International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 4 578 - 583

Watermarking for Security in Database

Prof. Manoj Dhande
Department of Computer Engineering of Shah and Anchor Kutchhi Engineering College, Chembur, University of Mumbai, India
manoj.dhande@gmail.com

Aishwarya Kotyankar
Department of Computer Engineering Shah and Anchor Kutchhi Engineering College, Chembur Mumbai, India
aishwaryakotyankar.2016@gmail.com

Nidhi Mannadiar
Department of Computer Engineering Shah and Anchor Kutchhi Engineering College, Chembur Mumbai, India
nidhimannadiar.2016@gmail.com

Prajkta Jagtap
Department of Computer Engineering
Shah and Anchor Kutchhi Engineering College, Chembur
Mumbai, India
prajktajagtap.2016@gmail.com

Abstract: Watermarking technology is used to embed copyright information in objects such as images, audio, video and databases. The increasing use of relational database systems in many real-life applications creates the need for database watermarking systems for protection of database. As a result, watermarking relational database system deals with the legal issue of copyright protection of database system. There are different types of databases like, Numerical and Categorical Databases. Working with numerical data is easier as compared to categorical databases which is much harder to work with. This report addresses a unique, robust copyright protection scheme for Relational Database. Watermark (Characteristic code) is a binary string calculated through the characteristic operation on the original database. A watermark is called robust if it resists a designated class of transformations. Robust watermarks may be used in copy protection applications to carry copy and access control information. The algorithm correlates characteristics according to the content of the databases, which can resist invertibility attack efficiently.Invertibility attack on database is being considered in this paper.

Keywords: *Digital Watermarking, Authentication, Relational Databases, Invertibility Attack, Relating Characteristic.*

*****

I. INTRODUCTION

Digital multimedia watermarking embeds the copyright information in digital objects like images, audio and video.

Digital watermarking embeds information into a digital signal in a way that is difficult to remove. Where, the signal can be an audio, pictures or video. If the signal is copied, then the information is also carried in the copy. A signal may carry several different watermarks at the same time.

A. Types of Watermarking:

1. Visible Watermarking

In visible watermarking, the information is visible in the picture or video. The information is text or a logo which identifies the owner of the media. When a television broadcaster adds its logo to the comer of transmitted video, this is also a visible watermark.

2. Invisible Watermarking

In invisible watermarking, information is added as digital data to audio, picture or video, but it cannot be perceived as such (although it may be possible to detect that some amount of information is hidden). Database watermarking comes under this category.

3. Dual Watermarking:

It is a combination of both visible and invisible watermarking.

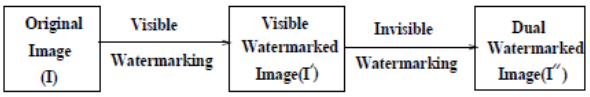

Figure 1. Dual Watermarking Process

Applications:

- Source Tracking.
- Broadcast Monitoring.
- Content protection for audio and video content.
- Forensics and piracy deterrence.
- Communication of ownership and copyright.
- Document and image security.

B .Classification Model of Database Watermarking:

1. Data Type: Different schemes are designed for watermarking different types of data, including numerical data and categorical data.
2. Distortion to underlying data: While some watermarking schemes inevitably introduce distortions/errors to the underlying data, others are distortion-free.
3. Sensitivity to database attacks: A watermarking scheme can be either robust or fragile to database attacks. A scheme is robust (fragile, respectively) if it is difficult to make an embedded watermark undetectable (unchanged, respectively) in database

attacks, provided that the attacks do not degrade the usefulness of the data significantly.

4. Watermark information: The watermark information that is embedded into a database can be a single-bit watermark, a multiple-bit watermark, a fingerprint, or multiple watermarks in different watermarking schemes.
5. Verifiability: A watermark solution is said to be private if the detection of a watermark can only be performed bysomeone who owns a secret key and can only be proven once to the public (e.g., to the court). After this one-time proof, the secret key is known to the public and the embedded watermark can be easily destroyed by malicious users. A watermark solution is said to be public if the detection of a watermark can be publicly proven by anyone, as many times as necessary.
6. Data structure: Different watermarking schemes are designed to accommodate different structural information of the underlying data, including relational databases (with or without primary keys), data cubes, streaming data, and XML data.

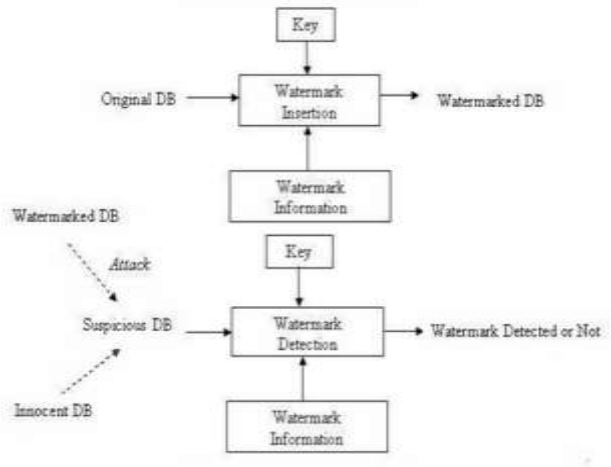

Figure 2. Watermarking Process

C.PROPOSED ALGORITHM PROCESS

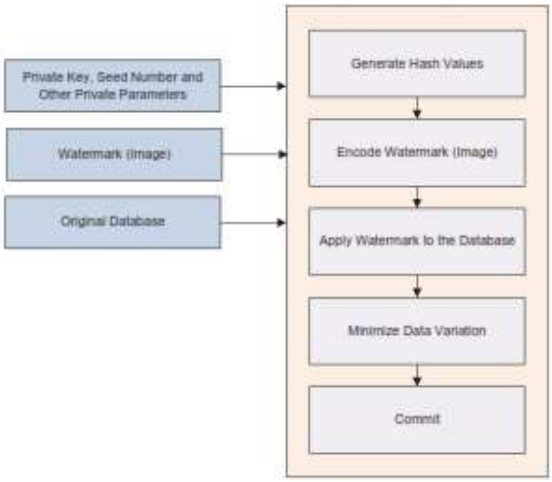

Figure 3. Detailed Process [8]

1. Generate Hash Value:

A database is embedded with a watermark using a private secret key. A new virtual primary key is generated using a cryptographic secure one-way hash function, which consists of with a secret key (KEY) in conjunction with the database table primary key (pk). Hashing algorithm-MD-5 is used.

2. Encode Watermark:

To embed mark bits. All bits will be XORed with securely selected MSB (Most Significant Bit) of selected attribute among changeable candidate attributes, according to generated key. This approach will increase mark imperceptibility.

3. To apply watermark to the Database:

Two dimensions are used to refer the image directly, and does not map the image from two dimensions to a vector. A hash function is used to generate values for a tuple. We then use the generated values and use it for elements coordinates and the value of pixel in this coordinates will be embedded into the relevant tuple.

4. Minimize Variation of Data:

Altering the value of attributes is the most important point in database during watermarking and should be noted. The vital issue is to limit modifications in the tuples, to an acceptable range.

5. Selection of LSB:

In this method one of the specified LSBs will be selected and modified.The attribute will be changed as well. If thefirst LSB is the selected bit, the value will be changed in minimum rate, but if the chosen bit is the second or a higher bit, the value will be changed by more than one unit.

6. Commit:

All modifications will be saved into the database by issuing Commit command, provided all of the above steps are performed without any error.

D. Analysis of Invertibility Attack Resistance

An attacker may take following steps to find the fictitious watermark.

(1) Choose the parameters randomly and use them to do extraction operation on the watermarked database $R\_$;

(2) From each tuple an index number can be computed (index (ri) =rand (K, $ri.P$)). And in this way he can choose

(3) Calculate the watermarked attributes and the LSBs (attribute index x=index (ri) mod υ, bit index j= index (ri) mod ε);

(4) At last, extract the watermark information from the effective bits and obtain the fictitious watermark $WM\_$.

Then he can forge the original database according to the fictitious watermark *WM_* by using the same parameters to fictitious watermark.

Once a forged original database $R*$is obtained, the attacker may declare his copyright of database *R_* by providing forged original database $R*$, watermark *WM_* and watermarked database *R_*.

To confirm the copyright, a legitimate owner can providethe original database *R*, watermark *WM*, watermarkeddatabase *R_*. According to the proposed algorithm, characteristic code for *R* and $R*$can be computed respectively:

Table 3: Values of Similarity 2 after Invertibility Attacks

Parameters	ρ_2
Different K	0.45
Different γ	0.55
Different ε	0.64
Different υ	0.53
Different K and γ	0.47
All different	0.49

All different 0.49 restore the watermarked database. The original bit before embedding is either 1 or 0. He can randomly change the data at the calculated location with 0 or 1, according to the $O(R) = WM$", $O(R*) = WM*$
Calculate the similarity $\rho 1$ and $\rho 2$:

$\rho 1 = sim\ (WM, WM$"$)$, $\rho 2 = sim\ (WM\_, WM*)$
Since the fictitious watermark *WM_* is randomly found,
The similarity $\rho 2$ is relatively low while the similarity $\rho 1$ is higher.

In the experiment, an invertibility attack was simulated.
The similarity $\rho 1$ and $\rho 2$ were calculated respectively. Inthe experiment, $\rho 1 = 0.99$, greater than the given threshold α. We also calculated the values of $\rho 2$ according to different values of parameters, shown as Table 3.

Table 3 shows that if one of the four parameters K, γ, υ, ε selected by the attacker is different with the legitimateowner, the similarity will be much less than the giventhreshold in watermark detection. The probability of selecting all the parameters same with the legitimate owner's by an attacker is shown as formula (5):
$P = 1$
$K \times \gamma \times \varepsilon \times \upsilon$

(5)Assuming $K \in [1, 1000]$, $\gamma \in [1,1000]$, and $\varepsilon \in [1,4]$, there are 50 numerical attributes to be embedded watermark. The probability of selecting all the parameters which are the same with the legitimate owner's by an attacker is shown as formula (6):
$P = 1$
$1000 \times 1000 \times 4 \times 50$
$= 0.5 \times 10{-}8$ (6)
So it can be concluded that it is impossible for an attacker to select all the parameters same with those of the legitimate owner.
At the same time, we can conclude from Table 3 that different parameters have different effect on watermark detection. The effect of ε for watermark detection is stronger.

III.PROPOSED SYSTEM

To identify the tuples in the relational database and embedding a watermark in them in spatial domain. This watermarked relational database should not be vulnerable to attacks like invertibility attack. The end result will be a relational database with copyright protection that cannot be tampered with or modified without the author 's permission or access rights. The used method will insert text and/or image watermark, detect the important tuples that need to be watermarked so that the entire database gets protected against unauthorized access. We also perform encryption and decryption of the database to enhance security. We also provide a query tuner to optimize the queries. We perform fragmentation and decomposition where we propose a futuristic view to watermark the fragmented and decomposed database.

1.The Query tuner optimizes a query by prompting the appropriate solution.

Figure 1: Query Prompt

2.The database is encrypted using a key and the same key is used to decrypt the database

2.1 Data Encryption Standard or DES Algorithm:

1. Firstly, we need to process the key.
2. Process a 64-bit data block.
3. Get a 64-bit data block

 ISSN: 2321-8169
578 - 583

4. Perform the following permutation on the data block.
5. Split the block into two halves. The first 32 bits are called L[0], and the last 32 bits are called R[0].
6. Apply the 16 subkeys to the data block.
7. Expand the 32-bit R[i-1] into 48 bits according to the bit-selection
8. Exclusive-or E(R[i-1]) with K[i].
9. Break E(R[i-1]) xor K[i] into eight 6-bit blocks
10. Substitute the values found in the S-boxes for all B[j].
11. until all 8 blocks have been replaced.
12. Repeat the same for L[i-1]
13. Perform the permutation on the block R[16] L[16].

2.2 Decryption Algorithm :

1. To decrypt, use the same process as encryption with one change.
2. Use the keys K[i] in reverse order.
3. Instead of applying K[1] for the first iteration, apply K[16], and then K[15] for the second, on down to K[1].

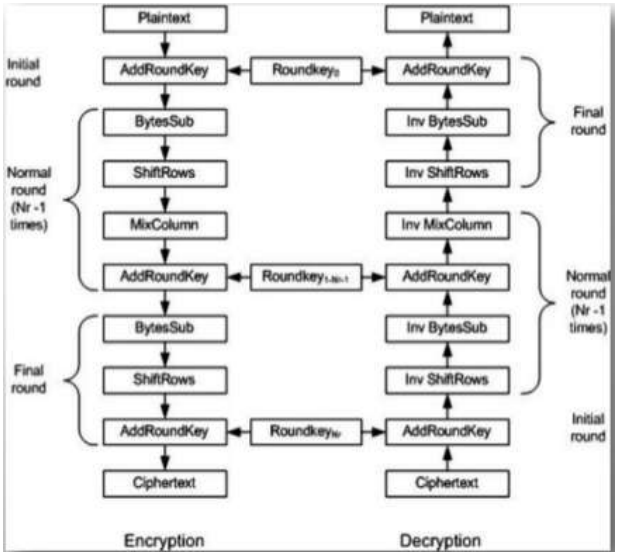

Figure: 2.1 AES encryption and Decryption process

Figure 2.2: Encryption of Database

Figure 2.3: Encrypted Database

Figure 2.4: Decryption of Database

Figure 2.5: Decrypted Database

3.Horizontal and vertical Fragmentation .

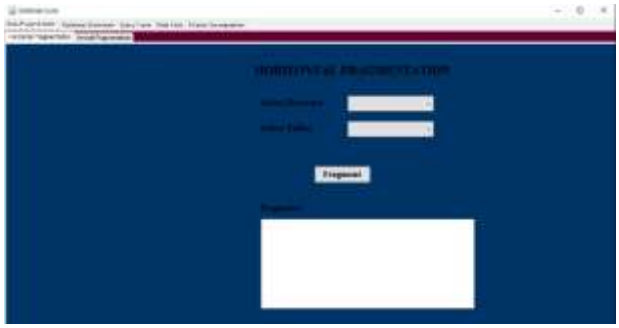

Figure 3.1 Horizontal Fragmentation

Figure 3.2 Vertical Fragmentation

4. Decomposition provides .mdl file which shows cardinality between entities of schema.

Figure 4 : Schema Decomposition

5.Watermark.wes file contains the database watermarked by text and watermarking.wes file contains database watermarked by image.

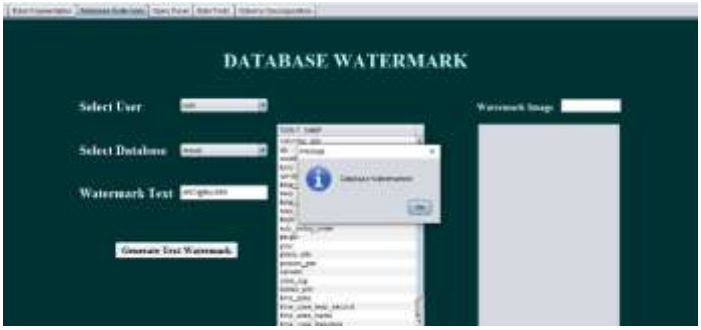

Figure 5.1 : Database Watermarking using text

Figure 5.2 : Database Watermarked with text

Figure 5.3 : Database Watermarking using image

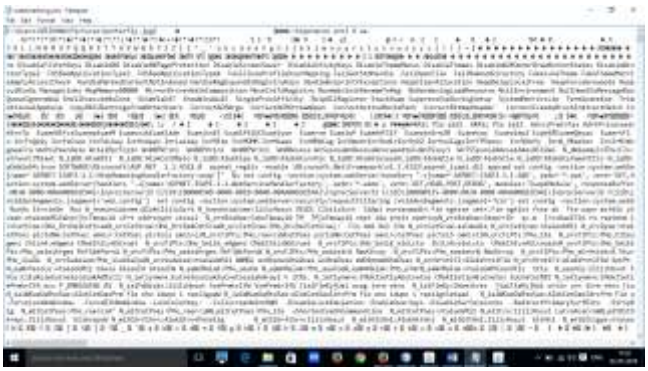

Figure 5.4 : Database Watermarked with image

IV.CONCLUSION:

The algorithm presented in this report is a new, novel and robust watermarking technique for watermarking relational database. The implementation results clearly show that the use of LSB bit replacement technique proves for an efficient way to watermark relational database for providing copyright protection to the database owner. We also encrypt and decrypt the database using DES algorithm. This feature provides for enhanced security. The query tuner provides for optimized querying of the relational database. The fragmentation and decomposition of database provides for a futuristic scope wherein we propose the watermarking process to not be limited to embedding the watermark to the entire database but also should provide for embedding the watermark in the desired fragments of the database as per user convenience and desirability. Results and relative analysis show that database is authenticated by embedding watermark text or image and its security is enhanced by encryption processes .The basic system encompasses of a query tuner, Schema decomposer, a provision to watermark the database using text and image .Database can be encrypted and decrypted in the system. The main advantage of this method is that it is robust and can withstand various attacks. The system therefore provides for an efficient system that confers security to database with authentication

.In the experiments, database is used as the host, whereas text and .jpeg images are used as watermark image.

V. ACKNOWLEDGEMENT

A Paper is a teamwork and reflects the contribution of many people .We would like to thank everyone who has contributed to this effort by sharing their time and taking interest in our work and encouraging us all the way through.In particular, we thank our guide Mr. Manoj Dhande, for helping in our project. We thank him for his guidance, support and words of encouragement throughout the time of our project.

VI.REFERENCES

[1] R. Agrawal, P. J. Hass and J. Kiernan, Watermarking relationaldata: framework, algorithms and analysis, Proceedingsof the 28th VLDB Conference, 157-169, 2003.

[2] Y. Zhang, D. Zhao, and D. Li, Digital watermarking for relational databases, Computer engineering and application, No 25, 193-195, 2003. (In Chinese)

[3] Y. Wang, G. Zhu and S. Zhang "Research on the Watermarking Algorithm based on Numerical Attributein the Relational Database", IEEE, ICCSEE, pp. 363-367, 2012.

[4] M. Atallah and S. Wagstaff. Watermarking with quadratic residues. In *Proc. of IS&T/SPIE Conference on Security and Watermarking of Multimedia Contents*, January 1999.

[5] K. Huang, M. Yue, P. Chen, Y. He and X. Chen "A Cluster-based Watermarking Technique for Relational Database", IEEE, IWDTA, pp. 107-110, 2009.

[6] Atallah M, Wagstaff S (1999) Watermarking with quadrati residues.In: Proceedings of IS&T/SPIE conference on security and watermarking of multimedia contents, January 1999

[7] Y. Li, H. Guo and S. Jajodia, Tamper Detection and Localization for Categorical Data Using Fragile Watermarks. Proceedings of the 4thACM Workshop on Digital Rights Management, Washington DC, USA, pp. 73-82, 2004.

[8] http://www.globalcis.org/jcis/ppl//JCIS1-088017IP-4.pdf.