ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

A Tidemark for Image Verification and Validation using UTC

V.Dhivya
Final year, Information Technology,
Sri Vidya College Of Engineering & Technology
Virudhunagar,
dhives95@gmail.com

K.Lakshmi
Final year, Information Technology,
Sri Vidya College Of Engineering & Technology
Virudhunagar,
amuabi1623@gmail.com

Ms.B.Vanaja
Assistant Professor,
Sri Vidya College Of Engineering & Technology
Virudhunagar,
vanajanjoy@gmail.com

Abstract - In this developing world information's are playing vital role so, it became a challenging task to make those information confidential. Watermarking method is to recognizable pattern used to identify authenticity. Intentionally introduced pattern in the data is hard to find and destroy, robust against malicious attack. The owner of the Relational Database embeds the data in the gray scale image and in verification side the image given by the user is verified with the image encrypted previously and finally decrypted with the help MD5 algorithm as a result, sharing of data between its owners and legitimate users. (i) Watermark encoding and decoding by accounting for the role of all the features in knowledge discovery; and, (ii) original data recovery in the presence of active malicious attacks. In this paper, a robust and semi-blind reversible watermarking (RRW) technique for numerical relational data has been proposed that addresses the above objectives. Experimental studies prove the effectiveness of RRW against malicious attacks and show that the proposed technique outperforms existing ones.
Keywords*:* Relational database, reversible watermarking (RRW), encryption, MD5

1. INTRODUCTION

Data mining, the extraction of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions.

In the digital world of today, data is excessively being generated due to the increasing use of the Internet and cloud computing. Data is stored in different digital formats such as images, audio, video, natural language texts and relational data. Relational data in particular is shared extensively by the owners with research communities and in virtual data storage locations in the cloud [1]. The purpose is to work in a collaborative environment and make data openly available so that it is useful for knowledge extraction and decision making. RRW mainly comprises a (1) data preprocessing phase, (2) watermark encoding phase, (3) attacker channel, (4) watermark decoding phase and (5) data recovery phase. Reversible watermarking is employed to ensure data quality along-with data recovery [1]. 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. LSB-based data hiding techniques are efficient, but an attacker is able to easily remove watermark by simple manipulation of data by shifting LSB. 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. A robust watermark algorithm is used to embed

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

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(Message Digest Algorithm) 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. Decoding phase consist also these process to extract the Watermarked content.

2. RELATED WORK

Copyright protection and authentication of digital contents has gained more importance with the increasing use of internet, e-commerce, and other efficient communication technologies. [5]. Authentication and copyright protection of digital images, audio, and video. Text is easier to copy, reproduce and tamper as compared with images, audio and video. Cryptography-based AWT and authentication signature (AS) is computed from the whole image data and inserted into the image itself. In cryptography, AS is called message authentication code (using secret-key) or digital signature (using public/private-key). AWTs for binary images present a watermarking to detect unintentional changes in halftone images, but this cannot be considered to be an AWT because it does not withstand intentional or malicious attacks. [4]In a cryptography-based AWT but it is especially suited for dispersed-dot halftone images and the visual quality for a generic binary image. Image watermarking techniques can be easily applied to a text document, but they introduce a very noticeable white noise in the text document. We present a new method for watermarking electronic text documents which is similar to the word-shift coding and line-shift coding methods described above [2]. We present a sequential quadratic programming (SQP) method for large-scale optimization problems involving general linear and nonlinear con- strains. SQP methods have proved reliable and efficient for many such problems [6]. Detecting the watermark and it works with documents that contain aligned left, centered, aligned right, or justified paragraphs as well as regular or irregular line spacing. Justified paragraphs are very common in electronic documents [5]. The authors proposed a robust, blind, resilient and reversible, image based watermarking scheme for large scale databases. Genetic algorithm based on difference expansion water- marking (GADEW) technique is used in a proposed robust and reversible solution for relational databases [1]. Fingerprinting, data hashing, serial codes are some other techniques used for ownership protection. Fingerprints also called transactional watermarks, are used to monitor and identify digital ownership by watermarking all the copies of contents with different watermarks for different recipients [8],[5]. Reversible water- marking technique introduces distortions as a result of the embedding process. Changes in the data are controlled by placing certain bounds on LSB [1].

3. REVERSIBLE WATERMARKING FOR RELATIONAL DATA

We implement a new approach to generate the watermark bits from UTC (Coordinated Universal Time) date and 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

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

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.

4. IMPLEMENTATION

MD5 is a widely used cryptographic function with a 128- bit hash value. MD5 has been employed in a wide variety of security applications, and is also commonly used to check the integrity of files. An MD5 hash is typically expressed as a 32-digit hexadecimal number. MD5 processes a variable-length message into a fixed-length output of 128 bits.

Steps for MD5:

1).The input message is broken up into chunks of 512-bit blocks (sixteen 32-bit little endian integers), the message is padded so that its length is divisible by 512. 2).The padding works as follows: first a single bit, 1, is appended to the end of the message.3).This is followed by as many zeros as are required to bring the length of the message up to 64 bits fewer than a multiple of 512. 4).The remaining bits are filled up with a 64-bit integer representing the length of the original message, in bits. 5).The MD5 algorithm uses 4 state variables, each of which is a 32 bit integer (an unsigned long on most systems). These variables are sliced and diced and are (eventually) the message digest. 6).output the hex values of each the state variables, least significant byte first.

EXAMPLE:

MD5 ("The quick brown fox jumps over the lazy dog") = 9e107d9d372bb6826bd81d3542a419d6

It is easy to compute the hash value for any given message, it is infeasible to find a message that has a given hash, it is infeasible to modify a message without changing its hash, and it is infeasible to find two different messages with the same hash. The security of the MD5 hash function is severely compromised. A collision attack exists that can find collisions within seconds on a computer with a 2.6 GHz Pentium4 processor (complexity of 2 24.1). A number of projects have published MD5 rainbow tables online, that can be used to reverse many MD5 hashes into strings that collide with the original input, usually for the purposes of password cracking. It is mainly depends on images rather than alphanumerical. The main argument here is that pass-images from the recognizing and memorizing pictures.

```
/*
 Message-digest routines:
 *To form the message digest for a message
M                    **
 **   (1) Initialize a context buffer context using
MD5Init          **
 **    (2) Call MD5Update on context and
M                      **
 **    (3) Call MD5Final on
context                          **
 ** The message digest is now in context->digest
[0...15]          **
 */
```

5. Block diagram and modules

5.1 Block diagram

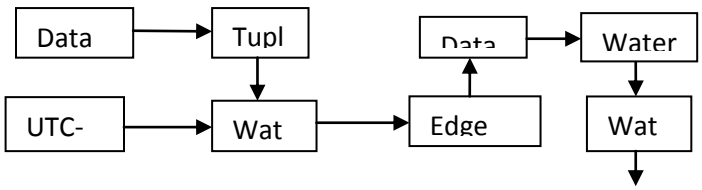

Fig 1: Block diagram

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

In the above shown block diagram, Data Partitioning comes under Watermark Encoding Phase which has been done by owner of the Database (i.e.) Admin. The data partitioning algorithm to partitions the data set. 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 in each attribute for tuples selection. Cryptographic hash function MD5 is applied on the selected data set to select only those tuples. The watermark generating function takes date-time stamp as an input and then generates watermark bits. These bits are given as input to the watermark encoding function. These watermark bits are embedded in the original data set by using watermark embedding algorithm. Edge detection Authentication is proposed as an alternative solution for text based.Watermark Extraction process in the Decoding phase. The main argument here is that pass-images from the recognizing and memorizing pictures. It is mainly depends on images rather than alphanumerical

5.2 MODULES

- Database Collection and Partitioning
- Tuples Selection for Watermarking.
- Watermark Embedding.
- Edge detection Authentication.
- Watermark Decoding.

5.2.1 Database Collection and Partitioning

In this module includes the Data partitioning Relational Numerical Database for Watermarking. Data Partitioning comes under Watermark Encoding Phase which has been done by owner of the Database (i.e.) Admin. The data partitioning algorithm to partitions the data set

$$Par(r) = H(KS||H(r.pk||ks)) \bmod m$$

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 in fig 2. Admin has to decide the group's length that is m.

Fig 2. Database collection and partitioning

5.2.2 Tuples Selection for Watermarking

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

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

computed threshold, it is selected for Encoding Process fig3. 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 in fig 4.Cryptographic hash function MD5 is applied on the selected data set to select only those tuples.

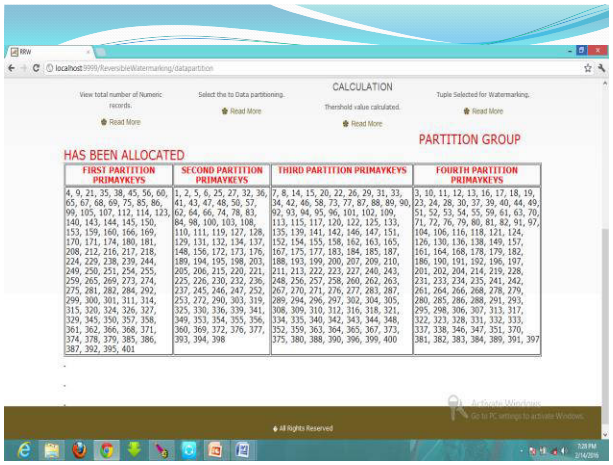

Fig 3. Threshold calculation

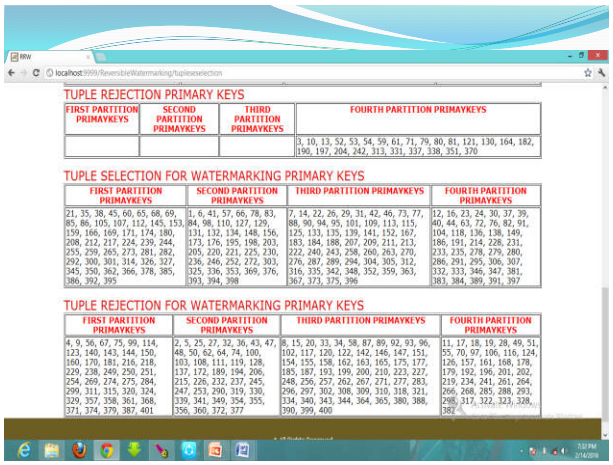

Fig 4. Tuple selection

5.2.3 Watermark embedding

To construct a watermarked data set, these watermark bits are embedded in the original data set by using watermark embedding algorithm. The watermark generating function takes date-time stamp as an input and then generates watermark bits. These bits are given as input to the watermark encoding function in fig 5.

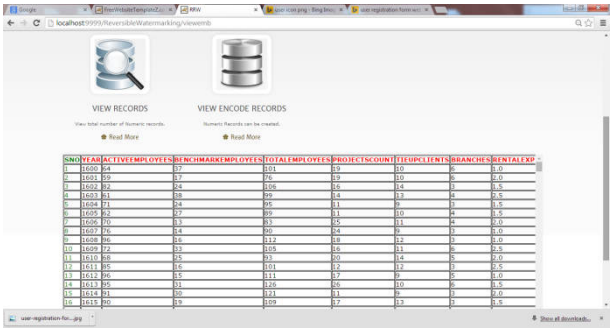

Fig 5. Watermark encode

5.2.4 Edge detection authentication

Edge detection Authentication is proposed as an alternative solution for text based. It is mainly depends on images rather than alphanumerical. The main argument here is that pass-images from the recognizing and memorizing pictures in fig.6. Watermark Extraction process in the Decoding phase.

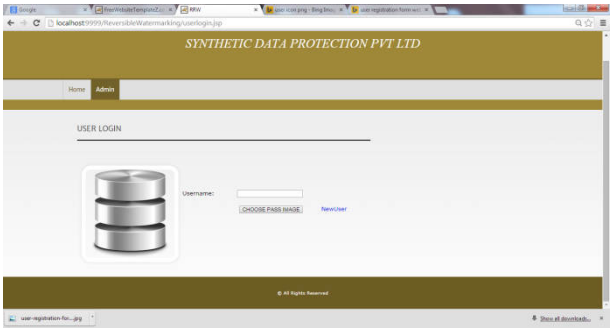

Fig 6. Pass images (user login).

5.2.5 Watermarking decoding

Watermark decoding is performed in the reverse order of the watermark embedding process; the last embedded bit is decoded first and so on. The original data is recovered along with watermark bits from watermarked data in fig.7(a)&(b).

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016

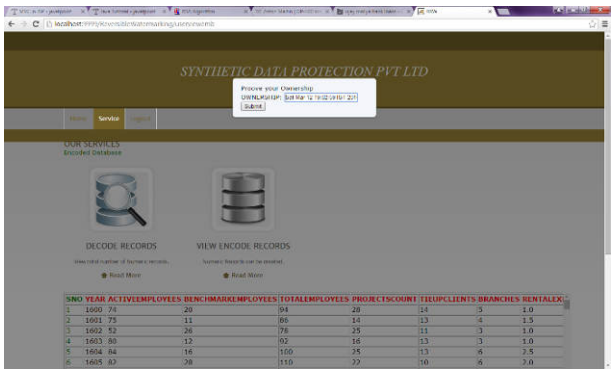

Fig 7(a).Watermark decoding

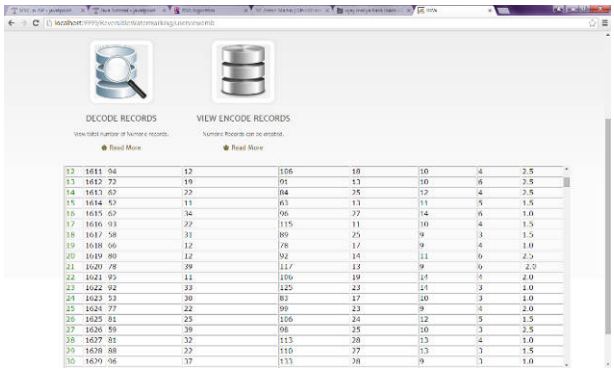

Fig 7(b).Watermark decoding

6. CONCLUSION

Irreversible watermarking technique makes changes in the content of data. It is reduce the quality of data. These techniques are not robust against malicious attacks. Reversible watermarking techniques are used to recover original data from watermarked data and ensure data quality to some extent. In this paper, a novel robust and reversible technique for watermarking numerical data of relational databases is presented. it allows recovery of a large portion of the data even after being subjected to malicious attacks .If an intruder deletes, adds or alters up to 50 percent of tuples, RRW is able to recover both the embedded watermark and the original data.

7. REFERENCES

1. Saman Iftikhar, M. Kamran, and Zahid Anwar," RRW—A Robust and Reversible Watermarking Technique for Relational Data", IEEE transactions on knowledge and data engineering, VOL. 27, NO. 4, APRIL 2015

2. Zunera Jalil1, Anwar M. Mirza1 and Maria Sabir2 "Content based Zero-Watermarking Algorithm for Authentication of Text Documents", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 2, February 2010.

3. Hae Yong Kim* and Ricardo L. de Queiroz** "A public-key authentication watermarking for binary images"

4. H. Y. Kim and A. Afif, "Secure Authentication Watermarking for Binary Images," in Proc. Sibgrapi - Brazilian Symp. on Comp. Graph. and Image Proc., pp. 199-206, 2003.

5. Adnan M. Alattar and Osama M. Alattar"Watermarking Electronic Text Documents Containing Justified Paragraphs and Irregular Line Spacing "

6. Philip E.Gill,Walter Murray,and Michael A.saunders" SNOPT: an sqp algorithm for large-scale constrained optimization" Vol. 12, No. 4, pp. 979–1006,2002

7. E. Sonnleitner, "A robust watermarking approach for large data- bases," in Proc. IEEE First AESS Eur. Conf. Satellite Telecommun., 2012, pp. 1–6.

8. S. Subramanya and B. K. Yi, "Digital rights management," IEEE Potentials, vol. 25, no. 2, pp. 31–34, Mar.-Apr. 2006.

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 15, March 2016