Transmission of Relational Data Set for Stream Records by Multisecure Attribution

Dr. Janani.V* M.E, Ph.D Leelavathi. S**, Lethraa Devi. M**, Saranya. M**

*Associate professor, Department of CSE,vajjiram.janani@gmail.com.

** Final year students, Department of CSE, leeluprem137@gmail.com

lethraadevi97@gmail.com, saranyamahi106@gmail.com

ADHIYAMAAN COLLEGE OF ENGINEERING, HOSUR

***Abstract*:-**Transmission of relational data set for stream records by multisecure attribution is a novel streaming for applying to large data set and includes three components: a new data management scheme for processing and storing the incoming data, a layout construction algorithm specifically designed for incrementally generating storylines from large data set, and a layout refinement for improving the legibility of the visualization. Secure Data Hiding Algorithm (SDHA) have emerged as an important building block that plays a crucial role in addressing the fake record (Prevent from fake data) problem. The system the novel problem of securely sending provenance for data security using J-bit encoding, DES algorithm and Secure Data Hiding Algorithm (SDHA), A SDHA describes information that can be used to prove the ownership of data such as the owner, origin, or recipient of the content. Secure embedding requires that the embedded hiding must not be easily tampered with, forged, or removed from the SDHA data. SDHA have been developed for video, images, audio, and text data and also for software and natural language text. SDHA embedding for relational data is made possible by the fact that real data can very often tolerate a small amount of error without any significant degradation with respect to their usability. In particular the proposed technique is resilient to tuple deletion, alteration, and insertion attacks.

Keywords*:-*SDHA,Layout construction and refinement algorithm, Watermarking, secure attribution

*****

Introduction

The rapid growth of the Internet and related technologies offered an unprecedented ability to access and redistribute digital contents. In such a context, enforcing data ownership is an one of the important requirement, which requires articulated solutions, encompassing technical, organizational, and legal aspects. Although the system is still far from such comprehensive solutions, in last years, watermarking techniques have emerged as an important building block that plays a crucial role in addressing the ownership problem. Such techniques allow the owner of the data to embed an imperceptible watermark into the data. A watermark describes information that can be used to prove the ownership of data such as the owner, origin, or recipient of the content. Secure embedding requires that the embedded watermark must not be easily tampered with, forged, or removed from the watermarked data. The Imperceptible embedding means the presence of watermark is unnoticeable in the data. Furthermore, the watermark detection is blinded, that is, the technique neither requires the knowledge of the original data nor the watermark. Watermarking techniques have developed for video, images, audio, and text data and also for software and natural language text. By contrast, the problem of watermarking relational data has not been given appropriate attention. There are, however, many application contexts for which data represent an important asset, the ownership of which must thus be carefully enforced. This is the case, for example, of weather data, stock market data, power consumption, consumer behavior data, and medical and scientific data. Watermark embedding for relational data is made possible by the fact that real data can very often tolerate a small amount of error without any significant degradation with respect to their usability. For example, when dealing with weather data, changing some daily temperatures of 1 or2 degrees is a modification that leaves the data still usable.

only a few approaches to the problem of watermarking relational data have been proposed. These techniques, however, are not very resilient to watermark attacks. The paper, presents a watermarking technique for relational data that is highly resilient compared to these techniques. In particular, The proposed technique is resilient to tuple deletion, alteration, and insertion attacks. The main contributions of the paper is to formulate the watermarking relational databases as a constrained optimization problem and discuss efficient techniques to handle the constraints

Problem statement.

LITERATURE SURVEY:

As increasing amounts of data are produced, packaged and delivered in digital form, in a fast,networked environment, one of its main features threatens to become its worst enemy: zero-cost verbatim copies. The ability to produce duplicates of digital Works at almost no cost can now be misused for illicit profit. This mandates mechanisms for effective rights assessment and protection. One such mechanism is based on Information Hiding. By concealing a resilient rights holder identity "signature" (watermark) within the digital Work(s) to be protected, Information Hiding for Rights Assessment (Watermarking) enables ulterior court-time proofs associating particular Works with their respective rights holders. One main challenge is the fact that altering the Work in the process of hiding information could possibly destroy its value. At the same time one has to be concerned with a malicious adversary, with the major incentives to remove or alter the watermark beyond detection – thus disabling the ability for court-time proofs – without destroying the value of the Work – to preserve its potential for illicit profit. The proposed system explains how Information Hiding can be deployed as an effective tool for Rights Assessment for discrete digital data. More specifically, numeric and categorical relational data is discussed.

The model aims at preserving "classification potential" of each feature and other major characteristics of datasets that play an important role during the mining process of data; as a result, learning statistics and decision-making rules also remain intact. We have implemented our model and integrated it with a new watermark embedding algorithm to prove that the inserted watermark not only preserves the knowledge contained in a dataset but also significantly that enhances watermark security compared with existing techniques[1]. The major contribution of this paper is a robust and efficient watermarking scheme for relational databases that is able to meet all above-mentioned four challenges. The results of our experiments prove that the proposed scheme achieves 100 percent decoding accuracy even if only one watermarked row is left in the database[2].

The proposed system explains a new lossless reversible watermarking approach that allows embedding of the message within categorical data of relational database. The reversibility property of our scheme is achieved by adapting the well known histogram shifting modulation. Based on this algorithm we derive a system for verifying the integrity of the database content, it means detecting addition, removal modification of tuples or attributes[3].

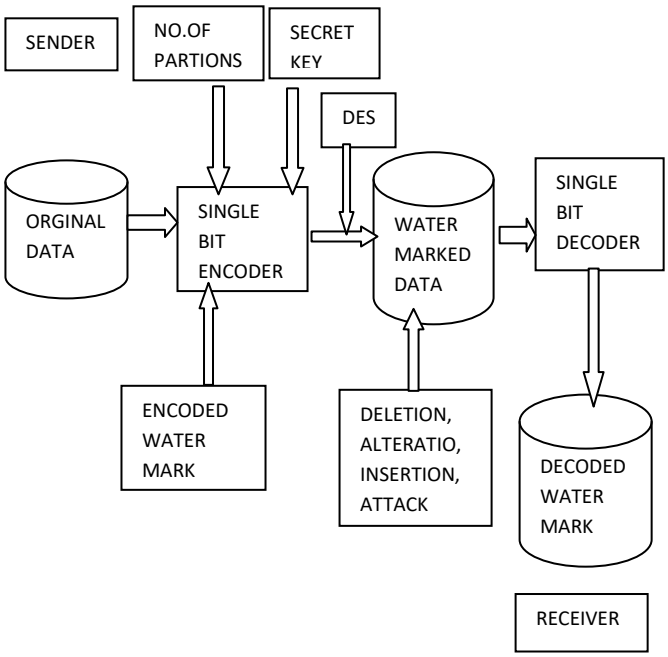
The proposed solution is based first on an adequate selection of XML locators,i.e, document fragments targeted to embed the watermark. Watermark retrieval is achieved thank to a set of personalized fuzzy queries that reconstruct locators that contain the watermark. We show theoretically the watermark robustness against possibleXML file transformations; then, we present some initial experiments that validate the approach[4].

The resulting scheme modulates the angular position of the circular histogram center of mass of one numerical attribute for message embedding. It verifies database authentication as for traceability when identifying database origin after it has modified. Beyond the application framework, the system theoretically evaluate the performance of the scheme in terms of capacity, distortion, robustness against common databasemodifications[5].

Existing system:

Only the encoding and decoding is used for security purpose.The streaming of data is in the block of data. Streaming transmission setup is considered where an encoder observes a new message in the beginning of each block and a decoder sequentially decodes each message after a delay of Time blocks. The technique does not provide mechanism for advanced securities.Only the encoding and decoding can be easily compromised by very trivial attacks. Attackers easily capture the whole data's by less security.It can handle only small size of dataset for transmission.

Architecture Diagram



Proposed system:

The proposed system address the novel problem of securely transmitting provenance for data streams in large data using DES algorithm.J-bit encoding is used for the encoding data process. Dataset will do Partition and

grouping for encoding for processing however, unlike traditional watermarking approaches, we embed provenance over the inter packet delays (IPDs) it avoids the problem of data degradation due to watermarking. The system propose a data hiding-based solution that embeds provenance over the inter packet delays. The security features of the scheme make it able to survive against various network or flow attacks.Here the embedding data is hidden the original content with cover text (Fake content) is invisible to the user.

Authentication:

The process of identifying an individual, usually based on a username and password. In security systems, authentication is distinct from authorization , which is the process of giving individuals access to system objects based on their identity. Authentication merely ensures that the individual is who he or she claims to be, but says nothing about the access rights of the individual.

Data Partioning:

A partition division logical database or its constituent elements into distinct independent parts. Database partitioning is normally done for manageability,performance or availability reasons, or for load balancing.

J-Bit Encoding:

J-bit encoding (JBE) works by manipulate bits of data to reduce the size and optimize input for other algorithm. The main idea of this algorithm is to split the input data into two data where the first data will contain original nonzero byte and the second data will contain bit value explaining position of nonzero and zero bytes. Both data then can be compress separately with other data compression algorithm to achieve maximum compression.

DES Encryption:

Encrypt the portioned and conceded data. The Data Encryption Standard (DES) is a symmetric-key block cipher published by the National Institute of Standards and Technology (NIST).DES is an implementation of a Feistel Cipher. It uses 16 round Feistel structure. The block size is 64-bit. Though, key length is 64-bit, DES has an effective key length of 56 bits, since 8 of the 64 bits of the key are not used by the encryption algorithm

SDHA SERIALIZATION:

This class serialize (Convert The Object into file) and sends it to the destination via network connections.“secure hiding information technique within a noise; a way to supplement encryption, to prevent the existence of encrypted data from being detected”. Encryption provides an approach to information security, and encryption programs readily available. However, encryption clearly marks are a message as containing “secret data” information, and the encrypted message becomes subject to attack. Many cases it is desirable to send secret information without anyone noticing that information has been sent is secret information.

Client (Destination) Module

Authentication

Here the username and password are validated by user inputs. The authentication that performsvalidation user is authenticated or not**.**

Data Partioning

This class eventually divides the table records and assigns partition number to it.

J-Bit encoding

This class encodes the partitioned file by adding one bit to each record. The main idea of this algorithm is to split the input data into two data.

DES Encryption

Encrypt the portioned and conceded data.It uses 16 round Feistel structure. The block size is 64-bit. Though, key length is 64-bit, DES has an effective key length of 56 bits, since 8 of the 64 bits of the key are not used by the encryption algorithm

SDHA SERIALIZATION

This class serialize(Convert The Object into file)and sends it to the destination via network connections.Encryption provides an approach to information security, and encryption programs readily available

CONCLUSION

The proposed project presentes a resilient watermarking technique for relational data that embeds watermark bits in the data statistics. The watermarking problem was formulated as a constrained optimization problem that maximizes or minimizes a hiding function based on the bit to be embedded. GA and PS techniques were employed to solve the proposed optimization problem and to handle the constraints. The proposed system developes an efficient

threshold-based technique watermark detection that is based on an optimal threshold that minimizes the probability of decoding error. The watermark resilience was improved by the repeated embedding of the watermark and using majority voting technique in the watermark decoding phase. Moreover, the watermark resilience was improved by using multiple attributes.A proof of concept implementation of our watermarking technique was used to conduct experiments using both synthetic and real-world data. A comparison our watermarking technique with previously posed techniques shows the superiority of technique to deletion, alteration, and insertion attacks.

References

[1] A formal usability constraints model for watermarking of outsourced datasets," IEEE Trans. Inf. Forensics Security, vol. 8, no. 6, pp. 1061–1072, Jun. 2013.
[2] A robust,distortion minimizing technique watermarking relational databases using once-for-all usability constraints," IEEE Trans. Knowl. Data Eng., vol. 25, no. 12, pp. 2694–2707, Dec. 2013.
[3] Lossless watermarking of categorical attributes for verifying medical data base integrity," in Proc. Annu. Int. Conf. IEEE EMBC, Aug./Sep. 2011, pp. 8195–8198.
[4] Query-preserve watermarking relational database XML documents," ACM Trans. Database Syst., vol. 36, no. 1, 2011, Art. ID 3.
[5] Sahai, "Anytime information theory," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), 2001.
[6] R. Sukhavasi and B. Hassibi, "Linear error correcting codes with anytime reliability," in Proc. IEEE Int. Symp. Inform. Theory (ISIT), Jul.-Aug. 2011, pp. 1748–1752.
[7] A. Khisti and S. C. Draper, "The streaming-DMT of fading channels," IEEE Trans. Inf. Theory, vol. 60, pp. 7058–7072, Nov. 2014.
[8] S. C. Draper and A. Khisti, "Truncated tree codes for streaming data: Infinite-memory reliability using finite memory," in Proc. International Symposium on Wireless Communication Systems (ISWCS), Nov. 2011, pp. 136–140.
[9] A. Sahai, "Why do block length and delay behave differently if feedback is present," IEEE Trans. Inf. Theory, vol. 54, pp. 1860–1886, May 2008.
[10] S.-H. Lee, V. Y. F. Tan, and A. Khisti, "Streaming data transmission in the moderate deviations and central limit regimes," IEEE Transactions on Information Theory, vol. 62, no. 12, pp. 6816–6830, Dec 2016.
[11] E. A. Haroutunian, "A lower bound of the probability of error for channels with feedback," Problemy Peredachi Informatsii, vol. 13, pp. 36–44, 1977.
[12] L. Zhou, V. Y. F. Tan, and M. Motani, "Second-order and moderate deviation asymptotics for successive refinement," IEEE Trans. Inf. Theory, accepted for publication, 2017.
[13] E. MolavianJazi and J. N. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," IEEE Trans. Inf. Theory, vol. 61, pp. 6719–6733, Dec. 2015.
[14] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Feedback in the non- ´ asymptotic regime," IEEE Trans. Inf. Theory, vol. 57, pp. 4903–4925, Aug. 2011.
[15] S. L. Fong and V. Y. F. Tan, "Asymptotic expansions for the AWGN channel with feedback under a peak power constraint," in Proc. IEEE Int. Symp. Inform. Theory (ISIT), Hong Kong, June 2015, pp. 311–315