International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Implementation of Efficient Privacy Preserving Data Analysis for Secure Data Transmission

Shital Gawahale[1], Rekha Jadhav[2]

Computer Department, G.H. Raisoni Institute Engineering & Technology, Pune, India

Professor, Computer Department, G.H.Raisoni Institute Engineering & Technology, Pune, India

Abstract: ***Privacy preserving is one of the most important research topics in the data security field and it has become a serious concern in the secure transformation of personal data in recent years. For example, different credit card companies and disease control centers may try to build better data sharing or publishing models for privacy protection through privacy preserving data mining techniques (PPDM). A model has been proposed to design the effective Privacy Preserving Mining Framework for secure private information transformation and Publishing. The system considers a distributed database such as bank database that is used to construct the incentive model. Basically incentive data are used to check the user knowledge that is the online processing user is correct person or not. Incentive Compatible Model based secure code computation process and PPDM techniques like Association rule mining, Randomization method and Cryptographic technique. An Encryption algorithm is used to identify which data sets need to be encrypted for preserving privacy in data storage publishing. Secure multi-party computation (SMC) has recently emerged as an answer to this problem but SMC model they had generate only one key for the each card and after some extent time there may be chance of hacking the password by the hackers. To overcome the above said problem in our proposed system we are using symmetric key cipher algorithm to increase the security measures. In this system it will encrypt the security code so if any hacker check for the key it will be in encrypted format so no one can hack the password. And to send the keys to the mail we are using Java Mail API directly to communicate with the Gmail server.***

Keywords: PPDA, PPDM, SMC, TTP etc

1. Introduction

Today's scenario we have E-Commerce, E- Governance and personal data is distributed online, privacy of data is become the most important issue. The information found in mining can be sensitive or it can be misuse by anyone. Privacy and security, particularly maintaining confidentiality of data, have become a challenging issue with advances in information and communication technology. the ability to communicate and share data has many benefits. and the idea of an omniscient data source carries great value to research and building accurate data analysis models. For example, for credit card companies to build more comprehensive and accurate fraud detection system, credit card transaction data from various companies may be needed to generate better data analysis models.

2. Related Work

Author Name	*Survey*
Murat Kantarcioglu and Wei Jiang,	Privacy-preserving data analysis protocols are designed using cryptographic techniques for security and privacy. Data are generally divided into two types of data collection form to be either vertically or horizontally partitioned. Different sites collect the same set of information about different attributes in the case of horizontally partitioned data.
Dr. K. P. Thooyamani, Dr.V.khanaa,	proposed procedures to mine distributed association rules on horizontally partitioned data and showed that distributed association rule mining can be done efficiently under reasonable security assumptions
Vaidya et	The concept of the heuristic approach method is the way to hide sensitive rules which is used to be mined from the dataset while maximizing the outcome of the released data. The second approach is Cryptography based method, This approach has been developed to solve the problem such as SMC.

Objective

1) The main objective of the paper is to maintain the confidentiality of the data in secure transactions.
2) We are providing a secure SMC detection to more precisely detect the possible attackers and question generation module to challenge the suspected requesters who are detected by the detection module.
3) When client attacks on server system our system detects that attack

Design Goal

The incentive compatible privacy preserving model has to interact with the participating parties to verify the transaction making use of the user's knowledge. The E-Shopping is a service oriented application, which provides a user interaction interface that provides more security for individual details transformation compared with the other privacy preserving models.

Scope

Privacy preserving data mining (PPDM) is how to protect the sensitive information or private knowledge from leaking in the mining process, meanwhile obtain the accurate results of data mining.
The privacy preserving data mining is divided into two levels

- First level of PPDM is focus on protecting the sensitive data such as id, name, address and other sensitive information.
- Second level of PPDM is focus on protecting the sensitive knowledge which is showed by data mining.

Problem Statement

Secure multi-party computation (SMC) has recently emerged as an answer to this problem but SMC model they had generate only one key for the each card and after some extent time there may be chance of hacking the password by the hackers.

3. Existing System

In an existing system they first develop key theorems and then base on these theorem, they analyzed what types of privacy-preserving data analysis tasks could be conducted in a way that telling the truth is the best choice for any participating party. Secure multi-party computation (SMC) has recently emerged as an answer to this problem but SMC model they had generate only one key for the each card and after some extent time there may be chance of hacking the password by the hackers[1].

Disadvantages

- Participating parties can't provide truthful input data.
- Security system not depend upon the truth full data
- Fraud accept the credit card easily

To overcome the above said problem in our proposed system we are using symmetric key cipher algorithm to increase the security measures. In this system it will encrypt the security code so if any hacker check for the key it will be in encrypted format so no one can hack the password. And to send the keys to the mail we are using Java Mail API directly to communicate with the Gmail server

Implementation Idea

Privacy and security, particularly maintaining confidentiality of data, have become a challenging issue with advances in information and communication technology.. For example, for credit card companies to build more comprehensive and accurate fraud detection system, credit card transaction data from various companies may be needed to generate better data analysis models.

Methods

Various types of web service models are used in private data transformation applications. These types of security models are based on the various Privacy Preserving Data Mining (PPDM) techniques such as Randomization method, Secure Code Computation process and Encryption method [6] [9].

Motivation

Now days secondary use of data become very common. Secondary use of data means data is used for some other purpose not for which data is collected initially. The potential misuse of personal information of public is increasing rapidly. The scope of sensitive data is not limited to medical or financial records it may be phone calls made by an individual, buying patterns and many more. No one wants that his/her personal data is sold to any other party without their prior consent. Some individuals become hesitant to share their information which results additional difficulty to obtaining correct information. Public awareness is so much important if private information is shared between different entities. Public awareness about privacy and lack of public trust in organization may introduce additional complexity to data collection

Background

SMC (Secure Multi Computing)

SMC technology used in distributed privacy preserving data mining areas that mainly consists of a set of secure sub-protocols, such as, secure sum, secure intersection, secure set union, secure comparison, dot product protocol and so on. In the following will be briefly describe the basic idea of two kinds of secure sub-protocols used in horizontally partitioned and vertically partitioned setting.

Proposed System

In the SMC model they will generate only one key for the each card and after some extent time there may be chance of hacking the password by the hackers so in the proposed system we are using symmetric key cipher algorithm to increase the security measures. In this system it will encrypt the security code so if any hacker checked for the key it will be encrypted format so no one can hack the password. And to send the keys to the mail we are using Java Mail API directly to communicate with the Gmail server[2].

Advantage

- Users give their truth full data for security system.
- User Only Knows the answers for security questions.
- Users Knows the Fraud entry.
- Fraud could be detected.

Design and Implementation Constraints

1. Our SMC is used to Mail server to send code and card no
2. For code encryption we are using Symmetric Algorithm.
3. Change password link is used to update old password
4. Apply Credit Card link is used to generate a Card No.
5. Preprocess all data is used for preprocessing purpose.
6. Use credit card link for Purchasing product online
7. The accuracy of the predicted model should be good as compared to other models.

Proposed Architecture

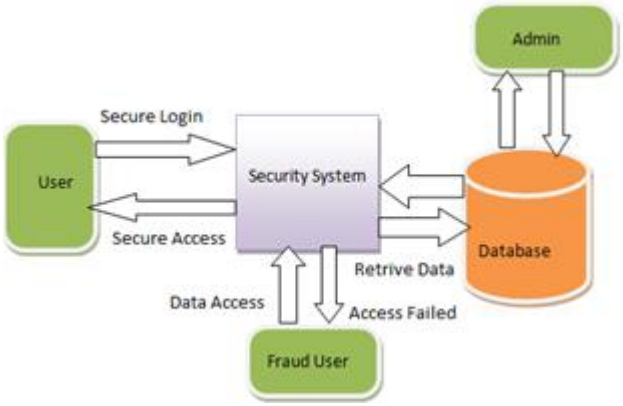

Figure 1: Proposed Architecture

4. Proposed Methodology

4.1 Java Mail API:

The Java Mail API provides a platform-independent and protocol-independent framework to build mail and messaging applications. The Java Mail API provides a set of abstract classes defining objects that comprise a mail system. It is an optional package (standard extension) for reading, composing, and sending electronic messages.

Java Mail provides elements that are used to construct an interface to a messaging system, including system components and interfaces. While this specification does not define any specific implementation, Java Mail does include several classes that implement RFC822 and MIME Internet messaging standards. These classes are delivered as part of the Java Mail class package.

Following are some of the protocols supported in Java Mail API:

- SMTP: Acronym for Simple Mail Transfer Protocol. It provides a mechanism to deliver email.
- POP: Acronym for Post Office Protocol. POP is the mechanism most people on the Internet use to get their mail. It defines support for a single mailbox for each user. RFC 1939 defines this protocol.
- IMAP: Acronym for Internet Message Access Protocol. It is an advanced protocol for receiving messages. It provides support for multiple mailbox for each user, in addition to, mailbox can be shared by multiple users. It is defined in RFC 2060.

SMPT server

To send emails, you must have SMTP server that is responsible to send mails. You can use one of the following techniques to get the SMTP server.

- Install and use any SMTP server such as Postfix server (for Ubuntu), Apache James server (Java Apache Mail Enterprise Server)etc.
- Use the SMTP server provided by the host provider for eg: free SMTP provide by Jango SMTP site is relay.jangosmtp.net
- Use the SMTP Server provided by companies e.g. gmail, yahoo, etc[2].

5. Analysis and Discussion

Privacy Preserving Association Rule Mining
The association rule mining and analyze whether the association rule mining can be done in an incentive compatible manner over horizontally and vertically partitioned databases. The Security Code is valid means retrieve the data otherwise you are a fraud user[3].

Modules

- Privacy-Preserving Data Analysis
- Non-Cooperative Computation
- Analyzing data analysis tasks ii the NCC model:
- Privacy Preserving Association Rule Mining

Privacy-Preserving Data Analysis
The privacy preserving data analysis protocols assume that participating parties are truthful about their private input data .
The techniques developed in assume that each party has an internal device that can verify whether they are telling the truth or not.
.

Privacy Preserving Data Publishing
The incentive compatible model is only concentrating on the secure data sharing process and does not consider the data storage publishing. All the users' private information's are stored in the particular database that is more securable one. Using a symmetric encryption algorithm, Triple Des algorithm is used to encrypt the all the users sensitive information in the secure database.

Privacy Preserving Association Rule Mining
The association rule mining and analyze whether the association rule mining can be done in an incentive compatible manner over horizontally and vertically partitioned databases. The Security Code is valid means perform the credit card operation otherwise not respond the system.

Randomized Incentive Compatible Model

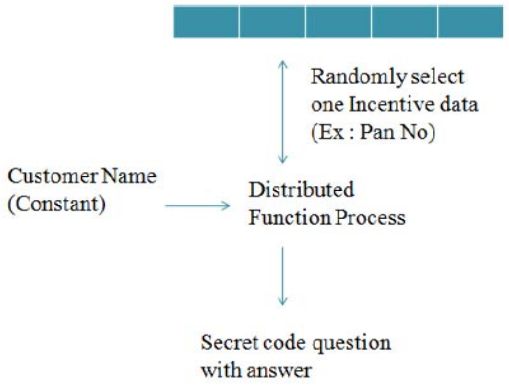

Figure 2: Randomized Incentive Compatible Model – Secure Code Computation Process

Secure Code Computation Process Incentive Compatible Secret Code

Question for the Non – Cooperative Privacy Model. The figure 2 represents the process of secure code computation for privacy preserving randomized incentive model.

The computation process theorem consists of following steps:
Step 1: Select two attributes from customer details from bank database as inputs for the distributed function.
Step 2: Here first attribute is constant and another one attribute is other personal details of customer information that is selected by random.
Step 3: Applied Privacy Preserving techniques such as partitioning data, secure sum and dot protect operation on the selected two attributes [3].

6. Algorithm

Symmetric Key Cryptosystem

Symmetric Key cryptosystem Symmetric encryption, also referred to as conventional encryption or single key encryption was the only type of encryption in use prior to the development of public-key encryption .

1) Plaintext: This is the original intelligible message or data that is fed to the algorithm as input.
2) Encryption algorithm: The encryption algorithm performs various substitutions and permutations on the plain text .
3) Secret Key: The secret key is also input to the encryption algorithm. The exact substitutions and permutations performed depend on the key used, and the algorithm will produce a different output depending on the specific key being used at the time.
4) Cipher text: This is the scrambled message produced as output. It depends on the plaintext and the key. The cipher text is an apparently random stream of data, as it stands, is unintelligible.
5) Decryption Algorithm: This is essentially the encryption algorithm run in reverse. It takes the cipher text and the secret key and produces the original plaintext [2].

Encryption Algorithm

Step 1: Generate the ASCII value of the letter
Step 2: Generate the corresponding binary value of it.
[Binary value should be 8 digits e.g. for decimal 32 binary number should be 00100000]
Step 3: Reverse the 8 digit's binary number
Step 4: Take a 4 digits divisor (>=1000) as the Key
Step 5: Divide the reversed number with the divisor
Step 6: Store the remainder in first 3 digits & quotient in next 5 digits (remainder and quotient wouldn't be more than 3 digits and 5 digits long respectively. If any of these are less then 3 and 5
digits respectively we need to add required number of 0s (zeros) in the left hand side. So, this would be the cipher text i.e. encrypted text.
Now store the remainder in first 3 digits & quotient in next 5digits.

Decryption Algorithm

Step 1: Multiply last 5 digits of the cipher text by the Key
Step 2: Add first 3 digits of the cipher text with the result produced in the previous step
Step 3: If the result produced in the previous step i.e. step 2 is not an 8-bit number we need to make it an 8- bit number
Step 4: Reverse the number to get the original text i.e. the plain text.

OUTPUT

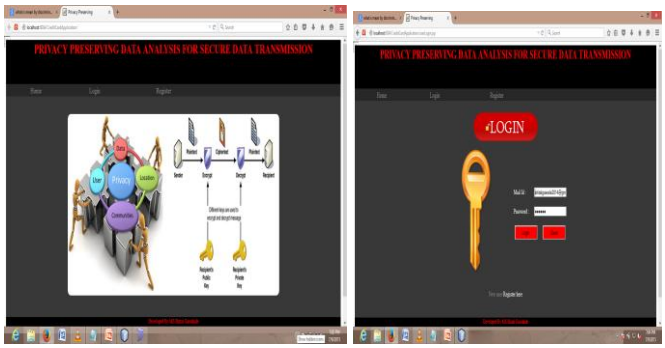

Figure 3: Login Screen

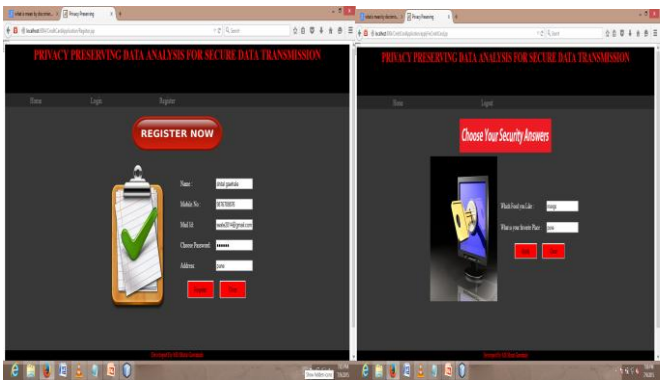

Figure 4: Registration Screen

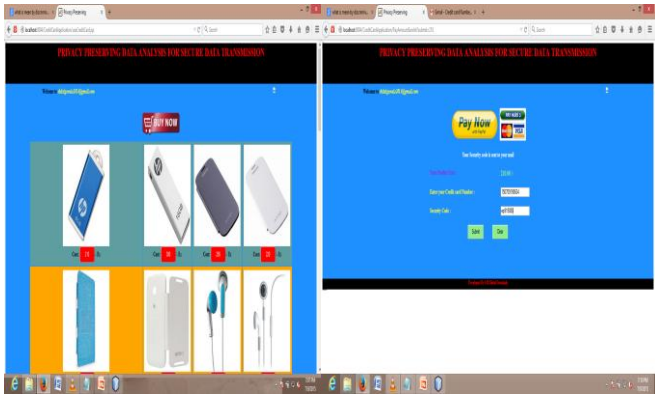

Figure 5: Payment Screen & Question Generation Screen

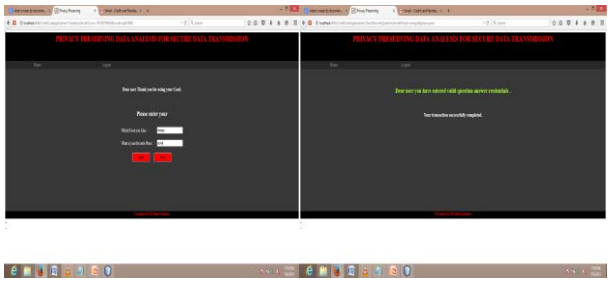

Figure 6: Detect Frud invalid or Valid user

7. Result

The Non - Cooperative Incentive Compatible Model is helpful for secure data sharing process in more online applications. This model is mainly used to maintain the confidentiality of the user private data and does not allow the malicious user to access the securable database. The Incentive Compatible Probabilistic NCC Model provides 95% success rate in user's authentication process compared to existing Deterministic NCC model that provide only 85% success rate..

8. Conclusion

Here we finally Conclude that The incentive compatible privacy-preserving data analysis technique has been developed to motivate the participating parties to provide truthful inputs. The privacy preserving data analysis task that provides a new model called Incentive Compatible The main advantage of this model is that to reduce the number of False Positive transactions. It tries to find any anomalies transaction based on the data analysis model. In proposed system we are using symmetric key cipher algorithm to increase the security measures. In this system it will encrypt the security code so if any hacker check for the key it will be in encrypted format so no one can hack the password. And to send the keys to the mail we are using Java Mail API directly to communicate with the Gmail server**.**

9. Future Enhancement

The incentive compatible model is helpful for more privacy preserving application approaches that are used to interact with user original knowledge. In future to provide more than two attribute based more securable privacy preserving model can be built. And also using various privacy preserving techniques the efficiency of the privacy preserving model can be improved.

References

[1] Murat Kantarcioglu and Wei Jiang, "Incentive Compatible Privacy-Preserving Data Analysis", IEEE transactions on knowledge and data engineering, vol. 25, no. 6, june 2013.
[2] Ms. Shital Gawhale,Prof.Rekha Jadhav," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)" Volume 4, Issue 1, January-February2015 ISSN 2278-6856.
[3] M. Dhanalak," Effective Incentive Compatible Model for Privacy Preservation of Information in Secure Data Sharing and Publishing" International Journal of Computer Applications (0975 – 8887) Volume 96–No.16, June 2014
[4] M. Kantarcioglu and C. Clifton, "Privacy-Preserving Distributed Mining of Association Rules on Horizontally Partitioned Data," IEEE Trans. Knowledge and Data Eng., vol. 16, no. 9, pp. 1026-1037, Sept. 2004.
[5] W. Du and Z. Zhan, "Building Decision Tree Classifier on Private Data," Proc. IEEE Int'l Conf. Data Mining Workshop Privacy, Security, and Data Mining, C.Clifton and V. Estivill-Castro, eds.,vol. 14, pp. 1-8,Dec. 2002.
[6] J. Vaidya and C. Clifton, "Privacy Preserving Association Rule Mining in Vertically Partitioned Data," Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD '02), pp. 639-644, July 2002.
[7] R. Agrawal and E. Terzi, "On Honesty in Sovereign Information Sharing," Proc. Int'l Conf. Advances in Database Technology, pp. 240-256, 2006.
[8] M. Kantarcioglu and R. Nix, "Incentive Compatible Distributed Data Mining," Proc. IEEE Int'l Conf. Soc. Computing/IEEE Int'l Conf. Privacy, Security,