A Survey On securing m-commerce through robust cryptographic and data mining concept

Kavita Rathod[1],IndrJeet Rajput[2]

[1]Research Scholar, Computer Engineering, Hasmukh Goswami Collage of Engineering, Ahmedabad, Gujarat, India

[2]Assistant Professor,Computer Engineering, Hasmukh Goswami Collage of Engineering , Ahmedabad, Gujarat, India

Abstract

This paper focuses mainly on the Cryptography is an art and science of achieving security by encoding message to make them non-readable. It converts the data from readable format that is known as plain text into unreadable format known as cipher text and vice versa. There are various types of cryptographic algorithms proposed over the years based on different techniques. These techniques use various approaches to implement the basic functionality of cryptography i.e. to hide the information from unauthorized user. This survey describes various aspects of cryptographic techniques and various issues related to cryptography. Along with it, a proposed work is there which addresses some of the core issues of cryptography along with their solutions. Data mining is a technique to dig the data from the large databases for analysis and executive decision making. Security aspect is one of the measure requirement for data mining applications. In this paper we present security requirement measures for the data mining.

Keywords- *Encryption, Decryption, Multilevel security , RSA,LZW,ECC.*

INTRODUCTION

1.0 Introduction to Security :

Information Security is the protection of information from unauthorized access, use, disclosure, disruption, changing or demolition in order to provide confidentiality, integrity and availability. It is the process of ensuring the protected innovation of the organization.

Information security is a multidisciplinary area of study and expert activity which is concerned with the improvement and usage of security mechanisms of all available types in order to keep information in all locations and consequently, information systems, where information is created, processed, then stored, transmitted and ultimately destroyed, free from threats.

1.1 Data mining:

Data Mining is an analytic process designed to explore data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. In other way, it is the extraction of hidden predictive information from large databases.

It is a powerful technology with great potential to help companies focus on the most information in their data warehouses. Data mining tools predict future trends and behaviors, allowing. Data mining has been very interesting topic for the researchers as it leads to automatic discovery of useful patterns from the database.

2.0Background Theory

2.1: Multilevel security or multiple levels of security (MLS):

It is the application of a computer system to process information with incompatible classifications (i.e..At different security levels), permit access by users with different security clearances and needs-to-know, and prevent users from obtaining access to information for which they lack authorization.

There are two contexts for the use of Multilevel Security.

Another context is to refer to an application of a computer that will require the computer enough to protect itself from subversion and possess adequate mechanisms to separate information domains, that is, a system we must trust. This distinction is important because systems that need to be trusted are not necessarily trustworthy.

2.2: Security concern in data mining :

Databases are important and essential components of different government and private organizations. To protect the data of the databases used in data warehouse and then data mining is central theme of security system.
The requirements of data mining security concerned with the following traits.

2.2.1.Physical Database Integrity:

This physical database integrity related with the power failure of the system. When power fails the intermediate records are not posted or retrieved correctly. Due to this the data mining becomes unable to predict pattern by given applications.

2.2.2.Logical Database Integrity:

This type of integrity indicates that modification of value of one field does not affect records. Whenever this occurs the data mining algorithm can not be able to predict correct information due to logical integrity anomalies with given database for data mining.

2.2.3.Element Integrity:

The integrity of each individual element is necessary for the database which is used for the data mining. If each element of database of data warehouse maintains the integrity, there is no chance for change by human mistake and by any other programs.

2.2.4.Auditability:

The modification of records and fields of the database are taken with OLTP (On line transaction processing applications and by the human operators or by database administrator. The date, time, fileds, records and the previous value of the records should have to be recorded under a log file. This ensures that the proper modification is taken on the database implemented under the data warehouse.

2.2.5.Access Control :

Database system has the capability for the access control. This access control ensures the access privileges of data items from the database. This means that who can read, modify, delete the records or individual fields of the database.
This access control is Security Measures in Data Mining defined by the database administrator for the users of the enterprise. If a user has only privilege to read the data items of database then he or she can only see the records but can not do anything others. The database administrator can have all types of privileges on the database .
It means he or she is database administrator then he or she can read, delete, modify the records, tables and others elements of the database

2.2.6.User Authentication:

Database management system requires the regroups user authentication. Without valid user identification number and password the database does not allow the user to do anything on data items of database. Each user has its own user authentication and identification entity. The user has to keep its user ID and password secret.

Related Work:

RSA algorithm:

RSA algorithm was introduced by three researchers fromMIT (Massachusetts Institute of Technology), namely RonRivest, AdiShamir, and Len Adleman in 1976. RSAencryption and decryption process based on the concept of prime numbers and moduloarithmetic. Both encryption anddecryption keys are both integers. The encryption key is keptsecret and is not known publicly (so-called public key), butkept secret key for decryption. Decryption keys are made ofmultiple primes together with the encryption key. To find thedecryption key, must be factored a composite number intoprime factors. In fact, factoring nonprime numbers into primefactors is not easy. There has been no efficient algorithmsfound in factoring it. The larger the number the more difficult to get the non prime factoring. The stronger is the RSA algorithm :

RSA algorithm having magnitudes as follows:

p and q are primes (secret)

n = pq (not secret)

(n) = (p - 1) (q - 1) (secret)

e (encryption key) (not secret)

d (decryption key) (secret)

m (plain text) (secret)

c (cipher text) (not secret)

Elliptical curve Cryptography(ECC)

Elliptic curve cryptography (ECC) is an approach to public-key cryptography based on the algebraic structure of elliptic curves over finite fields. Elliptic curves are also used in several integer factorization algorithms that have applications in cryptography, such as Lenstra elliptic curve factorization.The primary benefit promised by ECC is a smaller key size, reducing storage and transmission requirements, i.e. that an elliptic curve group could provide the same level of security afforded by an RSA-based system with a large modulus and correspondingly larger key – e.g., a 256-bit ECC public key should provide comparable security to a 3072-bit RSA public key.

Implimentation:

Key Generation

Key generation is an important part where we have to generate both public key and private key. The sender will be encrypting the message with receiver’s public key and the receiver will decrypt its private key.

Now, we have to select a number ‘d’ within the range of ‘n’.

Using the following equation we can generate the public key

Q = d * P

d = The random number that we have selected within the range of (1 to n-1). Pis the point on the curve.

‘Q’ is the public key and ‘d’ is the private key.

Encryption

Let ‘m’ be the message that we are sending. We have to represent this message on the curve. This have in-depth implementation details. All the advance research on ECC is done by a company called certicom.

Conside ‘m’ has the point ‘M’ on the curve ‘E’. Randomly select ‘k’ from [1 - (n-1)].

Two cipher texts will be generated let it be C1 and C2.

$C1 = k*P$

$C2 = M + k*Q$

C1 and C2 will be send.

Decryption

We have to get back the message ‘m’ that was send to us,

$M = C2 – d * C1$

M is the original message that we have send.

How does we get back the message,

$M = C2 – d * C1$

‘M’ can be represented as ‘$C2 – d * C1$′

$C2 – d * C1 = (M + k * Q) – d * (k * P)$ ($C2 = M + k * Q$ and $C1 = k * P$)

$= M + k * d * P – d * k *P$ (canceling out $k * d * P$)

= M (Original Message)

LZW(Lempel-Ziv-Welch Algorithm):

The LZW Compression Algorithm can summarised as follows:

```
w = NIL;
    while ( read a character k )
        {
          if wk exists in the dictionary
           w = wk;
          else
            add wk to the dictionary;
            output the code for w;
            w = k;
```

}

Advantages: LZW can be made really fast; it grabs a fixed number of bits from input stream, so bit parsing is very easy. Table lookup is automatic.

CONCLUSION

In this paper, the main approach is to classify the overall Financial Network information into different levels and to secure the classification through RSA and LZW algorithm. The entire Financial Network(m-commerce) System is adequately monitored through CPS for Fraud Detection And using clustering to classify the data in sequence manner using data mining concept.

Acknowledgements

I am deeply indebted & would like to express my gratitude to my thesis guide Prof. Indrajeet Rajput (Assistant Professor, HGCE) for his great efforts and instructive comments in the dissertation work.

He has devoted significant amount of his valuable time to plan and discuss the Literature review work. During the dissertation he provided me excellent guidance & support during discussion about my progress. He gave me a great deal of freedom to choose a research topic, focusing more on my research interests and building skills that will allow me to be a successful in my life, without his experience and insights, it would have been very difficult to do qualitative work.

I am obliged to Prof. Indrajeet Rajput, H.O.D of Computer Department of HGCE, for the valuable information provided by him in him respective fields. I am grateful for him cooperation during the period of my dissertation phase.

I would also like to thank all staff members and my colleagues who were always there at the need of hour and provided with all help and facilities, which I required. I would also like to thanks to all my friends for helping me in the hour of need and providing me all the help and supported me during this thesis work.

Last but not the least I am highly grateful to all my family members for their inspiration and ever encouraging moral support, which enables me to pursue my studies.

REFERENCE

[1] Sharada Varalaxmi Mangipudi, P. Suresh Verma, M. Srinivasa Rao” Securing Financial Network System through Multilevel Security Using Cyber- Physical System and Data Mining Concepts”978-1-4799-5958@2014 Ieee.

[2] Kyungroul Lee, Hyeungjun Yeuk, Habin Lim, and Kangbin Yim”Security threats to the platform idetification”978-1-4673- 8315-8@2015 Ieee.

[3] Prasad Seemakurthi, Shuhao Zhang, and Yibing Qi” Detection of Fraudulent Financial Reports with Machine Learning Techniques”978-1-4799-1832-4@2015 Ieee.

[4] Kaveh Paridari, Alie El-Din Mady, Silvio La Porta§, Rohan Chabukswar Jacobo Blanco, André Teixeira, Henrik Sandberg, Menouer Boubekeur” Cyber-Physical-Security Framework for Building Energy Management System”978-1-5090-1772-0@2016 Ieee.

[5] Dr.Vivek Kapoor ,Rahul Yadav"A Hybrid Cryptography Technique to Support Cyber Security Infrastructure."ISSN:2278-1323@2015IJARCET.

[6] Anish Gupta, Vimal Bibhu, Md. Rashid Hussain"Security Measures in Data Mining"@MECS2012.

[7] Luna M. Zhang” : Genetic Deep Neural Networks Using Different Activation Functions for Financial Data Mining”978-1-4799-9926-2@2015Ieee