International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Email:editorijoer@gmail.com http://www.ijoer.in

Vol.4., Issue.1., 2016
(Jan-Feb)

RESEARCH ARTICLE

ISSN: 2321-7758

RRW - A RESILIENT REVERSIBLE WATERMARKING TECHNIQUE FOR THE PRECLUSION OF INFORMATION FROM CYBER PUNKERS

T.THILAGAM[1], R.VINOTH[2]

[1] Assistant Professor, CSE Gojan school of business and Technology
[2] Assistant Professor, IT Gojan school of business and Technology

T.THILAGAM

R.VINOTH

ABSTRACT

Securing the ownership and controlling the copies of Information have become very important issues in Internet-based applications. Reversible watermark technology allows the distortion-free recovery of relational databases after the embedded watermark data are detected or verified. In this paper we propose a robust technique of embedding reversible watermark in a relational database with non-numeric attributes. In this paper, we propose a robust, resilient and reversible watermarking scheme for non-numeric data that can be used to provide proof of ownership for the owner of a relational database. We implement a new approach to generate the watermark characters from UTC (Coordinated Universal Time) date time equivalent ASCII characters which is the primary time standard used to synchronize the time all over the world. In order to analyze these techniques, a classification has been performed on the basis of (i) the extent of modifications introduced by the watermarking scheme in the underlying non-numeric data and (ii) the robustness of the embedded reversible watermark technique for preclusion of information in relational database against variety of cyber punkers or data attacks.

Key Words: Reversible Watermarking, Genetic Algorithm, Robustness, Non-Numerical Data

I. INTRODUCTION

Reversible Watermarking is a technique which enables information to be authenticated and then restored to their original form by removing the watermark and replacing the data that had been overwritten [5]. Security is of increasing concern with databases for database's high added values and extensive installation in modern information systems. In addition to encryption, watermarking techniques is practically proven as another possible solution to enhance databases' content security especially for copyright protection and data tampering detection [1]. Unlike encryption or hash description, typical watermarking techniques modify original data as a modulation of the watermark information, and inevitably cause permanent distortion to the original data, and therefore cannot meet the integrity requirement of the data in some applications. This underlying defect can be relieved by reversible watermarking techniques by their reversibility in both robust watermarking and fragile watermarking

The direct beneficiary from this reversibility is those applications requiring zero permanent distortions such as medical imaging, military imaging, forensics of documents and art work

authentication. On the other hand, the perfect restoring ability realizes watermarking based lossless authentication which accounts for the major part of earlier algorithms. In recent years, researches on reversible watermarking center on increasing embedding capacity to meet requirements of large volume data embedding.

II. RELATED WORK

Approached proposed by [3] uses non numeric attribute to compute Eigen matrix and Eigen value to generate secrete key and to identify position to watermark. This method limits itself as fewer Eigen matrixes may have fewer than real roots, or no real roots at all and ASCII value cannot be computed for Hindi phonemes.

Using non numeric attribute another method is proposed by [4] perform matrix operation by forming the matrix using number of vowels, consonants and ASCII value of each character the result is used to scale the image related to that record (profile image). In this method it is not necessary that every relational database has an image associated to each record. Again use of ASCII is problem and restricted by 3*3 matrix only.

Using predefined signals of each ASCII character and a set of abbreviation of words a novel method is proposed by [5]. But this method depends on signal processing tools and to maintain and compute data in signal form is bulky and time consuming. Another approach of embedding watermarks in non- numeric attribute is proposed by [5], where hash functions and secrete key is used to generate watermarks.

Genetic Algorithm based on Difference Expansion watermarking (GADEW) technique is used in a proposed robust and reversible solution for relational databases [8]. GADEW improves upon the drawbacks mentioned above by minimizing distortions in the data, increasing watermark capacity and lowering false positive rate. To this end, a GA is employed to increase watermark capacity and minimize introduced distortion. This is because the watermark capacity increases with the increase in number of features and the GA runs on more features to search the optimum one for watermarking. However, watermark capacity decreases with the increase in watermarked tuples. GADEW used the distortion measures (AWD and TWD) to control distortions in the resultant data. In this context, the robustness of GADEW can be compromised when AWD and TWD are given high values.

III. OUR APPROACH

In our approach, unlike previous approaches we concentrate on tuples with their entirety rather than a subset of their attributes. Our approach aims to select tuples and embedding non-numeric data as reversible watermark for data then insert them erroneously into the database.

In the proposed system we implement a new approach to generate the watermark characters from UTC (Coordinated Universal Time) date time equivalent ASCII characters 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.

In order to analyze these techniques, a classification has been performed on the basis of (i) the extent of modifications introduced by the watermarking scheme in the underlying non-numeric data and (ii) the robustness of the embedded reversible watermark technique for preclusion of information in relational database against variety of cyber punkers or data attacks.

It is a big challenge encrypt and decrypt selected tuples should be inserted into the relation. This is because marks should not by any means degrade the quality of the data. For the number of selected tuples from partitioned data, we implement approach RRW for reversible watermarking of relational databases that improves data recovery ratio. The main architecture of RRW is presented in figure 1. RRW includes the following four major phases: (1) Secret key generation; (2)

watermark encryption; (3) watermark decryption; and (4) data recovery.

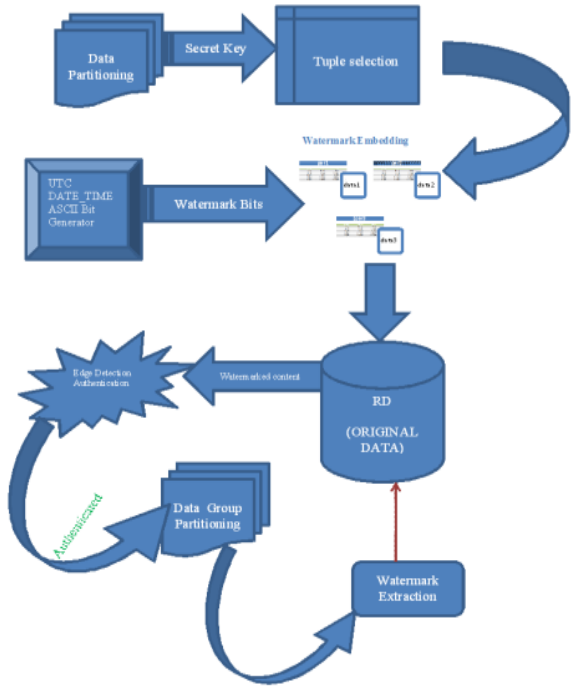

Figure 1. Main Architecture of RRW

The Watermarking process includes encryption and decryption Phase. The encryption phase consist of Data partitioning, Selection of data set for watermarking, Watermark embedding process. Decryption phase consist also these process to extract the Watermarked content.

A).Feature Analysis and Selection

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.

par(r)=H(ks||H(r.Pk||ks))mod 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 groups length that is m.

B).Secret Key Generation

In the Secret Key Generation phase, two important tasks are accomplished: (1) selection of a suitable tuples for watermark embedding; (2) calculation of an optimal watermark with the help of an optimization technique.

TABLE 1: Notations Used in the Paper

Symbol	Description
D	Original database
D_W	Watermarked database
R	Total number of tuples/rows/records in a table (or dataset)
A	a feature/column/attribute selected for watermarking (D)
min(a)	The minimum value of a feature
RRW	proposed Robust and Reversible Watermarking technique
w	Watermark bits
W_D	Decoded watermark
D_W	D watermarked by the proposed scheme
^{D}r	Recovered Data
MI_W	Mutual information of watermarked data
b	The watermark bit
r	A tuple in the database table
r	A matrix containing percent change in data values
$_r$	The difference between the changes detected in the value of a feature during the encoding and decoding process
dtW	The watermark decoder used for watermark decoding
$_D$	Mean of the original data of RRW
D_W	Mean of the watermarked data of RRW
$^0_{D_W}$	A watermarked database after the malicious attacks

The watermark bits can generates from UTC (Coordinated Universal Time) date time equivalent ASCII characters using Genetic Algorithm (GA).

For the creation of optimal watermark information, that needs to be embedded in the original data, we use an evolutionary technique; GA.

. The GA preserves essential information through the application of basic genetic operations to these chromosomes that include: selection, crossover, mutation and replacement. The GA evaluates the quality of each candidate chromosome by employing a fitness function. The evolutionary mechanism of the GA continues through a number of generations, until some termination criteria is met.

C).Watermark Encoding

A Tuple is one record or one row in a Relational Database. In this phase to select the

Particular tuples for embedding Watermarked Content Threshold Computation is a method 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

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.

Algorithm 1 Watermark Encoding

Output: D_W,r

for w= 1tol

do

//loop will iterate for all watermark bits w from 1 to length l of the watermark

forr= 1toR **do**

//loop will iterate for all tuples of the data

if$b_{r;w}$== 0 **then**

// the case when the watermark bit is 0 changes are calculated by using equation 6 data is watermarked by using equation 8 insert $_r$ into r

end if

if$b_{r;w}$== 1 **then**

// the case when the watermark it is 1 changes are calculated by using equation 6 data is watermarked by using equation 7 insert $_r$ into r

end

if end for

end for return D_W ; r

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; and 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.

D).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 multi bit 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.

E).Edge detection Authentication and Watermark Decoding

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.

Algorithm 2 Watermark Decoding

Input: D_W or D_W^0,r,l **Output:** W_D

for r= 1 to R **do**
//loop will iterate for all tuples of the data **for** b=l to 1 **do**
//loop will iterate for all watermark bits b from 1 to length l of the watermark
$d_r^{(D}{}_W{}^0{}_{(r)r}$ (d_r r
0 **then**
detected watermark bit (dtW) is 1
else if $_r$>0 and $_r$1 **then** detected watermark bit (dtW) is 0
end if end for
end for
W_D(mode(dtW (1; 2; :::; l))
return W_D

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. CONCLUSIONS AND FUTURE WORK

In this paper, 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. RRW is also evaluated through attack analysis where the watermark is detected with maximum decoding accuracy in different scenarios. A number of experiments have been conducted with different number of tuples attacked. Proposed method does not depend on primary key because Secret key is created. It depends on non-numeric attribute. Secret key is stored with trusted party and/or once again can be computed with original values. It is robust against various malicious attacks. we implement a new approach RRW using Genetic Algorithm and discussed the Watermark encryption and watermarking decryption algorithms in details. Our goal has been to ensure that our Reversible watermarking approach satisfies edge authentication detection and secure original information from cyber punkers.

This study can be further extended to any text data. Further approaches can be derived from more secure and proven algorithm in cryptography.

ACKNOWLEDGEMENT:

We thank our Management for supporting and encourage doing our research work. We thank our team members and family members for supporting us.

V. REFERENCE

[1]. Cox, M. Miller, J. Bloom, and M. Miller, Digital watermarking. Morgan Kaufmann, 2001.

[2]. M. E. Farfoura, S.-J. Horng, J.-L. Lai, R.-S. Run, R.-J. Chen, and M. K. Khan, "A blind reversible method for watermarking relational databases based on a time-stamping protocol," Expert Systems with Applications, vol. 39, no. 3, pp. 3185–3196, 2012.

[3]. R Bedi, A Thengade& V M Wadhai, "A New Watermarking Approach for Non-numeric Relational Database", International Journal of Computer Applications, 13(7)., 2011, pp. 37-40.

[4]. Bedi, R., Wadhai, V.M., Sugandhi, R., Mirajkar, A.: Watermarking Social Networking Relational Data using Non-numeric Attribute. International Journal of Computer Science and Information Security (IJCSIS) 9(4), 2011, pp 74–77.

[5]. Chin-Chen Chang, Thai-Son Nguyen, and Chia-Chen Lin, "A blind reversible robust watermarking scheme for relational database", The Scientific World Journal, vol 2013, Article ID 717165, 12 pages.

[6]. NahlaEl_Hahhar, M. M. Elkhouly, Samah S. Abu El Alla, "Blind watermarking technique for relational database", COMPUSOFT, an International Journal of Advanced Computer Technology, 2(3), May 2013, pp. 121-126.

[7]. M. Mitchell, "An introduction to genetic algorithms mit press," Cambridge, Massachusetts. London, England, 1996.

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

[9]. SamanIftikhar, M. Kamran and Zahid Anwar, "RRW – A Resilient and Reversible

Watermarking Techniques for Relational Data" 1041-4347 (c) 2013 IEEE.

[10]. VahabPournaghshband, " A New Watermarking Approach for Relational Data ",University of California, Berkeley

[11]. Agrawal, R., Haas, P., and Kiernan, J. 2003. Watermarking relational data: framework, algorithms and analysis. *TheVLDB Journal* 12, 2 (Aug. 2003), 157-169.

A Brief Bio of Authors

T.Thilagam is an assistant professor in the Department of Computer Science and Engineering, Gojan School of Business and Technology, Anna University. She received B.E. (2010) in Computer Science and Engineering, Gojan School of Business and Technology, Anna University and M.E. (2012) in Computer Science and Engineering, St.Peter's University, Chennai, India respectively. Her current research interests include digital image processing, network security and privacy, cloud computing and Distributed Data mining.

R.Vinoth is an assistant professor in the Department of Information Techonolgy, Gojan School of Business and Technology, Anna University. He received B.E. (2012) in Computer Science and Engineering, Anna University and M.E. (2014) in Computer Science and Engineering, , Anna University, Chennai, India respectively. His current research interests include Cryptography and privacy, Information security, Distributed Data mining and Network Security.