Secure and Reliable Watermarking in Relational Databases

G.Shyamala[1], I.Jasmine Selvakumari Jeya[2], M.Revathi[3]
[1](PG Student, Dept. of CSE, Hindusthan College of Engineering and Tech, Coimbatore, Tamil Nadu, India)
[2,3](Assistant Professor, Dept. of CSE, Hindusthan College of Engg and Tech, Coimbatore, Tamil Nadu, India)

ABSTRACT : *The rapid growth of the Internet and related technologies has offered an unpredicted ability to access and redistribute original digital contents. Watermarking techniques have emerged as an important building block that plays a crucial role in addressing the ownership problem. A watermark is a owner information that is securely and robustly embedded into original information such as an text, image, video or audio signal, producing a watermark information. In this paper, it provides mechanism for proof of ownership based on the secure embedding of a robust imperceptible watermark in relational data. This technique is resilient to watermark synchronization errors because it uses a partitioning approach that does not require marker tuples. It is resilient to tuple deletion, alteration and insertion attacks. This approach is used in watermarking database to provide security to owner information and to identify the unique characteristics of relational data and provide desirable properties of a watermarking system for relational data. The watermarking technique can be applied to any database relation having attributes in which few tuples altered based on the threshold value do not affect the database values. This approach ensures that some bit positions of some of the attributes of some of the tuples contain watermark bits. The tuples division is done and after finding the threshold value watermark bits are added for efficient and secure watermark embedding and attacker may not find the original watermarked data. Only if user has access to the private key, then he can decode the database with high probability.*

Keywords *- Attacks, database security, ownership protection, secret key, watermarking.*

I. INTRODUCTION

In general, watermarking can be used for different format such as text, audio, video relational databases. The database watermarking techniques consist of two phases: Watermark Embedding and Watermark Verification [10]. In watermark embedding phase, a private key K (known only to the owner) which is a secret key is used to embed the watermark W into the original database. The watermarked database is then made publicly Watermark detection is blinded, that it does not require the original information or original data for the watermark decoding watermark. 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 [1]. Only the owner of the database itself should be able to extract and verify a given watermarked relation. This is accomplished using a secret key in order to determine watermarked tuples, partitions, attributes and the watermark itself.

II. RELATED WORKS

Mohamed Shehab, Elisa Bertino & Arif Ghaffor [9] proposed a watermarking technique, which is constrained optimization problem and efficient techniques to handle the constraints. Genetic algorithm and pattern search techniques are used to solve optimization problem. Hamed khataeimaragheh and Hassan Rashidi [4] proposed a novel watermarking scheme for detecting and recovering distortions in database tables. A hash value is computed over a relational database and then is signed with the owner's private key.

M.Kamaran and Fahim Arif proposed [8] watermarking technique for relational databases with emphasis on re-watermarking attack. It uses bi-level security principle, thus provides robustness against re-watermarking attack. Bi-folded security scheme to detect and resolve conflicting ownership issues in case of re-watermarking attack (also called additive or secondary attacks). Sukriti Bhattacharya and Agostino Cortesi [10] introduce a distortion free watermarking technique that strengthens the verification of integrity of the relational databases by using a public zero distortion authentication mechanism.

Vahab Pournaghshband [11] presents an effective watermarking technique for relational data that is robust against various attacks. While previous techniques have been mainly concerned with introducing errors into the actual data, our approach inserts new tuples that are not real and we call them "fake" tuples, to the relation as

watermarks. Vidhi Khandujaand O. P. Verma [12] proposes a secure robust and imperceptible algorithm. In this paper they proposed a technique to securely and randomly select any number of attributes out of selected candidate attributes for embedding watermarks in varying number of least significant bits.

Abdullah Bamatraf, Rosziati Ibrahim and Mohd. Najib Mohd. Salleh [1] proposed an algorithm is simple, more robust against Salt and Pepper Noise than LSB only watermarking techniques. Yossra H. Ali and Bashar Saadoon Mahdi [14] provides the effective watermarking technique to protect valuable numeric relational data from illegal duplications and redistributions as well as to claim ownership, the robustness of proposed system depending on using new hybrid techniques, first technique MAC (Message Authentication Code) that used one way hash function SHA1, second technique is threshold generator base on simple combination of odd number of register and by using secret key in proposed system. Also Sion et al. introduce a watermark technique for numerical data. This technique is dependent on a secret key, instead of primary key uses the most significant bits of the normalized data set, divides the data set into partitions using markers, and varies the partition statistics to hide watermark bits.

Hazem M. El-Bakry [5] proposed technique is available for any relational database. No delay and no additional time required till the normal calculation end. R.Balasubramaniam [2] present watermarking embeds ownership information in digital content Watermarking of relational databases as a constrained optimization problem and discus efficient techniques to solve the optimization problem and to handle the constraints. Jiri Fridrich and Miroslav Goljan [6] proposed robust watermarking for relational databases. In this paper, we discuss methods how such robust hash functions can be built. We describe an algorithm and evaluate its performance.

S.W. Weng, Y. Zhao and J.-S. Pa [13] proposed a reversible digital watermarking scheme based on a block-wise Difference Expansion (DE) method is proposed here. The proposed scheme differs from previous ones in its robustness against cropping and collage attacks. Goljan et al [5] proposed a two cycles flipping permutation to assign a watermarking bit in each pixel group. Celik et al. presented a high capacity, reversible data-embedding algorithm with low distortion by compressing quantization residues. The embedding process is locating patterned embedding.

III. PROPOSED SYSTEM

Watermarking embeds the ownership information in the form of digital content and this technique used to hide a small amount of digital data in a it can't be detected by a the attacker or hackers. Watermark describes information that can be used to prove the ownership of relational database. Multiple embedding of watermark bits in the dataset increases additional security. Fig.1 shows the stages of watermark encoding and decoding.

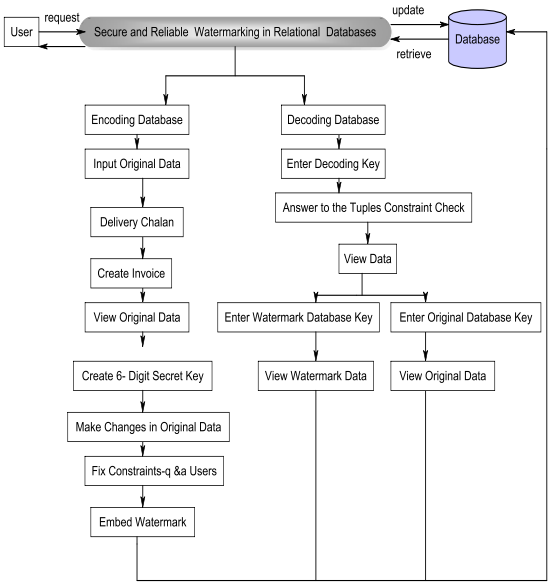

Fig 1.Stages of Watermark Encoding and Decoding

The original database has a 6-digit original database key and watermark key to embed the watermark in the database. In decoding phase the user has a secret login password which is generated based on the date-time stamp and after the authenticated login, user needs to answer to tight constraint questions which are fixed by the owner of the database. Only after the successful login the user can view the original database and the watermarked tuples. This technique is resilient against different type of attacks and more secure. The advantages of the proposed system are:

- The proposed system provides less distortion.
- The system provide 100 accuracy for decoding
- The system also helps to resolve ownership conflicts over watermarked data set in case of additive attacks

- Security is high because the watermarking depends on the secret key.

3.1 Encoding Phase-Input Original Data

In Encoding phase the first step is only an authorized owner can enters the encoding phase. First the owner enters the stock information in the Input original data of Stock database module admin will insert details about the company such as Invoice number Docket number, number of items purchased, amount of items sold from the companies. Each and every changes of the stock details should updated by admin. This module is completely handled by the admin side.

After this process the owner needs to embed the database using a 6-digit secret key which can be used only once. The owner creates a New DC (Data Count) and the lists of items are added and the DC no. is saved. The owner can search for a particular item, edit, Update DC and all the information are recorded in the Original database. A new Invoice number is created for each DC and it is unique id.

3.2 View Original Data- Tuples Division

In this module, after entering the stock details that information are updated in the original database. This is very important that tuples are divided into partition and the owner enters a 6-Digit original database key and the watermark key. The original database tuples are divided into cells and the owner enters the watermark bits.

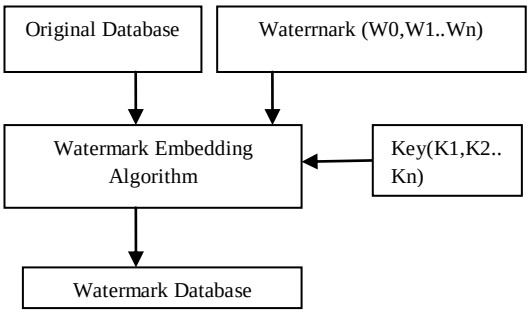

Fig 2. The Process of Watermark Embedding

3.3 Fix Constraints

The owners fix tight constraints for the users by adding secret questions and the destination mail-id. For eg: the Owner assigns two questions as shown below:

1. What is your first Vehicle number? 74****
2. What is your favorite color? Red

These questions and answers are entered by the owner and only answering these questions only user can enter the database. These fields are said to be case-sensitive by three wrong addresses, it will block the IP address of the users.

3.4 Embed Watermark

This module plays a main role of this project which uses the Algorithm 1.

Input: Original database D
Output: Watermarked database Dw
1. Retrieve O_k ← Original Data
2. Embed Watermark cells dynamically, ω ← Original Data
3. for i=0 to ColumnCount-1
4. for j=0 to RowCount-1
5. Watermark Data←BindCells(i,j, ω)
6. Next j
7. Next i
8. Enter 6-Digit ω_{key}
9. Input Watermark Constraints
10. Answer1 ←Question1
11. Answer2 ←Question2
12. Input Destination Mail Id, D_{mail}

Algorithm 1. Embed Watermarking

The watermark technique is applied to the input. After apply the watermark technique the data will be hided. The hided data we will send through network with watermark key. As the final step the watermarked data key and the original database key will send to destination the mail through the network. In this module after adding watermark bits the owner embeds the watermark database and decoding secret key and sent to the user.

3.5 Decoding Database

Decoding the database is an important module. In decoding phase the user can enter only by entering the secret key which is generated by using the Algorithm 2 and this secret key is sent to the user's mail id.

1. Assign S1,S2,S3,Secret Key S_{Key}
2. S1← UT Seconds
3. S2← UT Minutes
4. S3← UT Milliseconds
5. Secret Key S_{Key} ← S1&S2&S3
6. return

Algorithm 2. Secret Key Generator

After the successful login of the secret key the user has to answer for all the tight constraint questions which are generated by the admin and it

is case-sensitive. After answering for all the three questions only, user can enter the decoding phase and any one or more wrong answers will not allow the end users to enter the decoding phase. Only by entering the Decoding database key the user can enter the decoding phase. If hackers want to access our data in-between the network, our application will check the access permission and the key. Even if they try to change the watermark information it will not affect the end users, since they have the comparison of both original and watermark database. By three wrong secret key input the user ip address get blocked. It is more secure against attackers. In this module only by entering the decoding secret key, which is generated based on date-time stamp based on Algorithm 2, the user can enter the decoding phase and user needs to answer the tight constraints questions and answers. Two keys are sent to user one is the watermark key and other one is the database key. Algorithm 3 explain that after entering correct decoding secret key, user enters the decoding phase and needs watermark key and Database key to view the watermarked data.

Input: Decoding secret key S_{key}
Output: Watermarked dataset
1.Enter Secret key S_{Key}.If Succeed goto (2)
2.Enter Answer1 ←Question1
3.If Succed (2) then goto (3)
4.else repeat (2)
5. Enter Answer2 ←Question2
6.If Succed (3) then goto (4)
7. else repeat (3)
8. Enter O_k to decode Original Tuples
9. Enter ω_{key} to view Watermarked Data.
10. Count attacker channel $C_{ac}\varepsilon\omega_{tuples}$
11. Return

Algorithm 3. Decoding Secret Key

3.6 ORIGINAL AND WATERMARK DECODING

User enters the Watermark key to view the watermark tuples and original database key to view the original data. The watermarked tuples are marked in red to indicate the watermarking and visible only in the watermark database.

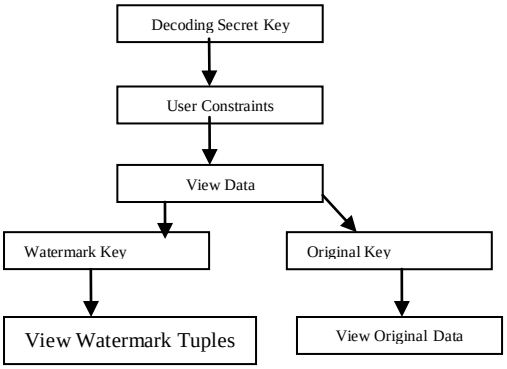

Fig 3. The Process of Watermark Decoding

IV. EXPERIMENTAL RESULTS

The experiments were performed in Windows 7, i5 Processor with Microsoft .net 4.0 as front end and MS-Access has back end. Fig IV.1 shows the input form of a Stock database where new dc and new invoice are created and added into the database. Fig IV.2 shows the process of adding original database key to the database.

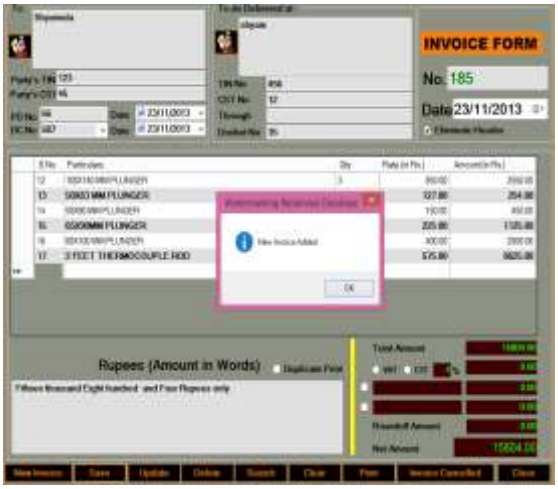

Fig IV.1 Shows the process of adding data to the Stock database

Fig IV.2 Shows the Process of Adding Database key to Original database

Fig IV.3 explains the process of tuple division for watermark insertion and fix tight constraints to the user. Fig IV.4 shows only the user who has access to the decoding secret key can enter into the decoding phase. The secret key is generated based on time and it is valid only for few minutes. For the next time the user needs to get new secret key from the owner. Fig IV.5 explains that only by answering to the tight constraints, the user can view the watermarked database. The IP

address gets blocked by itself with a maximum of nine improper attempts.

Fig IV.3 Owner Enters 6-Digit Key for Original Database

Fig IV.4 Shows the Process of Watermark embedding and fix constraints.

Fig IV.5 Shows that user entering the decoding secret key to enter decode phase

Fig IV.6 shows the watermarked database marked in red. The user needs the watermark key which is sent by the owner to the user mail-id. Only by giving that key the user can view the watermarked database. This paper provides high security to the owner's database by providing three key such as Secret Key for secret login of user decoding database, Watermark key and Original database key.

Fig IV.6 Shows the user answering to the tight constraints.

Fig IV.7 Shows the Watermarked Database Marked in Red.

V. CONCLUSION

The Watermarking technique results "maximum possible robustness" to the owners and "minimum distortion" in the original dataset to the user. It provides both authentication and integrity. It is based on partition tuple division and able to detect and locate modifications as so that it is easy to trace group which is possibly affected when a tuple is tampered. Neither watermark generation nor detection depends on any correlation or costly sorting among data items. Each tuple in the table is independently processed and therefore it is efficient for tuple oriented database operations.

REFERENCES

[1] Abdullah Bamatraf, Rosziati Ibrahim and Mohd. Najib Mohd. Salleh, "A New Digital Watermarking Algorithm Using Combination of Least Significant Bit (LSB) and Inverse Bit". *Journal Of Computing*, *Issn*, Volume 3, Issue 4, April 2011, 2151-9617.

[2] R.Balasubramaniam, "Data Security In Relational Database Management System", *International Journal of Computer Science and Security (IJCSS*), Volume (6): Issue (4): 2012.
[3] S. Bhattacharya and A. Cortesi, "A distortion free watermark framework for relational databases," *Proceedings of the 4th International Conference on Software and Data Technologies (ICSOFT 2009)*, pp. 229–234, 2009.
[4] Hamed khataeimaragheh and Hassan Rashidi, "A Novel Watermarking Scheme For Detecting And Recovering Distortions In Database Tables" *International Journal of Database Management Systems (IJDMS)* Vol.2, No.3, August 2010.
[5] Hazem M. El-Bakry, "A New Watermark Approach for Protection of Databases", Mansoura University, *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 116–129, 2007.
[6] Jiri Fridrich and Goljan, "Robust Hash Functions for Digital Watermarking", *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 3, No 1, May 2012.
[7] M. Kamran, Sabah Suhail, and Muddassar Farooq, "A Robust, Distortion Minimizing Technique for Watermarking Relational Databases Using Once-for-all Usability Constraints", *IEEE Transactions On Knowledge And Data Engineering* Digital Object Indentifier 10.1109.
[8] Sabah Suhail1, M. Kamran and Fahim Arif, "Watermarking of Relational Databases with Emphasis on Re-watermarking Attack", *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 3, No 1, May 2012.
[9] M. Shehab, E. Bertino, and A. Ghafoor, "Watermarking relational databases using optimization-based techniques," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 116–129, 2008.
[10]Shantanu Pal, Raju Halder, Agostino Cortesi "Watermarking Techniques for Relational Databases Survey, Classification and Comparison" Journal of Universal Computer Science, vol. 16, no. 21 (2010), 3164-3190 submitted: 18/12/09, accepted: 29/11/10, appeared: 1/12/10 © J.UCS
[11] Vahab Pournaghshband , "A New Watermarking Approach for Relational Data" , 9th WSEAS International Conference on Applied Informatics And Communications (AIC '09) Vol.2, No.3, August 2010.
[12] Vidhi Khanduja, and O. P. Verma, "Identification and Proof of Ownership by Watermarking Relational Databases" International Journal of Information and Electronics Engineering, Vol. 2, No. 2, March 2012.
[13] S.W. Weng, Y. Zhao and J.-S. Pan "Reversible watermarking resistant to cropping attack", The Institution of Engineering and Technology 2007 (IET Inf. Secure), 2007, 1, (2), pp. 91–95.
[14] Dr. Yossra H. Ali & Bashar Saadoon Mahdi , "Watermarking for Relational Database by using Threshold Generator", Computer Sciences Department, *University of Technology Eng. & Tech*. Journal, Vol. 29, No. 1, 2011.