*International Journal Of*

**Recent Scientific Research**

**ISSN: 0976-3031**

**Volume: 7(4) April -2016**

# IMPLEMENTATION OF INVISIBLE WATERMARKING TECHNIQUE IN RELATIONAL DATABASE

**Radha R., Siva Sankari K and Sri Devi S**

THE OFFICIAL PUBLICATION OF
INTERNATIONAL JOURNAL OF RECENT SCIENTIFIC RESEARCH (IJRSR)
http://www.recentscientific.com/ recentscientific@gmail.com

ISSN: 0976-3031

*Available Online at http://www.recentscientific.com*

*International Journal of Recent Scientific Research*
*Vol. 7, Issue, 4, pp. 10034-10037, April, 2016*

**International Journal of Recent Scientific Research**

**Research Article**

# IMPLEMENTATION OF INVISIBLE WATERMARKING TECHNIQUE IN RELATIONAL DATABASE

**Radha R[1]., Siva Sankari K[2] and Sri Devi S[3]**

[1,2,3]Department of Computer Science and Engineering Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai-62

**ARTICLE INFO**

***Article History:***

Received 16th January, 2015
Received in revised form 24th February, 2016
Accepted 23rd March, 2016
Published online 28th April, 2016

***Keywords:***

watermark embedding, edge detection authentication, data extraction and robustness

**ABSTRACT**

Watermarking is a recognizable pattern used to identify authenticity and it's main aim is to maintain the ownership of relational database. WATERMARKING, without any exception, has been used for ownership protection of a number of data formats like images, video, audio, software, XML documents, geographic information system (GIS) related data, text documents, relational databases and so on and also that are used in different application domains .In this paper, we are applying watermarking techniques for relational database. As a result of it, sharing of data between its owners and legitimate users becomes secure. The watermarking technique is applied for only a particular set of data and make it difficult to identify watermarked content from hackers and it also reduce the number of tuples to be watermarked.



## INTRODUCTION

In the digital world of today, data is excessively being generated due to the increasing use of the Internet and cloud computing [1]. 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. 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. Take the case of Walmart, a large multinational retail corporation that has made its sales database available openly over the Internet so that it may be used for the purposes of identifying market trends through data mining. However these openly available datasets make attractive targets for attacks. 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. 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 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 (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*

The first irreversible watermarking technique for relational databases was proposed by Agrawal and Kiernan in [12]. Similarly, the first reversible watermarking scheme for relational databases was proposed in [22]. In this technique, histogram expansion is used for reversible watermarking of relational database. Zhang *et al*. proposed a method of distribution of error between two evenly distributed variables and selected some initial nonzero digits of errors to form histograms. Histogram expansion technique is used to reversibly watermark the selected nonzero initial digits of

**Corresponding author: Radha R*
*Department of Computer Science and Engineering Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai-62*

errors. This technique is keeps track of overhead information to authenticate data quality. However, this technique is not robust against heavy attacks (attacks that may target large number of tuples). Difference expansion watermarking techniques (DEW), [23], [24], [25] exploit methods of arithmetic operations on numeric features and perform transformations. Genetic algorithm based on difference expansion watermarking (GADEW) technique is used in a proposed robust and reversible solution for relational databases [27]. Prediction-error expansion watermarking techniques (PEEW) like [28], [29], [30], [31], [32], [33] incorporate a predictor as opposed to a difference operator to select candidate pixels or features for embedding of watermark information. Gupta and Pieprzyks' [23], proposed reversible watermarking technique introduces distortions as a result of the embedding process. The reversible watermarking techniques DEW, GADEW and PEEW, proposed in [23], [27], [28] respectively, are not robust and reversible against heavy attacks. In RRW, all the tuples of the selected feature can be marked thanks to the selection of a low distortion watermark; therefore, the attacker will have to attack all the tuples to corrupt the watermark to mitigate the effect of the majority voting scheme. Attacking all the tuples is not a viable option for the attacker because he has no knowledge of the original data or the usability constraints and that would completely compromise its usefulness. Moreover, since RRW can afford to embed watermark bits in all or a large fraction of the tuples of the selected feature; it achieves high robustness against heavy attacks.

### *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 authenticated 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. The problem I the existing system is, it is not robust against malicious attacks and do not provide any mechanism to selectively watermark a particular attribute.

### *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 (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.

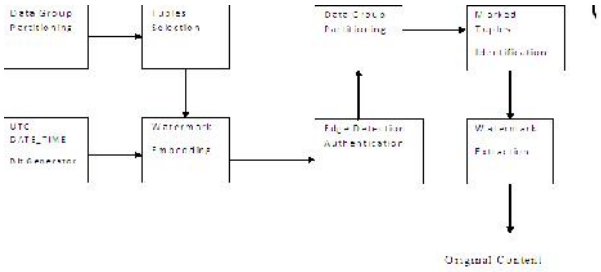

**Fig** System Architecture

## METHODOLOGY

### *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.par(r)=H(ks||H(r.Pk||ks))mod mwhere 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 of 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. 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. 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.

### *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.

***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 to done. Otherwise its not done.

## CONCLUSION AND FUTURE ENHANCEMENT

We achieved to maintain the ownership of Relational Database and also minimizing distortion in the watermarked content proves the security of relational database from being hacking. The watermarking technique used in this paper applied only for numeric database and future enhancement of this paper is, it can be applied to alphanumeric database, i.e., the whole database which includes both letters and numbers can be watermarked. The alphabetic database is watermarked by means of its ASCII value. This will further improve the security of the database from being hacked.

## References


1. Y.-C. Liu, Y.-T. Ma, H.-S. Zhang, D.-Y. Li, and G.-S. Chen, "Amethod for trust management in cloud computing: Data coloringby cloud watermarking," Int. J. Autom. Comput., vol. 8, no. 3, pp. 280–285, 2011.
2. (2012, Feb. 4).Walmart to start sharing its sales data. [Online]. Available: http://nypost.com/p/news/business/ walmart-opens-up
3. (2013, Apr. 11). Identity theft watch. [Online]. Available: http://scambook.com/blog/2013/ 04/identity-theft-watch-customerpasswords-stolen-from-walmart-vudu-video-service/
4. (2013, Feb. 26). Securing outsourced consumer data. [Online]. Available:http://databreaches.net/securing-outsourced-consumer-data/
5. (2012, Jun. 3). As patients' records go digital, theft and hacking problems grow. [Online]. Available:http://kaiserhealthnews.org/Stories/2012/June /04/electronic-health-records-thefthacking.aspx
6. I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673–1687, Dec. 1997.
7. I. Cox, M. Miller, J. Bloom, and M. Miller, Digital Watermarking. Burlington, MA, USA: Morgan Kaufmann, 2001.
8. P. W. Wong, "A public key watermark for image verification and authentication," in Proc. IEEE Int. Conf. Image Process., 1998, vol. 1,pp. 455–459.
9. P. W. Wong and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," IEEE Trans. Image Process., vol. 10, no. 10, pp. 1593–1601, Oct. 2001.
10. F. A. Petitcolas, "Watermarking schemes evaluation," IEEE Signal Process. Mag., vol. 17, no. 5, pp. 58–64, Sep. 2000.
11. J. T. Brassil, S. Low, and N. F. Maxemchuk, "Copyright protection for the electronic distribution of text documents," Proc. IEEE, vol. 87, no. 7, pp. 1181–1196, Jul. 1999.
12. R. Agrawal and J. Kiernan, "Watermarking relational databases," in Proc. 28th Int. Conf. Very Large Data Bases, 2002, pp. 155–166.
13. R. Sion, M. Atallah, and S. Prabhakar, "Rights protection for categoricaldata," IEEE Trans. Knowl. Data Eng., vol. 17, no. 7, pp. 912–926, Jul. 2005.
14. S. Subramanya and B. K. Yi, "Digital rights management," IEEE Potentials, vol. 25, no. 2, pp. 31–34, Mar.-Apr. 2006.
15. P. E. Gill, W. Murray, and M. A. Saunders, "Snopt: An sqpalgorithm for large-scale constrained optimization," SIAM Rev., vol. 47, no. 1, pp. 99–131, 2005.
16. K. E. Parsopoulos and M. N. Vrahatis, "Particle swarm optimization method for constrained optimization problems," Intel. Technol.–Theory Appl. New Trends Intell. Technol., vol. 76, pp. 214–220, 2002.
17. R. Hassan, B. Cohanim, O. De Weck, and G. Venter, "A comparison of particle swarm optimization and the genetic algorithm," in Proc. 46th AIAA/ASME/ASCE/AHS/ASC Struct., Struct. Dyn. Mater. Conf., 2005, pp. 1–13.
18. Y.-R. Wang, W.-H. Lin, and L. Yang, "An intelligent watermarking method based on parti le swarm optimization," Expert Syst.Appl., vol. 38, no. 7, pp. 8024–8029, 2011.
19. M. Kamran and M. Farooq, "An information-preserving watermarking scheme for right protection of EMR systems," IEEE Trans. Knowl. Data Eng., vol. 24, no. 11, pp. 1950–1962, Nov. 2012.
20. T. M. Cover, J. A. Thomas, and J. Kieffer, "Elements of information theory," SIAM Rev., vol. 36, no. 3, pp. 509–510, 1994.
21. T. M. Cover and J. A. Thomas, Elements of Information Theory. New York, NY, USA: Wiley-Interscience, 2012.

22. Y. Zhang, B. Yang, and X.-M. Niu, "Reversible watermarking for relational database authentication," J. Comput., vol. 17, no. 2, pp. 59–66, 2006.
23. G. Gupta and J. Pieprzyk, "Reversible and blind database watermarking using difference expansion," in Proc. 1st Int. Conf. Forensic Appl. Tech. Telecommun., Inf., Multimedia Workshop, 2008, p. 24.
24. A. M. Alattar, "Reversible watermark using difference expansion of triplets," in Proc. IEEE Int. Conf. Image Process., 2003, pp. I–501, vol. 1.
25. G. Gupta and J. Pieprzyk, "Database relation watermarking resilient against secondary watermarking attacks," in Information Systems and Security. New York, NY, USA: Springer, 2009, pp. 222–236.
26. J.-N. Chang and H.-C. Wu, "Reversible fragile database watermarking technology using difference expansion based on SVR prediction," in Proc. IEEE Int. Symp. Comput., Consum. Control, 2012, pp. 690–693.
27. K. Jawad and A. Khan, "Genetic algorithm and difference expansion based reversible watermarking for relational databases," J. Syst. Softw., vol. 86, no. 11, pp. 2742–2753, 2013.
28. M. E. Farfoura and S.-J. Horng, "A novel blind reversible method for watermarking relational databases," in Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl., 2010, pp. 563–569.
29. D. M. Thodi and J. J. Rodriguez, "Prediction-error based reversible watermarking," in Proc. IEEE Int. Conf. Image Process. 2004, vol. 3, pp. 1549–1552.
30. D. M. Thodi and J. J. Rodriguez, "Reversible watermarking by prediction-error expansion," in Proc. 6th IEEE Southwest Symp. Image Anal. Interpretation, 2004, pp. 21–25.
31. D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," IEEE Trans. Image Process., vol. 16, no. 3, pp. 721–730, Feb. 2007.
32. 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 Syst. Appl., vol. 39, no. 3, pp. 3185–3196, 2012.
33. X. Li, B. Yang, and T. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," IEEE Trans. Image Process., vol. 20, no. 12, pp. 3524– 3533, Dec. 2011.
34. E. Sonnleitner, "A robust watermarking approach for large databases," in Proc. IEEE First AESS Eur. Conf. Satellite Telecommun., 2012, pp. 1–6.
35. K. Huang, H. Yang, I. King, M. R. Lyu, and L. Chan, "Biased minimax probability machine for medical diagnosis," AMAI, 2004.
36. K. Bache and M. Lichman. (2013). UCI machine learning repository [Online]. Available: http://archive.ics.uci.edu/ml.

*******

**How to cite this article:**

Radha R., Siva Sankari K and Sri Devi S.2016, Implementation of Invisible Watermarking Technique in Relational Database. *Int J Recent Sci Res.* 7(4), pp. 10034-10037.

TSSN 0976-3031