© 2015 IJSRSET | Volume 1 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Reversible Watermarking Technique based on Time Stamping in a Relational Data

Aishwarya C, Aishwarya S, Sathish Saravanan P
Information Technology, Dhanalakshmi College of Engineering, Chennai, Tamilnadu, India

ABSTRACT

The main aim of this project is to maintain the ownership of Relational Database and also minimizing distortion in the watermarked content. Reversible watermarking is employed to ensure data quality along-with data recovery. However, such techniques are usually not robust against malicious attacks and do not provide any mechanism to selectively watermark a particular attribute. Therefore, reversible watermarking is required that ensures watermark encoding and decoding by accounting for the role of all the features in knowledge discovery and original data recovery.

Keywords: Reversible Watermarking, Data Recovery, Data Quality, Robust

I. INTRODUCTION

Over the last years, usage of Internet and cloud computing is widely used; Data is stored in different formats such as images, videos, audios, texts, relational data. Relational data is widely managed by the owners of relational database.

Reversible watermarking is employed to ensure data quality along-with data recovery. However, such techniques are usually not robust against malicious attacks and do not provide any mechanism to selectively watermark a particular attribute by taking knowledge discovery into account.

II. METHODS AND MATERIAL

A. Problem Statement

Reversible watermarking is employed to ensure data quality along-with data recovery. However, such techniques are usually not robust against malicious attacks and do not provide any mechanism to selectively watermark a particular attribute by taking into account its role in knowledge discovery.

In existing system MAC is used for Hash Function. The parameters selection for watermarking is based on computing message authenticated code (MAC), where MAC is calculated using the secret key and the tuple's primary key. This technique assumes LSB manipulation during watermark embedding process. Though LSB data hiding techniques are efficient, an attacker can easily remove watermark by simple manipulation of data by shifting LSB. Hence alternate way has to be found.

B. Proposed System

In 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 (Ks) 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.

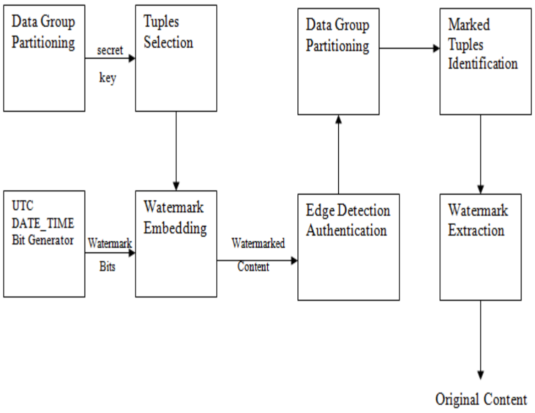

Figure1: Architecture diagram (overall representation)

III. RESULTS AND DISCUSSION

A. Feature Analysis And Selection

The initial step 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.

$$par(r)=H(ks||H(r.Pk||ks))\bmod m$$

Where r:PK is the primary key of the tuple r,H() is a cryptographic hash function Message Digest (MD5),|| is the concatenation, ks is a secret key. Logical groups or Partitions has been arrived after applied this algorithm. Admin has to decide the group's length that is m.

Algorithm

STEP 1. Append padding bits

The input message is "padded" (extended) so that its length (in bits) equals to 448 mod 512. Padding is always performed, even if the length of the message is already 448 mod 512.

STEP 2. Append length

A 64-bit representation of the length of the message is appended to the result of step1. If the length of the message is greater than 2^64, only the low-order 64 bits will be used.

The resulting message (after padding with bits and with b) has a length that is an exact multiple of 512 bits. The input message will have a length that is an exact multiple of 16 (32-bit) words.

STEP 3. Initialize MD buffer

A four-word buffer (A, B, C, D) is used to compute the message digest. Each of A, B, C, D is a 32-bit register. These registers are initialized to the following values in hexadecimal, low-order bytes first):

word A: 01 23 45 67
word B: 89 ab cd ef
word C: fe dc ba 98
word D: 76 54 32 10

STEP 4. Process message in 16-word blocks

Four functions will be defined such that each function takes an input of three 32-bit words and produces a 32-bit word output.

F (X, Y, Z) = XY or not (X) Z
G (X, Y, Z) = XZ or Y not (Z)
H (X, Y, Z) = X xor Y xor Z
I (X, Y, Z) = Y xor (X or not (Z))

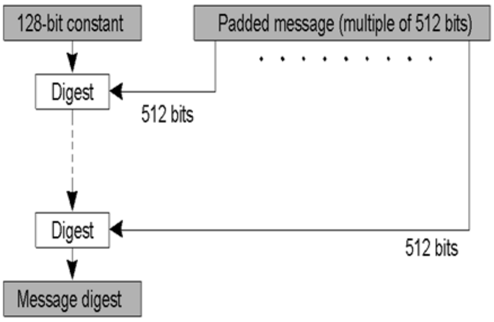

Figure 2: MD5 Algorithm Structure

B. Watermark Encoding

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

Where 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.

C. 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 re-watermark 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.

Approach Used

UTC divides time into days, hours, minutes and seconds. Days are conventionally identified using the Gregorian calendar, but Julian day numbers can also be used. Each day contains 24 hours and each hour contains 60 minutes. The number of seconds in a minute is usually 60, but with an occasional leap second, it may be 61 or 59 instead. Thus, in the UTC time scale, the second and all smaller time units (millisecond, microsecond, etc.) are of constant duration, but the minute and all larger time units (hour, day, week, etc.) are of variable duration.

D. Edge Detection Authentication and Watermark Encoding

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 to done. Otherwise it's not done.

IV. CONCLUSION

This paper formulates in ensuring data quality along-with data recovery provided with a mechanism to selectively watermark a particular attribute by taking knowledge discovery into account and by using UTC as a key. And thus this mechanism helps in providing robust and reversible watermarking to a Relational Data.

V. REFERENCES

[1] Y.-C. Liu, Y.-T. Ma, H.-S. Zhang, D.-Y. Li, and G.-S. Chen, " A method for trust management in cloud computing: Data coloring by cloud watermarking," International Journal of Automation and Computing, vol. 8, no. 3, pp. 280–285, 2011.

[2] Walmart to start sharing its sales data, http://www.nypost.com/p/news/business/walmart-opensup, last updated: February 4, 2012, last accessed: July, 20 2013.

[3] Identity theft watch, http://www.scambook.com/ blog/2013/04/identity-theft-watch-customer-passwords-stolen-fromwalmart-vudu-video-service/, last updated: April 11, 2013,last accessed: July, 20 2013.

[4] Securing outsourced consumer data, http://www.databreaches.net/securing-outsourced-consumerdata/,last updated: February 26, 2013, last accessed: July, 20 2013.

[5] As patients' records go digital, theft and hacking problems grow, "http://www.kaiserhealthnews.org/Stories/2012/June/04/ electronic-health-records-theft-hacking.aspx, last updated: June 03, 2012, last accessed: July, 20 2013.

[6] Y.-R. Wang, W.-H. Lin, and L. Yang, "An intelligent watermarking method based on particle swarm optimization," Expert Systems with Applications, 2011.

[7] M. Kamran and M. Farooq, "An information-preserving watermarking scheme for right protection of emr systems," Knowledge and Data Engineering, IEEE Transactions, 2012.

[8] T. M. Cover and J. A. Thomas, Elements of information theory. Wiley-interscience, 2012.

[9] K. Jawad and A. Khan, "Genetic algorithm and difference expansion based reversible watermarking for relational databases," Journal of Systems and Software, 2013.

[10] Message Digest algorithms designed by Professor Ronald Rivest of MIT (Rivest, 1992).