,Issue 4, Volume23,2019.
16. V.S.Suresh kumar , E.Prabhakar, Dr.S.Nandagopal Likihood weighted bagging ensemble approach to analyze public sentiment about covid -19 pandemic,International journal for mechanical engineering ,pp:11.01-1107,Issue-3,Volume-6,2021.