ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Reversible Water Marking Technique Based On A Time-Stamping In Relational Data.

J Jackulin Reeja*.

Faculty of Computing Sathyabama University, Chennai, Tamil Nadu, India.

ABSTRACT

Advancement in information technology is playing an increasing role in the use of information systems comprising relational databases. These databases are used effectively in collaborative environments for information extraction consequently; they are vulnerable to security threats concerning ownership rights and data tampering. Watermarking method is to recognizable pattern used to identify authenticity. Watermarking is advocated to enforce ownership rights over shared relational data and for providing a means for tackling data tampering. When ownership rights are enforced using watermarking, the underlying data undergoes certain modifications; as a result of which, the data quality gets compromised. 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. Therefore, reversible watermarking is required that ensures (i) watermark encoding and decoding by accounting for the role of all the features in knowledge discovery & (ii) original data recovery in the presence of active malicious attacks. In this , a robust and semi-blind reversible watermarking technique for numerical relational data has been proposed that addresses the above objectives. Experimental studies prove the effectiveness of Reversible watermarking against malicious attacks and show that the proposed technique outperforms existing ones..

Keywords: Reversible Watermarking, Coordinated Universal Time(UTC),Message Digest 5 Hash Function(MD5),Gray scale Edge Detection Image.

***Corresponding author**

INTRODUCTION

The main aim of this paper is to maintain the ownership of Relational Database and also minimizing distortion in the watermarked content Watermarking method is to recognizable pattern used to identify authenticity[1].Intentionally introduced pattern in the data is hard to find and destroy, robust against malicious attack. WATERMARKING, without any exception, has been used for ownership protection of a number of data formats—images, video, audio, software, XML documents, geographic information system (GIS) related data, text documents, relational databases and so on—that are used in different application domains. Recently, intelligent mining techniques are being used on data, extracted from relational databases, to detect interesting patterns (generally hidden in the data) that provide significant support to decision makers in making effective, accurate, and relevant decisions; as a result, sharing of data between its owners and legitimate users.

The owner of the Relational Database embeds the watermark data, the distortions in the original data are kept within certain limits, which are defined by the usability constraints, to preserve the knowledge contained in the data[2]. The proposed algorithm embeds every bit of a multibit watermark (generated from date-time) in each selected row (in a numeric attribute) with the objective of having maximum robustness even if an attacker is somehow able to successfully corrupt the watermark in some selected part of the data set.

RELATED WORK

In [9], the authors proposed a robust, blind, resilient and reversible, image based watermarking scheme for large scale databases. The bit string of an image is used as a watermark where one bit from the bit string is embedded in all tuples of a single partition and the same process is repeated for the rest of the partitions. As the watermarking is performed on the entire Database, so it become very easy for the attackers to identify whether the Database is watermarked or not.

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 authenticated code (MAC), where MAC is calculated using the secret key and the tuple's primary key[3]. This technique assumes unconstrained LSB manipulation during watermark embedding process.

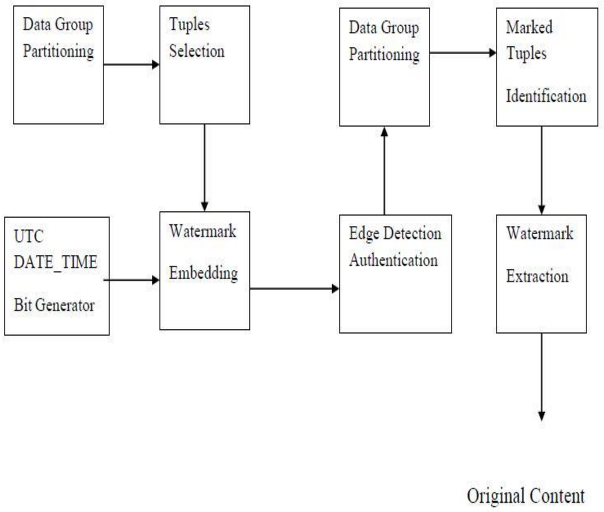
Although LSB-based data hiding techniques are efficient, but an attacker is able to easily remove watermark by simple manipulation of data by shifting LSB. The data partitioning concept is based on the use of special marker tuples, making it vulnerable to watermark synchronization errors.

PROPOSED WORK

This Proposed system 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 [7]. 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.

ARCHITECTURE

This architecture diagram shows how the tuples are selected from the data group partitioning in which data's are watermarked by using UTC and with the help of edge detection authentication reversible watermarking is executed.



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.

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 [5]. 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 decided the groups length that is m.

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 computed threshold, it is selected for Encoding Process[4]. The data selection threshold for an attribute is calculated by using the following equation:

$$T=c^* \text{ Mean}+ \text{ Standard Deviation}$$

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.

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[10]. 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.

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 [8]. 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.

EDGE DETECTION AUTHENTICATION AND WATERMARK XTRACTION:

Edge detection Authentication is proposed as an alternative solution to text based. It is mainly depends on images rather than alphanumerical[11]. 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[6]. 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.

CONCLUSION

In this paper we are using the UTC (coordinated universal time), which generate a unique secret key this makes the watermarking technique to be secure and confidential. And the edge detection authentication is used to impure the security of the paper.

Hence this can be used in the places which are based on numerical databases such as:

- Corporate sector
- Banking sector
- Education institutions

We have performed the watermarking technique only in the numerical database, the future enhancement of this paper will be performing the same action on the Alpha Numerical database.

REFERENCES

[1] P. W. Wong, "A public key watermark for image verification and authentication," in Image Processing,

[2] 1998. ICIP 98. Proceedings. 1998 International Conference on, vol. 1. IEEE, 1998, pp. 455–459.

[3] P. W. Wong and N. Memon, "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, 2001.

[4] F. A. Petitcolas, "Watermarking schemes evaluation," Signal Processing Magazine, IEEE, vol. 17, no. 5, pp. 58– 64, 2000.

[5] R. Agrawal and J. Kiernan, "Watermarking relational databases," in Proceedings of the 28th international conference on Very Large Data Bases. VLDB Endowment, 2002, pp. 155– 166.

[6] R. Sion, M. Atallah, and S. Prabhakar, "Rights protection for categorical data," Knowledge and Data Engineering, IEEE Transactions on, vol. 17, no. 7, pp. 912–926, 2005.

[7] S. Subramanya and B. K. Yi, “Digital rights management,” Potentials, IEEE, vol. 25, no. 2, pp. 31– 34, 2006.

[8] P. E. Gill, W. Murray, and M. A. Saunders, “Snopt: An sqp algorithm for large-scale constrained optimization,” SIAM review, vol. 47, no. 1, pp. 99– 131, 2005.

[9] K. E. Parsopoulos and M. N. Vrahatis, “Particle swarm optimization method for constrained optimization problems,” Intelligent Technologies– Theory and Application: New Trends in Intelligent Technologies, vol. 76, pp. 214–220, 2002.

[10] E. Sonnleitner, “A robust watermarking approach for large databases,” in Satellite Telecommunications (ESTEL), 2012 IEEE First AESS European Conference on. IEEE, 2012, pp. 1–6.

[11] R. Hassan, B. Cohanim, O. De Weck, and G. Venter, “A Comparison of Particle Swarm Optimization and The Genetic Algorithm,” in 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. American Institute of Aeronautics and Astronautics, 2005, pp. 1–13.

[12] J. T. Brassil, S. Low, and N. F.Maxemchuk, “Copyright protection for the electronic distribution of text documents,” Proceedings of the IEEE, vol. 87, no. 7, pp. 1181–1196, 1999.