ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

Privacy Preserving for Data Perturbation Using Cryptography Technique

A.Malathi[1], Dr.M.Malathi[2]

Research Scholar, Department of Computer Science, Government Arts College(Autonomous), Salem, India[1].

Assistant Professor, Department of Computer Science, Government Arts College(Autonomous), Salem, India[2].

ABSTRACT: Insider attack is originating from an authorized node that had first passed in all the authority steps to access the networking and then involved in compromised. Insider-related research involving the distribution of kernel-based data mining is limited, resulting in substantial vulnerabilities in designing protection against collaborative organizations. Specifically, if more of our assets are going to reside in the cloud, and as increasingly our lives, enterprises and prosperity may depend upon cloud, it is imperative that we understand the scope for insider attacks so that we might prepare best defenses.we are going to use String Matching algorithmA substring is a sequence of consecutive contiguous elements of a string, we will denote the substring starting at i and ending at j of a string .If the string has same value then only user can send and receive message .It can check all types of string for authorized. Whenever a sender sending the message to the receiver a unique key will be generated each and every time. The encrypted message will also send with a key during the message transformation. The message will be received as cypher text. The generated key is the source for reading the cypher text. When sending a message no intruder can get the key value. So it can be the highly secured transformation.

KEYWORDS: Mail submission agent(MSA), Mail user agent (MUA), Simple Mail Transfer Protocol(SMTP), Message-Digest(MD), String Matching.

I. INTRODUCTION

Data mining the extraction of hidden predictive information from large database, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer business questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

Most companies already collect and refine massive quantities of data. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources, and can be integrated with new products and systems as they are brought on-line. In this existing system used Insider-related research involving the distribution of kernel-based data mining is limited, resulting in substantial vulnerabilities in designing protection against collaborative organizations. Homomorphism encryption algorithm, Prior works often fall short by addressing a multi factorial model that is more limited in scope and implementation than addressing insiders within an organization colluding with outsiders. Such a pragmatic model considers the insider as the key player in sharing data with an attacker, who can then recover the original data from the intermediary kernel values of the SVM model. This attack is more realistic because the attacker needs only to obtain a few data entries rather than the entire database from an organization to successfully recover the rest of the private data. In our proposed system used string matching for detects the error.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

II. RELATED WORK

s refers to privacy preserving data mining has been studied extensively, because of the wide proliferation of sensitive information on the internet. We discuss method for perturbation, K- Anonymization, condensation, and distributed privacy preserving data mining [1].

Navithapokale et al depicts to we considering this assumption for expand the scope of perturbation based PPDM to multilevel Trust(MLT PPDM). The less perturbed copy can access permissions to the more trusted a data miner. Under this, a malicious data miner have access to different copies of miner have access to different copies of the same data through various forms, and it combine these copies to jointly infer additional metadata about the original data that the data owner does not intend to release[2]

K-anonymity provides privacy protection by guaranteeing that each released record will relate to at least k individuals even if the records are directly linked to external information . Latanya Sweeney et al when these data re linked together, they provide an electronic image of a person that is as identifying and personal as a fingerprint even when the information contains no explicit identifiers, such as name and phone number[3]

Damian Vizar et al the main contribution of Gentry's thesis was, that it has proven, that it actually is possible to design a fully homomorphic encryption scheme. However ground-breaking Gentry's result was, the designs, that employ the bootstrapping techniques user from terrible performance both in key generation and homomorphic evaluation of circuits[4].

Keke Chen at al Perturbation techniques are often evaluated with two basic metrics: level of preserved privacy guarantee and the level of preserved data utility. A data perturbation procedure can be simply described as follows. Before the data owners publish their data, they change the data in certain ways to disguise the sensitive information while preserving the particular data property that is critical for building meaningful data-mining models. The rotation perturbation and random projection perturbation are all threatened by prior-knowledge enabled Independent Component Analysis [5].

Privacy and accuracy in case of data mining is a pair of contradiction DarshnRathodl at al Privacy Preserving Clustering of Data Streams (PPCDS) is proposed stressing the privacy preserving process in a data stream environment while maintaining a certain degree of excellent mining accuracy. PPCDS is mainly used to combine Rotation Based Perturbation, optimization of cluster enters and the concept of nearest neighbor, in order to solve the privacy preserving clustering of mining issues in a data stream environment [6].

SlawomirGoryczka at alThe privacy constraint against any group of up to m colluding data providers. When the data are distributed among multiple data providers or data owners, two main settings are used for anonymization. One approach is for each provider to anonymize the data independently, which result in potential loss of integrated data utility[7].

Mr. SwapnilKadam at al Preserving Privacy in Data Mining(PPDM) technique introduces uncertainty about individual values before data are published or released to third parties for data mining purposes. In the single level trust assumption, the data owner or admin can create only single perturbed copy of its data with a fixed amount of uncertainty[8].

Stanley R. M. et al The privacy problem against unauthorized secondary use of information. The introduce a family of geometric data transformation methods(GDTM) which ensure threat the mining process will not violate privacy up to a certain degree of security. We focus primarily on privacy preserving data clustering, notably on partition-based and hierarchical methods[9].

V.V. Nagendrakumar et al it deals with the issue of privacy preserving in data mining while collaborating n number of parties and trying to maintain confidentiality of all data providers details while collaborating their database. They two type of attacks are addressed "insider attack" and "outsider attack". In insider attack, the data providers use their own records and try to retrieve other data provider details [10].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

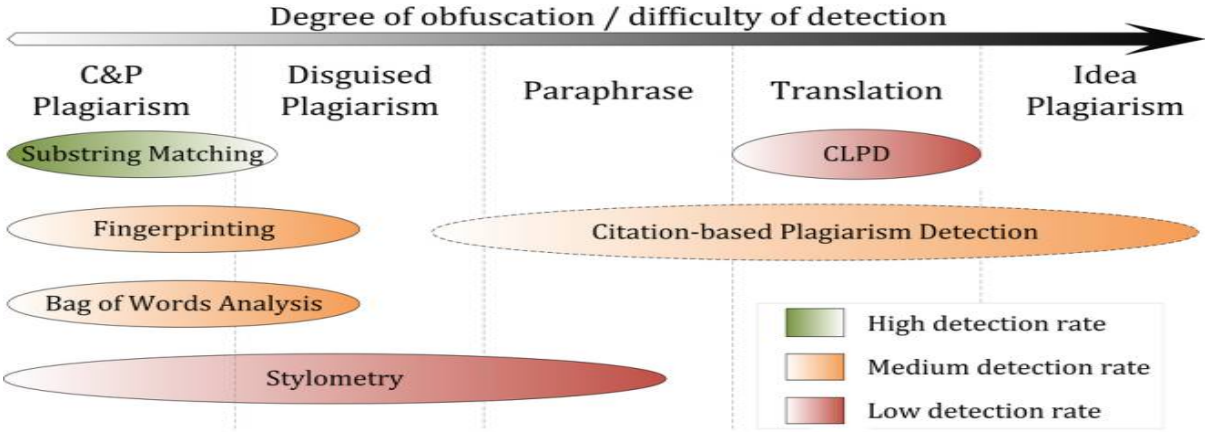
Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

III. METHODOLOGY

A.String Matching Algorithm

String Matching have a high detection rate by which you can easily detect a hacker.



A string is a sequence of characters. In our model we are going to represent a string as a 0-indexed array. So a string S =”Friend” is indeed an array [‘F’, ’r’, ’i’, ’e’, ’n’, ’d’]. The number of characters of a string is called its length and is denoted by |S|. If we want to reference the character of the string at position i, we will use S[i].

B.Authority Key Identification

Using String Matching, we match the particular id to create the regular identify an d the unique with password then only message transfer to other end for particular user. Other end user, login and then only they can study the private message from user identification. Every user has a separate random key .If that intruder not have that separate key, then that user unable to view message and send that message. Using MD5 we can terminate intruder without having key Value that intruder can’t view or send message Data.

C. Message Transfer

A message transfer agent receives mail from either that the regular transfer to across the simple mail transfer to obey the valid and to another MTA, a mail submission agent that which the agent of the regular (MSA), or a transfer of data in mail user agent (MUA). The transmission details in which that have specified by protocol the Simple Mail Transfer Protocol (SMTP). When a recipient mailbox of a message is not hosted locally, the message is relayed, that is, forwarded to another MTA. Every time an MTA receives a key with the generation to logic key to email message, it adds a received trace header field to find the local and specific creation of the specific task at the top of the header of the message, thereby building a sequential record to ideas of the particular notification of MTAs handling the message. The process of choosing a basic language in which the particular target MTA for the next hop and in the mid hop that which in the particular nodes to transfer to analysis in the problem of the maintain the regular modification of the specification also described in SMTP, but can usually have to regulate be overridden by ordering the configuring the MTA software in a proper manner and with specific routes.

D. Reducing the number of the insiders

In information security, intruder detection that the specific to modify the root of the key generation is the art of detecting and to produce the regular intervals of times intruders behind the node is the attacks as unique persons. This technique tries to identify from address of the particular mail that which the regular mail id the person behind an attack by analyzing their component the original key is to be in the computational behavior. This concept is sometimes that to have the modify specific and arrange of the confused with Intrusion Detection techniques which that can be able to carry the art of detecting intruder actions.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

E. Terminate Intruder

By Using MD -5 (Message-Digest) algorithms, Receiver gets the message and extracts the encrypted message digest. Then he computes his own message digest of the received message. He also decodes received message digest with sender's public key and gets decoded message digest. Then he compares both message digests when both message digests are equal, the message was not modified during the data transmission.

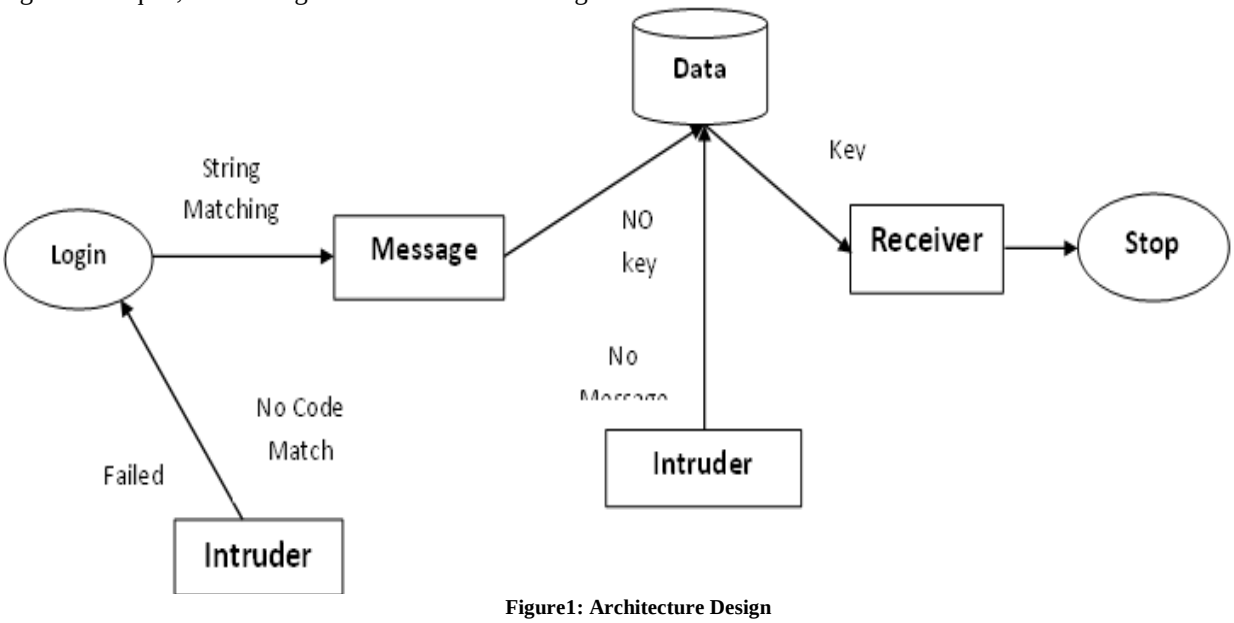


Figure1: Architecture Design

IV. RESULT AND DISCUSSION

It is the permanence analysis of security level in existing and proposed system.

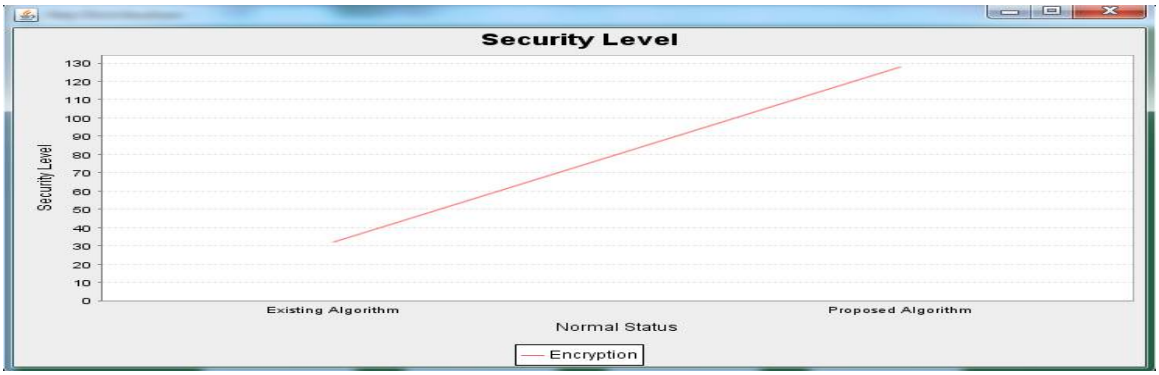


Figure2: Performance analysis chart

V. CONCLUSTION

We propose an insider collusion attack that is a and to maintain the security of the message can be get solved in the solution threat to most data mining systems the main analysis is to generate the basic key values of the using the solution of that operate on kernels and discuss how many insiders are attacks that which and to launch sufficient to launch this type of attack. We also present two privacy-preserving methods to defend against the attack is get defends the solution of the various necessary of the launch the solution of the key generation and attack.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

REFERENCES

[1] Ms. R. Kavitha, andProf.D. Vanathi," A Study Of Privacy Preserving Data Mining Techniques",, Volume 3, No.4, July - August 2014.
[2] Mayil.S, and Vanitha.M, "A Survey on Privacy Preserving Data Mining Techniques"International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014, 6054-5056
[3] Latanya Sweeney,"Achieving k-anonymity privacy Protection Using Generalization and Suppression", International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), 2002; 571-588.
[4] Damian Vizar and Serge Vaidemau "Cryptanalysis of Chosen symmentricHomomorphic Schemes" ,EPFL CH-1015 lausanne, Switzerland http://lasec.epfl.ch
[5] kekechen, and Ling Liu "Geometric data perturbation for privacy preserving outsourced data mining", Received: 10 March 2010 / Revised: 3 October 2010 / Accepted: 4 November 2010© Springer-Verlag London Limited 2010
[6] DarshnaRathodl, and AvaniJadeja "Geometric Data Perturbation using Clustering Algorithm", international journal of advances in cloud computing and computer science (ijacccs) issn(print): 2454-406x,(online): 2454-4078,volume-1,issue-1,2015
[7] SlawomirGoryezka, and Li Xiong,et al "m-Privacy for Collaborative Data Publishing" ,1http://www.hhs.gov/healthit/healthnetwork/background/
[8] Mr. Swapnilkadam, and Prof. NavnathPokale"Privacy Preserving through Data Perturbation Using Random Rotation Based Technique in Data Mining", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)Volume 5 Issue 1, January 2016
[9] Stanley R.M. Oliveria, and Osmar R. Zalane "Privacy Preserving Clustering By Data Transformation"
[10] V.V. Nagendrakumar, and C. Lavanya "Privacy -Perserving for Collaborative Data Publishing",(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4566-4569