International Journal of Advance Research in Science and Engineering
Vol. No. 5, Special Issue No. 01 , May 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

COMPARATIVE ANALYSIS OF RDBMS WATERMARKING TECHNIQUES

Dr. Nidhi H. Divecha[1], Bhavna Patel[2], Bijal Parmar[3]

[1] S. K. Patel Institute of Computer Studies, [2,3] B. P. College of Computer Studies,
Kadi Sarva Vishwavidyalaya,Gandhinagar, Gujarat, (India)

ABSTRACT

All commercial applications widely use Relational databases. Unauthorized modification to database can causes serious consequences and may be significant losses for the organization. An important area such as payroll, inventory, students' marks, defense and finance requires trustworthy method for checking data modification and integrity. Watermarking for relational databases emerged as a vital solution to provide copyright protection, tamper detection, traitor tracing and to maintain the integrity of data. This paper focuses on the comparative analysis of relational database watermarking techniques proposed by researchers also explains background knowledge of watermarking relational databases, such as types of attacks, requirements and basic techniques.

Keywords: Database, Security, Rdbms Watermarking, Database Copyright Protection, Secrete Key.

I. INTRODUCTION

Security and integrity of relational database is great concern now a day because of sharing of data over internet. Database creators provides services and make them available to users through internet, these services may attract more attacks. The database communication over internet also increase issues related to security of database. Ownership rights protection of relational database is crucial concern because unauthorized changes to data may have serious consequences and result in significant losses for the organization. Hence database creators need some technology that identifies the threats, pirated copies and unauthorized access to their databases. The increasing use of relational database creates a need for watermarking database. The purpose of Digital Watermarking is to protect a data from unauthorized duplication and distribution by enabling provable ownership over the data. Watermarking techniques allow the owner of the database to insert a watermark into the database. The ownership of data can be proved by a watermark. Secure watermark embedding requires that the watermark must not be easily forged, tampered with or removed from the watermarked database. The presence of watermark is unnoticeable in the database. Watermarking techniques have been developed for multimedia contents such as test, images, audio, video and suitable combination of them. It is also used for natural language text and software package. Relational database has exclusive and multifaceted requirement which makes relation data watermarking quite different from multimedia watermarking.

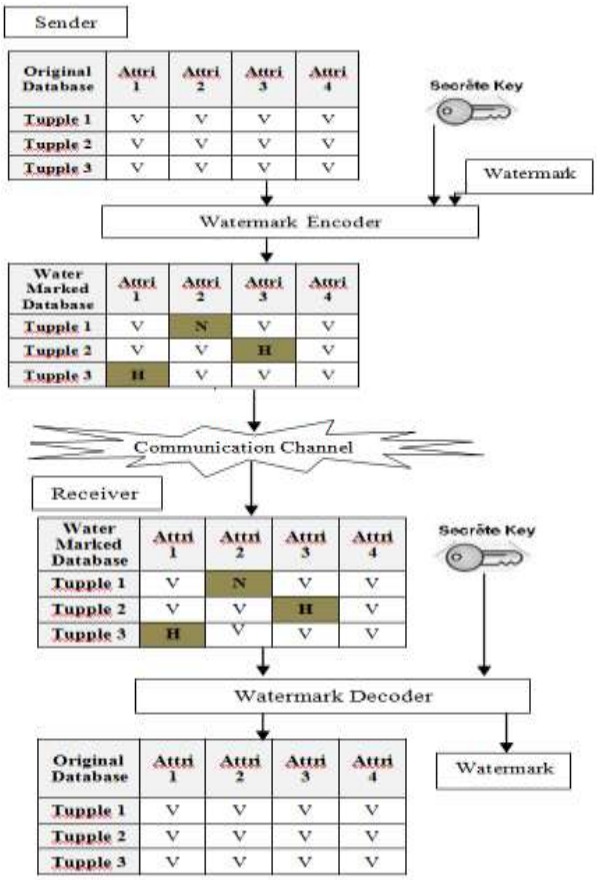
In relational database redundancy is very less which tends to narrow bandwidth of noise region available for embedding watermark information. It is very difficult to find out space to hide watermark within relational data. In addition,

database watermarking algorithms insist high standards of robustness because of normal update, manipulations or malicious attacks to database may result in loss of watermarking information.

Relational database watermarking technique is challenged for frequent updating, relational data out-of-order and data redundancy fewness. Due to such unique requirements and challenges, Watermarking relational databases literature is very limited and it focuses mainly on inserting binary bits in randomly chosen places in databases.

II. RELATIONAL DATABASE WATERMARKING MODEL

Relational database digital watermarking methods include two aspects: watermark encoding and watermark decoding. The elements in the process of watermark encoding are original database, watermarking information, key and an algorithm of



watermark encoding; simultaneously elements in the process of watermarking decoding include watermarked database, key and an algorithm of watermark extraction. These two processes are converse to each other as shown in Figure 1.

Figure.1. Database Watermark Encoding and Decoding Model

Watermark encoding phase includes a private or secrete key K (known only to the owner of database), which is used to embed the watermark bits information into the original database to shape up watermarked database. The watermarked database is then made available for public. To verify the right ownership of a doubtful database, the database verification process is performed. In this process the suspicious database is taken as input and by using the secrete key K which is used during the embedding phase, the embedded watermark (if present) is extracted from watermarked database and it is compared with the original watermark information.

Relational Database Watermarking Characteristics are:

Detect ability: The owner of the database should be able to detect the watermark by examining the tuples from the suspicious database.

Robustness: Watermarking scheme should be robust against different types of malicious attacks. The watermark should be embedded in such a way that it should be difficult for an attacker to delete or alter the watermark from the database without violating the knowledge of the data.

Usability: Data should be meaningful after watermark embedding. Watermarking technique should not result in distortion of data. Information and knowledge in the databases should be preserved.

Blindness: Watermark extraction should not require the knowledge of the original database and watermark itself.

Security: Database tuples, attributes, bit positions that are selected for embedding watermark bits should be kept secrete and it should be only known by having the knowledge of a secret-key means owner of the database.

Updatability: The watermark embedding should be done in such way that either the tuples of the relational database are inserted or deleted; the watermark value should not be changed.

III. RELATIONAL DATABASE WATERMARKING ATTACKS TYPE

The watermarked database may suffer from various kinds of intentional and unintentional attacks which may damage or remove the watermark, as described underneath.

Benign Update: The tuples or data of any watermarked relation are processed as usual. As a result, the marked tuples may be added, deleted or updated which may remove the embedded watermark or may cause the embedded watermark undetectable[1].

Deletion Attack: The Attacker deletes marked tuples from the relational database which leads to synchronization errors.

Alteration attack: Attacker alters the data values of the tuples which leads to disturbance in the watermark. Altering the data values violates the usability constraints and makes the data useless. For example an attempt of destroying watermark can be made by rounding all the values of numeric attributes.

Insertion Attack: Attacker inserts tuples to the data set hoping to disturb the embedded watermark which results in synchronization errors.

Subset Attack: By deleting or updating some of the tuples (subset) of the database attacker may try to destroy the watermark

Superset Attack: Some new attributes or tuples are added into the database.

Subset Reverse Order Attack: By changing the order or position of the tuples the attacker try to erase or disturb the watermark.

Mix-and-Match attack: The attacker collects related information from a different relation to build his own relation.

IV. COMPARATIVE ANALYSIS OF WATERMARKING RELATIONAL DATABASES

R. Agarwal et al. [1] proposed a watermarking technique that marks the only numeric attributes and tolerate some changes in some of the values. The basic marking idea is to allow errors in numeric attribute values. The process has mainly two major parts: first, based on the secrete key provided by user, the primary key value of tuple and percentage of tuples need to be marked, also determine elements to mark with the usage of one-way hash functions, such as SHA function and MDS function; second, identify attributes and positions of bit need to be marked based on the number of attributes and bits that could be marked and then mark part of certain attributes and bit positions which meet eligibility requirements as 0 or 1. The percentage of tuples, dimensions of attributes, Number of bits can be marked and the secrete key are only known by owner of the relational database. In that case, the security of database digital watermark algorithm has been largely improved. Watermarking information embedded in the database could be extracted only by getting the key. This algorithm embeds the watermark bits in the least significant bits(LSB) of selected attributes of a selected subset of tuples The one who has access to the primary key, can detect the watermark hence very high secure but can not provide robustness against common database attacks.

R. Sion et al. [2] proposed a watermarking algorithm that embeds the watermark into the relational database using data portioning technique. In data partitioning technique, all tuples securely divided into non-intersecting subsets. A single watermark bit embedded into tuples of a subset by modifying the distribution of tuples values. The same watermark bit embedded repeatedly across several subsets and the majority voting technique employed to detect the watermark. It is not resilient to synchronization errors and alteration attacks.

Z. Zhang et al. [3] had developed a watermarking algorithm in which an image is embedded into the relational data for representing the copyright information. The image contains bits which are used to represent the watermark bits. These bits are embedded in selected locations in database and if these bits are recovered correctly then it can be used to reconstruct an embedded image. This watermarking technique can be considered as a sub-class of the bit-level watermarking algorithm. This algorithm takes input as the relation R with the attributes as R(K.A0,A1,…..An) where K is the primary key of the database which is never marked. The pixel values of an image are I(v0,vi,….vm). They assumed that the database contains float attribute, the number of tuples must be greater than the numbers of image pixels (n>m). The relation R is divided into group of uniform size equal to the size of the image. The algorithm compares a pixel value with an attribute value of a tuple in a relation. Pixel value (0 to 255) is divided into 3 parts. Three types of watermarks are inserted into the relational data. The algorithm is robust against only subset selection attack.

M. Shehab et al [4] proposed a watermarking technique of relational databases which solves the optimization problem based on genetic algorithm and pattern search techniques. It is divided into two parts: Watermark encoding and decoding. Further, watermark encoding is done in three stages: Data set partitioning, watermark embedding and optimal threshold evaluation. Watermark bit(b1,b2,…,bm) is embedded in the m data set partitions (P1,P2,…,Pm). They also offered multiple embedding of bit by specifying length of watermark less than partition (m<n). While, watermark decoding is done in three stages such as data set partitioning, threshold based decoding and majority bit matching. The method is very robust to various attacks such as deletion,

insertion, alteration and synchronization errors. With the use of optimal threshold, it also minimizes the probability of decoding error.

A. Haj et al. [5] proposed for copyright protection of database. A binary image is used to watermark relational databases. The bits of the image are segmented into short binary strings. Each of binary string is transformed into its decimal equivalent, which is encoded in non-numeric, multi-word attributes of selected tuples of the database. The embedment process of each short string is based on creating a double space at a location determined by the decimal equivalent of the short string. Extraction of a short binary string is done by counting number of single spaces between two separated double-space locations. The image watermark is then constructed by converting the decimals into equivalent binary strings. The main advantage of using the space-based watermarking method is the large amount of bit capacity available for hiding the watermark. This facilitates embedment of large capacity watermarks or multiple small watermarks. The algorithm is robust as it resists attempts to remove or degrade the embedded watermark and it is blind as it does not require the original database in order to extract the embedded watermark.

B. Mehta et al. [6] proposed technique to generate robust and impersistant watermark for database. They used image as watermark. Image watermark is embedded over the database at two different attribute of tuple, one in the numeric attribute and another in the date attribute's time (seconds) field. The technique can be applied for numerical and categorical database. The technique is based on the modification of two algorithms given by R. Agrawal et al. [1] And A. Haj et al. [5]. That they are embedding the same watermark in two attributes. Therefore, it will be difficult for attacker to remove both watermarks from the database, based on extracted bits from both algorithms they can generate original watermark very easily, and can prove ownership of database. The technique is costly to implement as they have select embedment of two attributes but robust against common database attacks.

Comparative analysis of above listed database watermarking techniques is given in Table 1. Database attribute as a cover, watermark type, Granularity level, security of watermarking technique, watermarking application, blindness of the techniques and robustness against common database processing attacks are mentioned.

V. CONCLUSION

This paper proposed background and characteristics of relational database watermarking with the flow of watermark insertion and extraction from and to the database. It mainly focused on comparative analysis of different existing database watermarking techniques provided by different authors. Technique provided by B. Mehta[6] is more robust to variety of attacks as compare to other techniques. Further the security level of algorithm is very high as it uses two attributes to watermark.

REFERENCES

[1] R. Agarwal and J Kieman, "Watermarking relational databases" In Proceedings of 28th International In Proceedings of 28th International Conference on very large databases, Hong Kong, China, 2002.Table 1. Comparative Analysis of Different Database Watermarking Techniques

Comparative analysis of existing database watermarking techniques						
Existing Techniques	R. Agarwal[1]	R. Sion [2]	Z. Zhang [3]	M. Shaheb[4]	A. Haj [5]	B. Mehta [6]
Database Attribute as Cover	Numerical	Categorial	Numerical	Numerical	Categorial, Multi Words	Numerical and Categorial (Date)
Watermark Data	Pattern of Bits	Binary	Image	Pattern of Bits	Image	Image
Level of Hiding Details	Bit	Bit	Attribute	Attribute	Attribute	Multiple Attribute
Provide Security	Average	Good	Average	Good	Good	High
Application	Ownership Proof	Ownership Proof	Ownership Proof, Temper Detection	Ownership Proof, Temper Detection	Copyright Protection	Copyright Protection
Unambiguity	Poor	High	Average	High	High	High
Blind	Yes	Yes	Yes	Yes	Yes	Yes
Robust to Insertion Attack	No	No	No	Yes	Yes	Yes
Robust to Alteration Attack	No	Yes	No	Yes	Yes	Yes
Robust to Deletion Attack	No	Yes	No	Yes	Yes	Yes
Robust to Mix and Match Attack	Yes	No	No	No	No	Yes

[2] R. Sion, M. Atallah, and S. Prabhakar, “Rights Protection for Relational Data”, IEEE Transactions on Knowledge and Data Engineering, 16(6), June 2004.

[3] Z. Zhang , X. Jin and J. Wan , “Watermarking Relational Database Using Image”, IEEE Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004.

[4] M. Shehab, E. Bertino and A. Ghafoor, “Watermarking Relational Databases Using Optimization-Based Techniques”, IEEE Transaction on Knowledge and Data engineering, VOL. 20, NO. 1, JANUARY 2008.

[5] A. Haj and A. Odeh, “Robust and Blind Watermarking of Relational Database System,” Journal of Computer Science 4 (12), pp. 1024-1029, 2008.

[6] B. Mehta and U. Rao, “A novel approach as multi-place watermarking for security in databas", in International Conference on Security and Management, pp. 703-707, 2011.

[7] D. Pande, M. Upadhyay, and S. Pal, “Watermarking of Relational Databases Using Optimization Technique”, International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 4, April 2015

[8] G. Agila and N. Ananthanarayanan, “Watermarking Technique using UID for Relational Data Saving in Database”, International Journal of Computer Applications (0975 – 8887) Volume 132 – No.2, December 2015