ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

Association Rule Mining Method for Applying Encryption Techniques in Transaction Data

Karishma Chopda[1], Apurva Rote[2], Komal Gaikwad[3], Priyanka Gachale[4]
BE Student[1, 2, 3, 4]
Department of computer Engineering
LGNSCOE, Nashik, India

Abstract:
The techniques of knowledge discovery process is data mining and denotes to mining information from large amount of records. Association rule mining is one of the significant and popular data mining process which is used to find frequent patterns in the given dataset in which the Apriori algorithm are the most common for mining frequent item set. This paper explored DES algorithm to the client side for generating of the secret key to encrypt the items of the support table. Apriori algorithm is used at server side for getting transaction data from the client side by applying threshold or sigma value to filter out the item set whose frequency has to less than sigma value. Main focus of this paper is to achieve more security to client side.

Keywords: Privacy preserving data mining, Transaction database, Association rules mining, Data encryption standard

I. INTRODUCTION

Data mining is an interdisciplinary subfield in computer applications which is having computational process of discovering patterns of huge data sets involving methods at interaction of some machine learning and database system process applications. Data mining is analogical procedure to extract meaningful information from huge database. Data mining technology is an emerging process of identifying patterns from large quantities of data with relevant features in semantic data. Data mining is used to translate random data into meaningful information and show the huge data. There are two types of mining technique, Descriptive mining and Predictive mining. The descriptive mining distinguish the universal properties of information present in file. The predictive mining execute conclusion on existing information that arrange formulate calculations performing applying association rules mining. Association rule mining is solitary major method that treat in knowledge discovery in database. Association rule mining is generating the frequent item sets. Item set is frequent if the items in the group of sets transfer the data frequently. Association rule mining is one of the most important and well researched techniques of data mining. The aims to extracting the interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data set. Association rules are widely used in various fields such as telecommunication networks, market and risk management, inventory control etc. The frequently generated data required some amount of privacy. The goal is to provide the privacy to the convinced portions of the data, while preserving service. The transforming data in secure way is applying various encryption techniques proposed. The privacy calculation for using encryption methods and to achieve the privacy for variation of data. This method is reducing the granularity of arranging to magnify solitude. Encryption methods: The encryption method which enable optional privacy to be verified, and to confirm this copy in excess of inclusive actual existence operation databases. The client/owner encrypts its data by means of encrypt and decrypt Encryption or Decryption module. This module applying Data Encryption Standard (DES) algorithm for achieving security of the leaked data and another technique fake transaction for privacy preserving data miningin which sound is added to the information in arrange to k-1.

II. LITERATURE SURVEY

Data mining, otherwise known as knowledge discovery, attempts to answer this need. The field of privacy has seen rapid advances in recent years because of the increases in the ability to store data. In particular, recent advances in the data mining field have led to increased fears about privacy. While the topic of privacy has been traditionally studied in the context of cryptography and information hiding, recent emphasis on data mining has lead to renewed interest in the field. If there are multiple transaction all transactions is send to provider for mining association rule that are local to individual store or global rule for organization therefore effective data mining for distributed owner is done and the disadvantages is the data perturbation is less attractive and gives only approximate result. The one to one substitution method was applying so easy to guess to recover cipher text that's why to design DES algorithm and achieve more security as compare to 1-1 substitution method.

III. PROPOSED WORK

The proposed approach distributes each transaction D into partitions and in each local partition frequent item sets are find out at every site. After finding local frequent item sets all local frequents item sets are combined to find candidate item set. In last stage global frequent item sets are found. Data mining on a very large data set is a complex task. The rule discovery on this

large volume data becomes slow, since it is done serially on available big data sets. The large quantity of data records may overload a computer's memory and processor due to this the learning process becomes very slow. Mining on large data set may become impossible because of limitations on processor and memory. Distributed mining algorithm which is a combination of preserving privacy and fast distributed mining algorithm which is an unsecured distributed version of the Apriori algorithm. Its main idea is that any S-frequent item set must be also locally S-frequent in at least one of the sites. Hence, in order to find all globally S-frequent item sets, each player reveals his locally S-frequent item sets and then the players check each of them to see if they are S-frequent also globally.

IV. METHODOLOGY

A. System Block Diagram

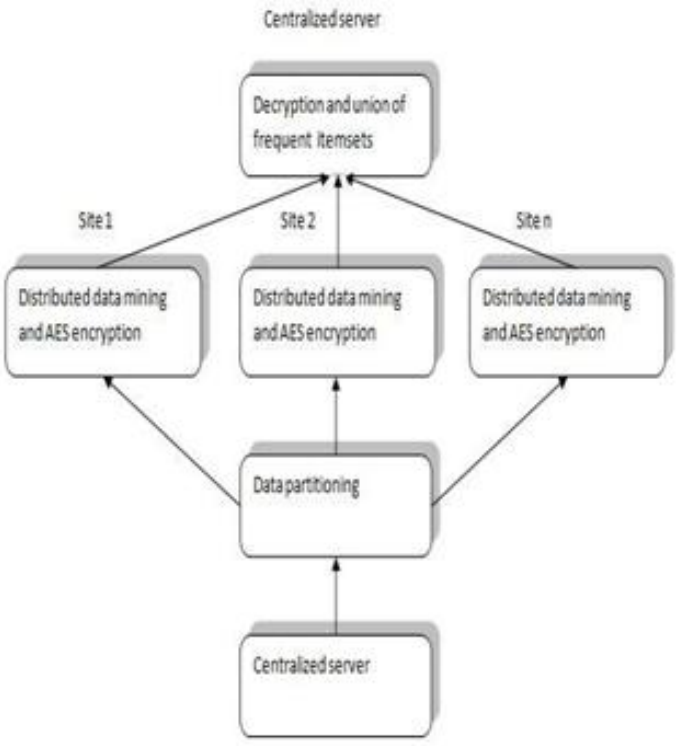
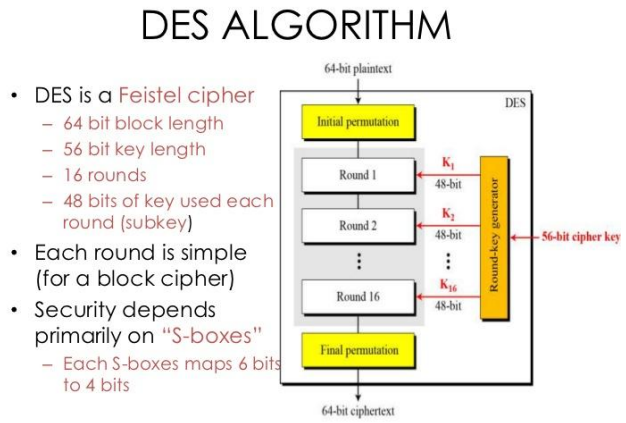


Figure shows how the system works in order to generate globally frequent item sets through mining process. It describes how the system flow is. Initially the centralized server takes transaction data set as input to system. The centralized server performs horizontal data partitioning through the application of distributed mining algorithm, this data is distributed among site1, site2 ...., Site N. While distributing data among sites the data is in the encrypted form so that the privacy is maintained during transfer of data, when this data reaches to particular client, the client again decrypts the data and extracts the original data. The encryption and decryption is performed using AES encryption/Decryption algorithm. The client receives this data in encrypted form, by applying decryption key, client extracts original data, after extraction of original data, client identifies locally frequent item sets by applying distributed mining algorithm, after identifying locally frequent item set it again encrypt these locally frequent item sets and send back to server. At server side these locally frequent item sets are encrypted and for merging of item sets the K&C algorithm is used at server side. This process is repeated until server finds globally frequent item set. Thus the output of system is globally frequent item sets which are obtained at centralized server. In the proposed system the problem of secure computation of union of private subsets available on sites is addressed and the database is horizontally Partitioning and distributed among various sites or player which are involved in transaction. Round robin technique is used for Horizontal distribution of Data sets to reduce the data skew. The problem of securely mining association rule in distributed environment is implemented here. In this system there are several sites that hold identical databases these databases are distributed horizontally over different sites participating in transaction. The goal is to mine these data sets for finding all association rules with support count at least s and confidence count at least c. The given minimal support count S and confidence size C, also hold for the unified database. The main target is to design an algorithm to enable handling of large data sets using available computing resources and to design the system to accelerate a mining process on large data sets. Achieving a speed up in computation process by utilizing resources available in distributed Environment provide more security in distributed computing environment. System is concerned with maintaining security while mining of association rules in distributed database, where input is synthetic database and output will be set of association rules. To design secure multi-party algorithms that computes the union of private subsets that each of the interacting sites hold, and to design algorithm to test the inclusion of set of an element held by one site in a subset held by another. The important objective of the proposed algorithm is to minimize the information disclosed about the private database held by the sites.

Algorithm



There are two main types of cryptography in use today symmetric or secret key cryptography and asymmetric or public key cryptography. Symmetric key cryptography is the oldest type whereas asymmetric cryptography is only being used publicly since the late 1970's 1. Asymmetric cryptography was a major milestone in the search for a perfect encryption scheme. It involves the use of only one key which is used for both encryption and decryption (hence the use of the term

symmetric). Figure depicts this idea. It is necessary for security purposes that the secret key never be revealed.

V. EXPERIMENTAL APPROACH AND RESULT

Figure.1. Login.

Figure.2. Dash-Board.

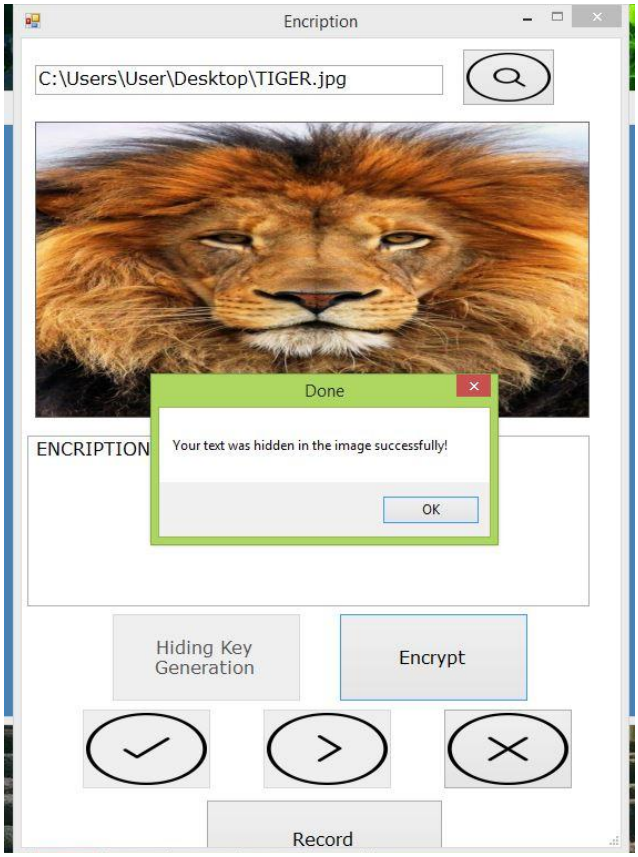


Figure. 3.Encryption.

Figure. 4. Decryption.

Time Complexity:
Encryption method O (n) and decryption method O (m). If n = m in that case we are using complexity O (2m).

Encryption:

```
for (int i = 1; i <=m; i += c) {
      // some O(1) expressions
}
```

Decryption:

```
 for (int i = 1; i <=n; i += c) {
      // some O(1) expressions
   }
```

Time complexity of above code is O(m) + O(n) which is O(m+n)

If m == n, the time complexity becomes O(2n) which is O(n).

VI. ACKNOWLEDGEMENT

It gives us great pleasure in presenting the preliminary project report on 'Association Rule Mining Methods for Applying Encryption Techniques in Transaction Dataset'. I would like to take this opportunity to thank my internal guide Prof. N. R. Wankhade for giving me all the help and guidance I needed. I am really grateful to them for their kind support. Their valuable suggestions were very helpful. I am also grateful to Prof. N. R. Wankhade, Head of Computer Engineering Department, College Name for his indispensable support, suggestions. In the end our special thanks to Other Person Name for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for this system.

VII. CONCLUSION

The proposed system is used to ensure security risk which helps to transfer data from sender to receiver. To send data user must be authorized user to send data and it has its own secret key that is used to transfer data to receiver. The receiver will get a notification viz email to decrypt data. Applying DES algorithm so achieve more security and prevent intruder attack. It also uses the Rob Frugal method to add the number of fake pattern. Fake transaction can be adding item based manner.

VIII. REFERENCES

[1]. FoscaGiannotti, Laks V. S. Lakshmanan, Anna Monreale, Dino Pedreschi, and Hui (Wendy) Wang," Privacy-Preserving Mining of Association Rules From Outsourced Transaction Databases", IEEE SYSTEMS JOURNAL, VOL. 7, NO. 3, SEPTEMBER 2013.

[2]. RakeshAgrawal,, IEEE, Tomasz Imielinski, and Arun Swami," Database Mining: A Performance Perspective", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL 5, NO. 6. DECEMBER 1993.

[3].RajkumarBuyya, Chee Shin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", The 10th IEEE International Conference on High Performance Computing and Communications.

[4].Alex Biryukov, Christophe De Cannière FDEF, Campus Limpertsberg, University of Luxembourg, Luxembourg Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven-Heverlee, Belgium, "Data Encryption Standard (DES)".

[5].Ling Qiu,Yingjiu Li and Xintao Wu," Preserving privacy in association rule mining with bloom filters". J IntellInfSyst (2007) 29:253–278.

[6].Justin Zhan,StanMatwin and LiWu Chang," Privacy-preserving collaborative association rule mining", ELSEVIER Journal of Network and Computer Applications 30 (2007) 1216–1227.

[7].Kshitij Pathak, Narendra S Chaudhari and Aruna Tiwari," Privacy Preserving Association Rule Mining by Introducing Concept of Impact Factor", IEEE 978-1-4577-2119-9/12/2011.

[8].HaishengLi, "Study of Privacy Preserving Data Mining", Third International Symposium on Intelligent Information Technology and Secuity.