SECURE DATA TRANSMISSION FOR BIG DATA RECORDS BY USING DES AND SDH ALGORITHMS

A.Theresa Vinothini[1], K.Kowsalya[2], M.Monika Victoriya[3], M.Parkavi[4]

[1]Assistant Professor, Dept of Information Technology, Jeppiaar SRR Engineering College, Tamilnadu, India
[2]Dept of Information Technology, Jeppiaar SRR Engineering College, Tamilnadu, India.

---***---

Abstract - *We formulate the several types of Relational Database securities as a constrained and an efficient techniques to solve the security problems. In existing system, no securities was provided to the relational datasets and only the classification and clustering of data was done .Data transmitted through the network from one server to another, so the attackers had possibilities to easily attack the dataset without any difficulties. We address this problem of securely sending provenance for data security using J-Bit Encoding (JBE), DES algorithm and Secure Data Hiding Algorithm (SDHA). MS SQL framework will be used in our proposed system which will be flexible to provide securities. In J-Bit Encoding each bit will be encoded and compression of data will be done. Then DES encryption (Data Encryption Standard) will be implemented. Finally the SDHA will be implemented to embed the hidden data using text. The text can be known as cover text which will be embedded with original data to prevent alterations to original data. These security features will provide an efficient way of providing security for data transmission without any loss.*

***Key Words**:* **Relational database, Classification Clustering , Encoding ,Encryption ,Framework Cover text , Data transmission.**

1. INTRODUCTION

Nowadays handling big data plays vital role in the technical world. But the rapid growth of the internet and its related technologies has offered an unprecedented ability to access the digital contents .There is a big deal to solve these security problems and have to provide a relational database which cannot be hack by anyone in the internet world. Thus having an ownership and encrypting the content is the ultimate solution to secure the data. Bid data is a dataset that are so voluminous and complex that compromises traditional data processing application software which are inadequate to deal with them. Here some of the technique and algorithm has been applied to big data. Currently most of the vendors like Bank, IT, Healthcare managing and maintaining their database on big data. Big Data increases both the number of data sources and variety and volume of data that is used for analysis.

1.1Data Streaming

Most of the machine learning problems are often characterized not only by a significant volume of data, but also by its velocity. Instances may arrive continuously in a form of a potentially unbounded data streams[4]. This poses new challenges for learning algorithms, as they must offer adaptation mechanisms for ever-growing dataset, being able to update their structure in accordance with the current state of a stream [2].

1.2 J-Bit Encoding

Mostly People tend to store lots of files in their storage. when the storage nears it limit, then they try to reduce those file size to minimum by using compression software .here we introduce a new algorithm for data compression called J-Bit Encoding(JBE) which compress and encode data without any loss.

1.3 Data Encryption Standard

Data Encryption Standard is a symmetric-key algorithm for the encryption of data. Totally,16 rounds takes place in 64-bit block length of 56-bit key length which makes the data highly secure.

1.4 Secure Data hiding Algorithm

Data hiding techniques have been widely used to transmission of hiding secret message for a long time. Ensuring data security is a big challenge for computer users and also for the transmission of big data. Here the encrypted data be hidden with a cover text.

2. PROBLEM IDENTIFIED

In early days, Data mining concepts has been used to mine a massive datasets within a server. Some mining techniques are used to retrieve a data from a volume of data in a particular server. No security issues were created while analyzing the storage of server. Nowadays moving of an enormous amount of data to another server with security is quite difficult. Transmitting huge amount of data from server to another server can cause performance delay or server may get jammed. Some data's might also get lost due to streaming into a buffer of destination server. Lack of security paves a way for many attacks.

Nearest neighbor algorithm to classify high-speed and massive data streams is implemented[4]. However, due to their lazy learning nature and high computational cost they have not gained significant attention in the domain of data stream analysis. Data searches implemented in Apache Spark[2], includes a distributed metric-space ordering to perform faster searches. Fully focus on the data storage reduction for data streaming within the storage. Partitioning of the data streams was performed by using Spark environment. These operations are designed to transform datasets by locally executing tasks within the data partitions, thus maintaining the data locality.

3. PROPOSED MODEL

The Architecture Diagram shown depicts the transfer of data from one server (sender) to another server(receiver) via Network. Server Authentication is done on both sides (receiver and sender) to check whether the user is valid. It is done by using login page which contains username and password. After authentication process Data Partitioning is done, in which data is partitioned based on Row Range Partitioning. Row Range Partitioning is the division of rows into set of rows. Set of rows consist of three rows. Followed by Data Partitioning J-Bit Encoding is done. J-Bit Encoding is the compression of data and loss of data will be reduced. The divided data will be compressed to specific size. The Compressed Data then will undergo the Encryption Process in the DES Algorithm which is a Symmetric Key Algorithm. The key size used here is 56-bits, block size is 64 bits and we have totally 16 rounds in this Algorithm. The encrypted data can be hided into a text which will be invisible to the hacker. This text is called Cover Text. This is done by the algorithm called Secure Data Hiding (SDH Algorithm). In this algorithm data can be hided within text, image, audio and video. Here we hide it using text. We use four different techniques in Sender side. In Receiver side, the reverse process of these techniques is done. After Server Authentication (Receiver) the reverse of SDHA is done where the hided data will be revealed. The Decryption process will be done to the revealed data using the DES Algorithm where the output is received in fixed length. The Decrypted data will undergo the Decoding process which is called the J-Bit Decoding. Here the Compressed data will be Decompressed. Finally Reverse Partitioning of data is done where the partitioned data will be combined. Thus the data will be received by the receiver .The Main Advantage of this Project is, we use three different techniques to the partitioned data so that the security is enhanced.

4. PROPOSED DIAGRAM

The proposed model block diagram is represented as follows:

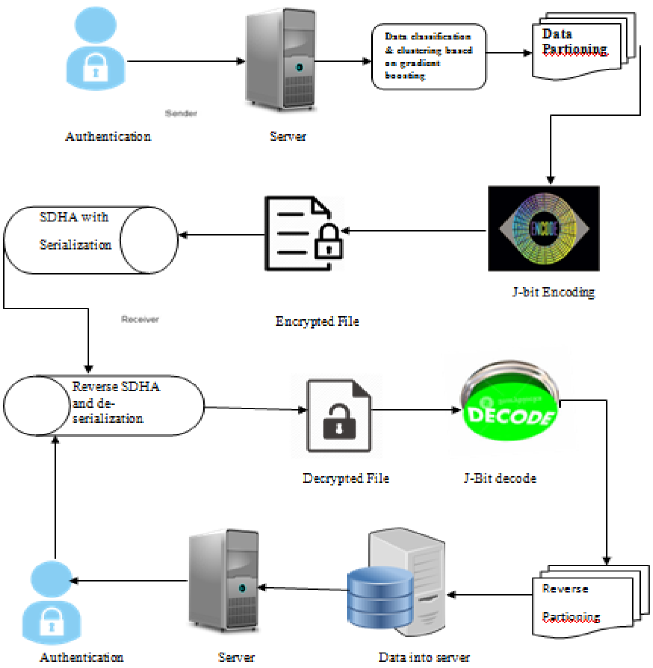

Fig1: Architecture Diagram

5. CONCLUSION

The final result is transmission of data from one server to other server without any data loss. Secure data streaming transmission is done. Loss of data can be reduced. Speed of transmission of big data stream records is increased. The multiple securities are provided for security purpose and to prevent the attacks from the attackers. Data loss can be prevented by using J-Bit Encoding. Using SDH Algorithm hides the data which confuses the Hackers. Although DES has been used in many existing paper, it adds an additional security to this project. Data Partitioning is done initially which actually strengthens the security and makes the hacking process complex. Therefore, usage of J-Bit Encoding, DES Algorithm, SDH Algorithm, and the process of Data Partitioning will increase the security in transmission of large amount of data and the receiver will receive the actual (original) data.

6. REFERENCES

[1] W.-P. Ding, C.-T. Lin, M. Prasad, S.-B. Chen, and Z.-J. Guan, "Attribute equilibrium dominance reduction accelerator (DCCAEDR) based on dis-tributed coevolutionary cloud and its application in medical records," IEEE Trans. Syst., Man, Cybern., Syst., vol. 46, no. 3, pp. 384–400,Mar. 2016.

[2]X. Meng et al., “Mllib: Machine learning in apache spark,” J. Mach.Learn. Res., vol. 17, no. 1, pp. 1235–1241, 2016.

[3]Y. Xun, J. Zhang, and X. Qin, “FiDoop: Parallel mining of frequent itemsets using MapReduce,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 46, no. 3, pp. 313–325, Mar. 2016.

[4]D. Han, C. G. Giraud-Carrier, and S. Li, “Efficient mining of high-speed uncertain data streams,” Appl. Intell., vol. 43, no. 4, pp. 773–785, 2015.

[5]H. Karau, A. Konwinski, P. Wendell, and M. Zaharia, Learning Spark: Lightning-Fast Big Data Analytics. Sebastopol, CA, USA: O’Reilly Media, 2015.