IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 08, Issue 1, January 2018, ||V-3|| PP 40-44

Reversible Watermarking Technique Based On a Time- Stamping In Relational Data

[1]Dr. A.N. Nanda Kumar, [2]Mr. S. Niresh Kumar, [3]Mrs. M. Revathi,
[1]*Dhanalakshmi Srinivasan College of Engineering and Technology*
[2]*Dhanalakshmi Srinivasan College of Engineering and Technology*
[2]*Asan Memorial College of Engineering and Technology*

ABSTRACT: *Watermarking method is to recognizable pattern accustomed identify authenticity. Intentionally introduced pattern within the data is difficult to hunt out and destroy, robust against malicious attack. Watermarking, with none exception, has been used for ownership protection of style of knowledge formats that are utilized in several application domains. The watermark embedding algorithm takes a secret key (Ks) and thus the watermark bits (W) as input and converts a knowledge set D into watermarked data set DW. A cryptographic hash function MD5 is applied on the chosen data set to decide on out only those tuples which have an honest hash value. The proposed algorithm embeds as of a multi bit watermark (generated from date-time) in each selected row (in a numeric attribute) which generates the UTC Date_Time key for authentication with the target of getting maximum robustness whether or not an attacker is somehow able to successfully corrupt the watermark in some selected a component of the data set. The watermark embedding has been tired numeric database instead of alpha numeric database. Security of the database are improved.*

I. INTRODUCTION

AIM

The main aim of this project is to maintain the ownership of Relational Database and also minimizing distortion in the watermarked content.

DATA MINING:

Data mining is the process of extracting patterns from data.

As more data are gathered, with the amount of data doubling every three years, data mining is becoming an increasingly important tool to transform these data into information.

One of the common algorithms is Clustering. It is like classification but the groups are not predefined, so the algorithm will try to group similar items together. Regression is an attempt to find a function which models the data with the least error. A common method is to use Genetic Programming

WATERMARKING:

Watermarking method is used to maintain ownership in Relational database without any exception. It has been used for ownership protection of a number of data formats such as images, video, audio, software, XML documents, geographic information system (GIS) related data, text documents, relational databases and so on which are used in different application domains.

There are two types of watermarking

1. Visible Watermarking
2. Invisible Watermarking

The owner of the Relational Database embeds the watermark data, the distortions in the original data1 are kept within certain limits, which are defined by the usability constraints, to preserve the knowledge contained in the data.

The proposed algorithm embeds every bit of a multibit watermark. Watermark Embedding Algorithm embeds watermark with the selected tuples. The tuples are selected based on the threshold value and the MD5 Hash Function. The threshold value is compared with the original data and if it has an even hash value then the tuple is selected for watermarking.

The selected tuple is converted into binary values and with the UTC key, the password should be given. The grey scale password will be generated. Then the grey scale image should be chosen. Finally, the selected data will be watermarked. To see the original data, again the user should enter the password. Then the original data will be displayed.

The data should be created. The created data should be viewed. Then the tuple selection process should be done. The tuple will be selected randomly. By using the threshold formula, the values will be computed. This

method is used to avoid data loss during data hide. After watermarked, the decode process should be done, Admin can view the original data without any authentication.

II. SYSTEM ANALYSIS

EXISTING SYSTEM

A bit-resetting algorithm that employs the principle of setting the least significant bit (LSB) of the candidate attribute of the selected subset of tuples.

In Existing System MAC is used for Hash Function. The parameters selection for watermarking is based on computing message authentication code (MAC), where MAC is calculated using the secret key and the tuple's primary key. This technique assumes unconstrained LSB manipulation during watermark embedding process.

Although LSB-based data hiding techniques are efficient, but an attacker is able to easily remove watermark by simple manipulation of data by shifting LSB.

The data partitioning concept is based on the use of special marker tuples, making it vulnerable to watermark synchronization errors.

DISADVANTAGES

Watermarking on alpha numeric database are not robust against malicious attacks.
Attacker is able to easily remove watermark by simple manipulation of data.

PROPOSED SYSTEM

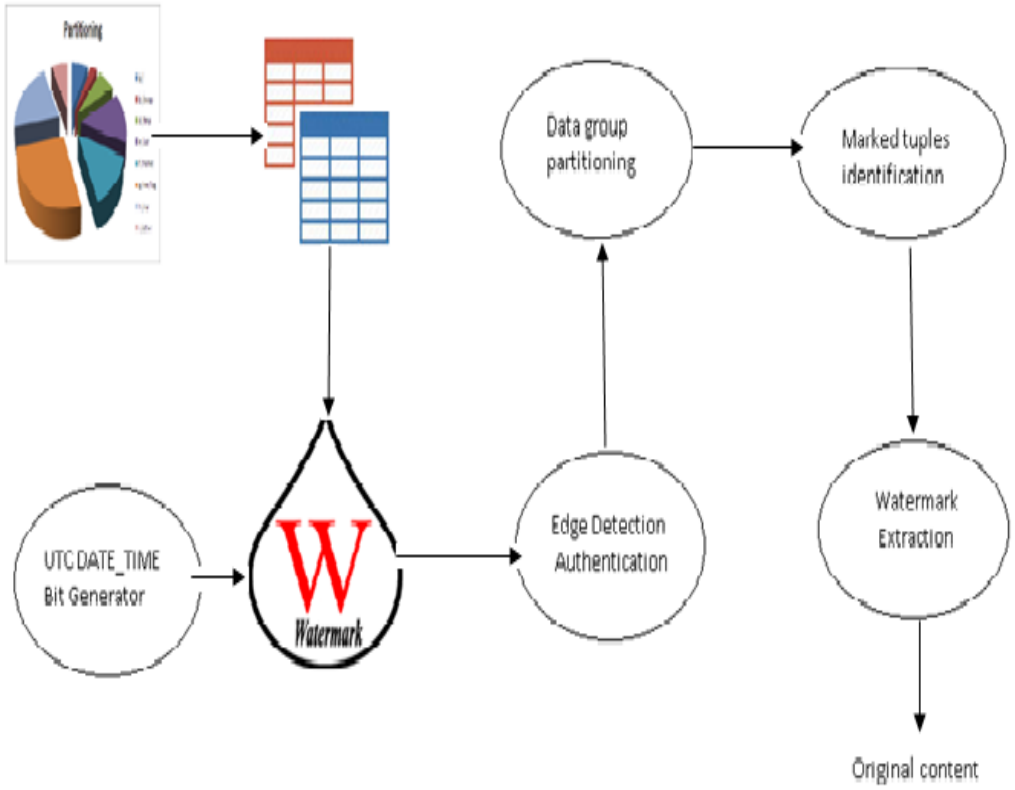
This Proposed system we implement a new approach to generate the watermark bits from UTC (Coordinated Universal Time) date time which is the primary time standard used to synchronize the time all over the world. A robust watermark algorithm is used to embed watermark bits into the data set of Database Owner. The watermark embedding algorithm takes a secret key (K_s) and the watermark bits (W) as input and converts a data set D into watermarked data set DW. A cryptographic hash function MD5 is applied on the selected data set to select only those tuples which have an even hash value. The Watermarking process includes Encoding and Decoding Phase. The Encoding phase consist of Data partitioning, Selection of data set for watermarking, Watermark embedding process, Decoding phase consist also these process to extract the Watermarked content.

ADVANTAGES

The new approach to generate the watermark bits from UTC.
Random mode of selection of dataset for watermarking.
.

III. SYSTEM DESIGN

The tuples in the database are partitioned and the tuples are selected to watermark in the tuple selection process. The selected tuples are embedded with the watermark bits. The UTC Date_Time bit generator generates the UTC key which is used as a authentication key for the user. In Edge Detection and Authentication process the authentication of the user is verified. The tuples in the partitioned data group is identified. In the final stage the embedded watermark bits are extracted from the original data and finally the original content is obtained.



IV. MODULES

1. Data Group Partitioning
2. Tuples Selection for Watermarking
3. Watermark Embedding
4. Edge detection Authentication and Watermark Extraction

1. Data Group Partitioning

In this module includes the Data partitioning Relational Numerical Database Watermarking. Data Partitioning comes under Watermark Encoding Phase which has been done by owner of the DataBase (ie) Admin. The data partitioning algorithm partitions the data set into logical groups by using data partitioning algorithm.

$$\mathbf{par(r)=H(k_s||H(r.P_k||k_s))mod\ m}$$

where $r.P_K$ is the primary key of the tuple r,

H() is a cryptographic hash function Message Digest (MD5),

|| is the concatenation,

k_s is a secret key.

Logical groups or Partitions has been arrived after applied this algorithm.Admin has to decided the groups length that is m.

2. Tuples Selection for Watermarking

A Tuple is one record or one row in a Relational Database. In this phas to Select the Particular tuples For embedding Watermarked Content.Thresold Computation is a mehod computed for each attribute. If the value of any attribute of a tuple is above its respective computed threshold, it is selected for Encoding Process. The data selection threshold for an attribute is calculated by using the following equation:

T=c * Mean+ Standard Deviation

Here, c is the confidence factor with a value between 0 and 1.

The confidence factor c is kept secret to make it very difficult for an attacker to guess the selected tuples in which the watermark is inserted. We select only those tuples, during the encoding process, whose values are above T. Collect Selected tuples for Encoding and apply Hash Value Computation.

In this step, a cryptographic hash function MD5 is applied on the selected data set to select only those tuples which have an even hash value. This step achieves two objectives:

1) it further enhances the watermark security by hiding the identity of the watermarked tuples from an intruder;

2) it further reduces the number of to-be-watermarked tuples to limit distortions in the data set .If the Hash Value Computation
Is Satisfied Select the tuples for Watermarking bits from Selected tuples for Encoding process.

3. Watermark Embedding

The watermark generating function takes date-time stamp as an input and then generates watermark bits b1b2 . . . bn from this date-time stamp. These bits are given as input to the watermark encoding function .The date-time stamp might also help to identify additive attacks in which an attacker wants to rewatermark the data set. To construct a watermarked data set, these watermark bits are embedded in the original data set by using watermark embedding algorithm. The proposed algorithm embeds every bit of a multibit watermark generated from date-time in each selected row. The watermark bits are embedded in the selected tuples using a robust watermarking function. Our technique embeds each bit of the watermark in every selected tuple of each partition.

4. Edge detection Authentication and Watermark Extraction

Edge detection Authentication is proposed as an alternative solution to text based. It is mainly depends on images rather than alphanumerical. The main argument here is that pass-images from the challenge set and then he/she will be authenticated users are better at recognizing and memorizing pictures. During Registration phase Admin has to provide some images to the user. In the registration phase the user is supposed to choose the pass-images for the verification phase. That image has to be Stored in Server For that Specific User.

During Login phase Admin has to converting the raw image to a gray scale followed by Edge detection image. The idea here is the user will have a challenge set which contains decoy and pass-images. The decoy images are randomly generated by the scheme during the verification process. On the other hand, pass-image will be the users selected images. Basically authentication is simple; a legitimate user needs to correctly identify pass-images from the challenge set and then he/she will be authenticated.

Watermark Extraction process in the Decoding phase. The Watermarked content has to be extracted only by legitimate user to give the proper ownership. If the user ownership content is matched by the admin generated content , Decoding process has done. Otherwise it is not done.

V. CONCLUSION

Reversible watermarking techniques are used to cater to such scenarios because they are able to recover original data from watermarked data and ensure data quality to some extent. The main contribution of this work is that it allows recovery of a large portion of the data even after being subjected to malicious attacks. Reversible watermarking technique is compared with recently proposed state-of-the-art techniques such as DEW, GADEW and PEEW to demonstrate that reversible watermarking outperforms all of them on different performance merits.

FUTURE ENHANCEMENT

Applying watermark to alpha numeric database. Hiding in alpha numeric database is applicable. We also plan to extend reversible watermarking for non-numeric data stores.

REFERENCES

[1]. Saman Iftikhar,M.Kamaran and Zahid Anwar,June 2014"RRW- A Robust and Reversible Watermarking Technique for Relational Data.
[2]. R. Agrawal and J. Kiernan,Dec 2002,"Watermarking relational databases," in Proceedings of the 28th international conference on Very Large Data Bases. VLDB Endowment, pp. 155–166.
[3]. J. T. Brassil, S. Low, and N. F. Maxemchuk,Oct 1999 "Copyright protection for the electronic distribution of text documents," Proceedings of the IEEE, vol. 87, no. 7, pp. 1181–1196.
[4]. P. E. Gill, W. Murray, and M. A. Saunders,Jan2005 "Snopt: An sqp algorithm for large-scale constrained optimization," SIAM review, vol. 47, no. 1, pp. 99–131.
[5]. R. Hassan, B. Cohanim, O. De Weck, and G. Venter,Mar 2005 "A Comparison Of Particle Swarm Optimization And The Genetic Algorithm," in 46 th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. American Institute of Aeronautics and Astronautics, pp. 1–13.
[6]. K. E. Parsopoulos and M. N. Vrahatis, Nov 2002"Particle swarm optimization method for constrained optimization problems," Intelligent Technologies–Theory and Application: New Trends in Intelligent Technologies, vol, pp. 214–220.
[7]. P. W. Wong, Nov1998"A public key watermark for image verification and authentication," in Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, vol. 1. IEEE, pp. 455–459.

[8]. P. W. Wong and N. Memon, June2001"Secret and public key image watermarking schemes for image authentication and ownership verification," Image Processing, IEEE Transactions on, vol. 10, no. 10, pp. 1593–1601.