A MULTI-ATTRIBUTE WATERMARKING TECHNIQUE ON RELATIONAL DATA FOR OWNERSHIP PROTECTION

Miss. Shubhangi Manohar Wadekar[#1], **Mr. Makarand Samvatsar**[*2]
*1(PG Student, Information Technology Department, Patel College of Science and Technology, Indore, India)
*2 (Associat. Professor, Information Technology Department, Patel College of Science and Technology, Indore)

Abstract— ***The property of personnel or organisation can be prove by using watermarking. Original data is encoded in the watermark bit and after the data decoding it is handover to the data owner in reversible watermarking technique. This technique can protect the ownership of data and data recovery after decoding of data. For the watermark data encoding our proposed technique uses feature selection and partitioning of data. This technique selects the multiple attribute from original database for the watermark encoding. Data distortion and information loss can be minimizing by use of data partitioning and pre-processing of data. The proposed technique result shows it is more efficient, scalable, and reversible against various attacks for the relational database as compare to the existing techniques.***

Keywords— *Watermarking, Reversible watermarking, non-numerical. Relational Databases, Ownership*

I. Introduction

In today's digital world, Internet is being use very rapidly in day by day, due to this data increasing and that data is stored in the different format i.e images, audio, video, natural language, texts, relational data. With the research association owners, the relational data is highly shared in particular manner. The Relational Databases contain the valuable information of various organizations and institutes. Sharing data online over the internet or virtual storage space is become vital activity for organization and, which may include importing and exporting of relational data [1]. In cloud environments, from un-trusted parties security threats arise. Watermarking is used to protect the ownership of data. Watermark techniques provide the protection to data by encoding. Watermarking can be threatened of malicious attack which may effect as alteration, deletion, or false insertion. The robust watermarking scheme possesses the exact recovery in the existence of active malicious attack [2].The watermarking techniques which are irreversible may causes modification of data undergoes through it at the certain extent [3].

1.1 Objectives-:

a. For the relational databases to provide robust security technique against such at-tacks.
b. To minimizes the information loss to maintain the data quality to designs a novel strategy for watermarking relational databases
c. To compare performance of proposed system with existing system and validate the same.

The remainder of the paper is organized as follows: Section 2 a brief overview of the different stages of the proposed watermarking technique is described and provides a detailed description of the proposed technique. Sections 3.To compare performance of proposed system with existing system by analysis of result. Finally, we conclude the paper with an outlook to future work.

II. PROPOSED SYSTEM

The data recovery radio can be improved by proposed watermarking architecture. The architecture is shown in Fig.1. Watermarking architecture includes the following major phases:
(1) Watermark pre-processing;
(2)Data Partitioning
(3) watermark encoding;
(4) watermark decoding;
(5) Data recovery.

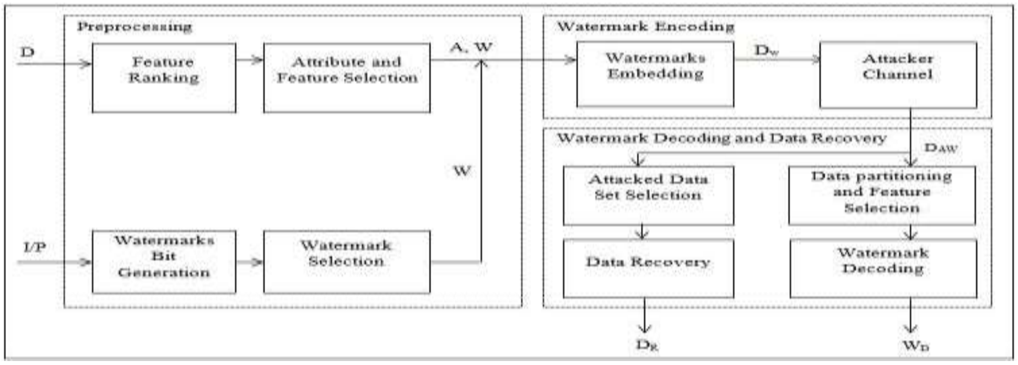
Preprocessing
D
Feature Ranking
Attribute and Feature Selection
A, W
W
I/P
Watermarks Bit Generation
Watermark Selection
Watermark Encoding
Dw
Watermarks Embedding
Attacker Channel
Watermark Decoding and Data Recovery
DAW
Attacked Data Set Selection
Data partitioning and Feature Selection
Data Recovery
Watermark Decoding
DR
WD

Fig 1. System Architecture

(1) Watermark Pre-processing Phase

In the watermark pre-processing phase, two important tasks are executed: (a) for watermark embedding to select a suitable feature; (b) calculation of an optimal watermark. For developing an information model of various features of the dataset, all the features are ranked according to their importance in information extraction, it based on mutual dependence on other features.

(2) Data Partitioning

In data partitioning phase data is logically partitioned into some clusters which are not having any common tuple. The partitioning algorithm can be given as:

Algorithm 1. Data Partitioning

```
Step 1:    Start
Step 2:    Read the values of D, m, $K_s$; P K
Step 3:    Repeat for all r from 1 to R in D
Step 4:    $par_r H(K_s(r:P\ K\ k\ K_s))modm$
Step 5:    Return $S_0$; :::; $S_{m\ 1}$
Step 6:    Stop
```

(3) Watermark Encoding

The main focus of watermark encoding phase is to embed watermark information in such a way that it does not affect the quality of data.The watermark encoding process uses the watermark bit generated from data and time.

(4) Watermark Decoding

The watermark decoding is performed on a single at a time by watermark decoder [2].In watermark decoding phase all the pre-processing phases done for encoding are per-formed on the watermarked data.

(5) Data Recovery-:

In the data recovery phase, data owner get original data from the attacked watermarked data. The main responsibility of post-processing is to use the decoded watermark bits, and convert these bits into the watermark information that was embedded as the watermark.

III. RESULT ANALYSIS

The proposed system is tested against the previous system based on various parameters like standard deviation, mean, variance etc. This study shows the measurable changes between the two systems. In the graph given below shows the comparison.

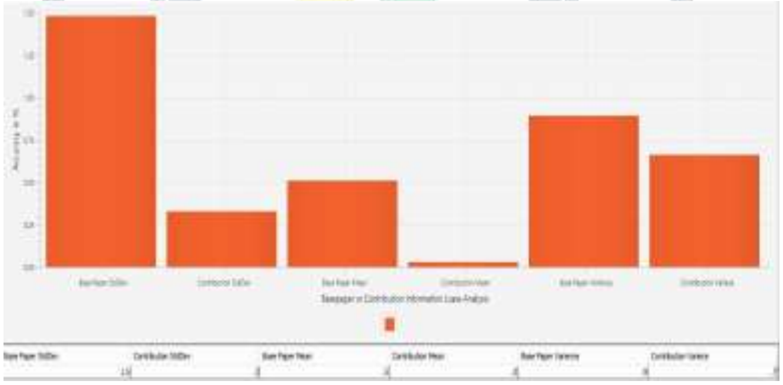

Fig. 2: Information Loss

- **Result Analysis of Existing System**

The Mutual Information between the two attributes before watermark and after water-mark is plotted in the form of graph. The graph shows that the co-relation between two attributes before watermark is greater than the co-relation between two attributes after watermark which means that the data is lost after watermark encoding.

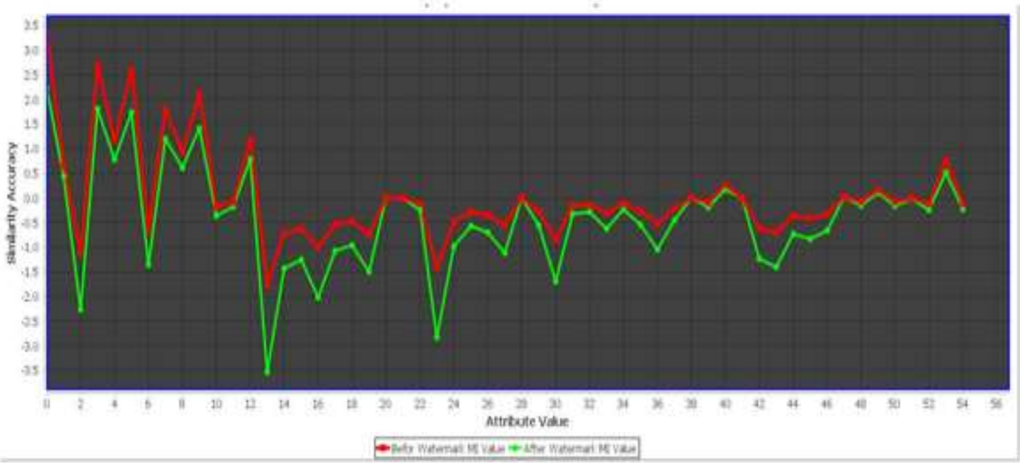

Fig. 3: Result Analysis for MI (Existing System)

➢ **Result Analysis of Proposed System**

The Mutual Information between the two attributes before watermark and after water-mark is plotted in the form of graph. The graph shows that the co-relation between two attributes before watermark and after watermark is not drastically changed, which means that there is minimum data loss after watermark encoding.

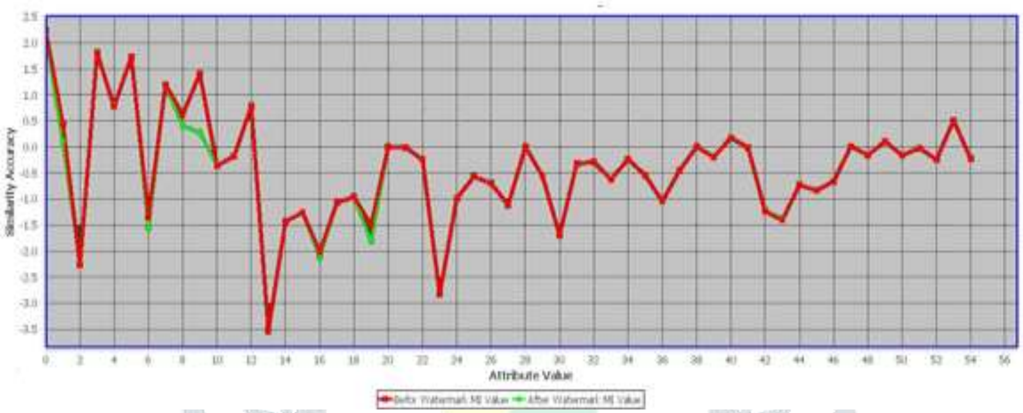

Fig 4: Result Analysis for MI

➢ **Watermark Decoding Accuracy Vs Deletion Attack**

The watermark decoding accuracy is tested on the dataset by deletion of tuples. The following graph shows the decoding accuracy in percentage.

Table 1: Watermark Decoding Accuracy Vs Deletion Attack (in Percentage)

Deleted Tuples	Decoding Accuracy
10	100
30	100
60	100
70	100
80	90
85	80
87	75
90	70

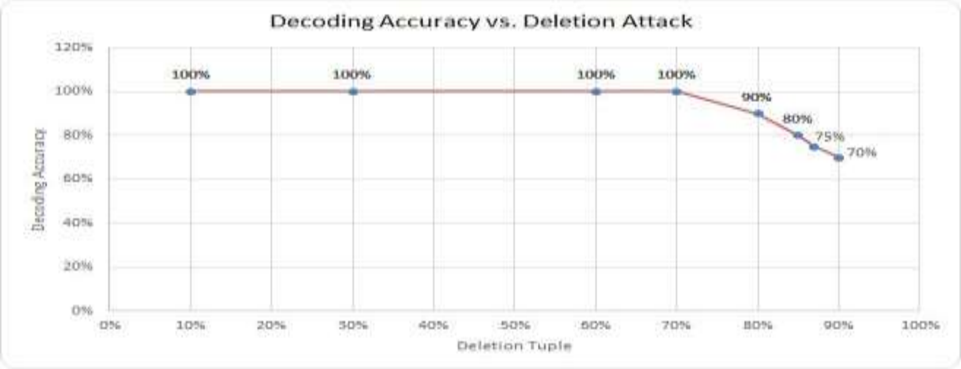

Fig. 5. Watermark Decoding Accuracy Vs Deletion Attack

- **Watermark Decoding Accuracy Vs Insertion Attack**

By inserting few into the original dataset the decoding accuracy of the proposed system is tested. The graph shows the decoding accuracy against the insertion attack.

Table 2: Watermark Decoding Accuracy Vs Insertion Attack (in Percentage)

Inserted Tuples	Decoding Accuracy
10	100
20	100
30	100
40	100
50	100
60	100
70	100
80	100
90	100
100	100

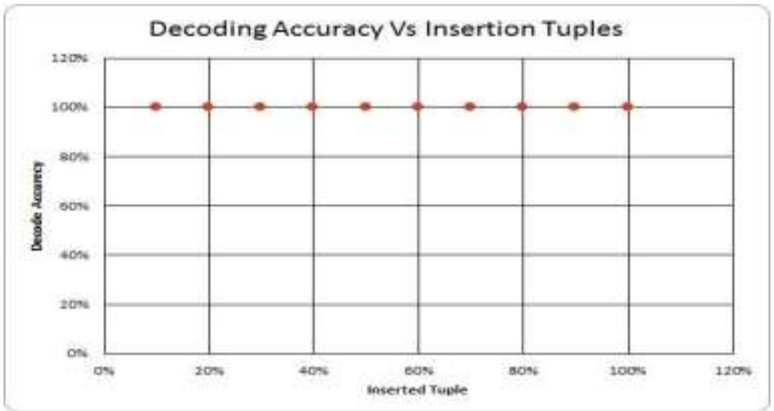

Fig. 6: Watermark Decoding Accuracy Vs Insertion Attack

- Advantages

1. Strong Technique of Numerical data Watermarking

2. Minimum Data Distortion

3. Original data can be recovered

4. Minimum Information Loss

IV. CONCLUSISON AND FUTURE WORK

In this paper, we proposed a multi-attribute, reversible and scalable watermarking technique for numerical data of relational databases. The advantage of this technique is that, if data is infected from to malicious attacks, our technique will allow recovery of a large portion of the data not only embedded watermark but also the original data. The experimental results against various types of attacks like deletion attack, insertion attack are tested. For deletion attack the decoding accuracy is 100 percent for 80 percent tuple deletion. Similarly for insertion attack 100 percent decoding accuracy is achieved. The system experientially proved better against existing system for the data loss issue. The difference between the MI values before watermarking and after watermarking is negligible. Hence we are succeeded to keep data more usable after watermarking. This research solely concentrates on using unsigned numeric data as a input. So, In future, we will extend this research towards signed numeric data. Future work also includes extending our embedding scheme so that it marks both non-numeric and numeric attributes.

REFERENCES

[1] M.E. Farfour, S.-J. Horng, J.-L.Lai, R.-S Run, R.-JChen, and M. K. Khan, “A blind reversible method for watermarking relational databases based on timestamping protocol,” Expert Syst. Appl., vol. 39, no. 3, pp. 3185–3196, 2012.

[2] SamanIftikhar, M. Kamran, and Zahid Anwar, ―RRW—A Robust and Reversible Watermarking Technique for Relational Data,‖ In IEEE Trans. on knowledge and Data Engineering, VOL. 27, NO. 4, APRIL 2015.

[3] R.Agrawal and J. Kiernan, Watermarking relational databases, in Proc. 28th Int. Conf. Very Large Data Bases, 2002, pp. 155166

[4] Saman Iftikhar, M. Kamran, and Zahid Anwar, "RRWA Robust and Reversible Watermarking Technique for Relational Data", In IEEE Trans. on knowledge and Data Engineering, VOL. 27, NO. 4 , APRIL 2015.

[5] J. T. Brassil, S. Low, and N. F. Maxemchuk, "Copyright protection for the elec-tronic distribution of text documents", In Proc. IEEE, vol. 87, no. 7, pp. 11811196, Jul. 1999.

[6] F. A. Petitcolas, Watermarking schemes evaluation", IEEE Signal Process.Mag., vol. 17, no. 5, pp. 5864 , Sep. 2000.

[7] A. M. Alattar, Reversible watermark using di erence expansion of triplets", In Proc. IEEE Int. Conf. Image Process., pp. I501, vol. 1, 2003.

[8] D. M. Thodi and J. J. Rodriguez, Prediction-error based reversible watermarking", in Proc. IEEE Int. Conf. Image Process. vol. 3, pp. 15491552, 2004.

[9] R. Sion, M. Atallah, and S. Prabhakar, "Rights protection for categorical data", In IEEE Trans. Knowl. Data Eng., vol. 17, no. 7, pp. 912 926, Jul. 2005.

[10] S. Subramanya and B. K. Yi, Digital rights management, In IEEE Potentials,vol. 25, no. 2, pp. 3134, Mar.-Apr. 2006.

[11] Y. Zhang, B. Yang, and X.-M. Niu, "Reversible watermarking for relational database authentication", J. Comput., vol. 17, no. 2, pp. 5966, 2006.

[12] D. M. Thodi and J. J. Rodriguez, Expansion embedding techniques for reversible watermarking, In IEEE Trans. Image Process., vol. 16, no. 3, pp. 721730, Feb. 2007.

[13] G. Gupta and J. Pieprzyk, "Reversible and blind database watermarking using difference expansion", In Proc. 1st Int. Conf. Forensic Appl. Tech. Telecom-mun., Inf., Multimedia Workshop, p. 24, 2008.

[14] 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., pp. 563569, 2010.

[15] X. Li, B. Yang, and T. Zeng, E cient reversible watermarking based on adaptive prediction-error expansion and pixel selection, In IEEE Trans. Image Process.,vol. 20, no. 12, pp. 3524 3533, Dec. 2011.

[16] M. Kamran and M. Farooq, An information-preserving watermarking scheme for right protection of EMR systems, In IEEE Trans. Knowl. Data Eng., vol.24, no. 11, pp. 19501962, Nov. 2012.

[17] J.-N. Chang and H.-C. Wu, Reversible fragile database watermarking technology using di erence expansion based on SVR prediction, in In Proc. IEEE Int.Symp. Comput., Consum. Control, pp. 690693., 2012.

[18] 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. 31853196, 2012.

[19] E. Sonnleitner, "A robust watermarking approach for large databases", In Proc.IEEE First AESS Eur. Conf. Satellite Telecommun., pp. 16, 2012.

[20] K. Jawad and A. Khan, "Genetic algorithm and di erence expansion based reversible watermarking for relational databases", J. Syst. Softw., vol. 86, no. 11, pp.27422753, 2013.

[21] M. Kamran, Sabah Suhail, and Muddassar Farooq, "A Robust, Distortion Minimiz-ing Technique for Watermarking Relational Databases Using Once-for- All Usability Constraints" In IEEE Transactions on Knowledge and Data Engineer-ing, VOL. 25, NO. 12, DECEMBER 2013.