ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

A Survey on Applying Metric Preserving Transformation of Robust Watermarking for Relational Database

T. Jaison Vimalraj, S. Praveena, S. Vishnupriya

Teaching Fellow, Department of IT, Anna University (BIT Campus), Trichy, Tamilnadu, India

UG Student, Department of IT, Anna University (BIT Campus), Trichy, Tamilnadu, India

UG Student, Department of IT, Anna University (BIT Campus), Trichy, Tamilnadu, India

ABSTRACT: In this paper, we show another powerful database fragile watermarking scheme, the innovation of which stands on a semantic control of the information bending and on the expansion of quantization Index Modulation (QIM) to round histograms of numerical qualities. With the arrival of Index based Hashing technique, it is easier to retrieve data. Thereafter prioritize it by Metric Preserving Transformation. The semantic contortion control of the installing process we propose depends on the ID of existing semantic connections in the middle of estimations of properties in a tuple by method for a meta-physics. Thusly, we maintain a strategic distance from indistinguishable or extremely uncommon record events which might inclination information translation on the other hand sell out the vicinity of the watermark. In a brief moment time, we adjust QIM to database watermarking. Watermark implanting is directed by adjusting the relative rakish position of the round histogram focal point of mass of one numerical characteristic. We hypothetically exhibit the power execution of our plan against most regular assaults (i.e., tuple insertion and erasure). This makes it suitable for copyright insurance, proprietor recognizable proof, or deceiver following purposes. We promote check tentatively these hypothetical points of confinement inside of the system of candidate passport details. Under the supposition forced by the focal cut off hypothesis, exploratory results fit the hypothesis. We too contrast our methodology and two proficient plans to demonstrate its advantages.

KEYWORDS: Robust Watermarking, QIM, Index based hashing, Metric preserving Transformation

I. INTRODUCTION

The most recent couple of years have seen a noteworthy increment in the development, exchange and sharing of databases. This is predominantly because of the fortification of their conservative quality furthermore, decisional interest, the last being connected to some degree to the advancement of information mining and examination instruments. In any case, these new get to abilities incite in the meantime security dangers, as information records might be redistributed or altered without consent. A few samples of data releases seem each year, even in touchy ranges like resistance [1] or social insurance [2]. Secure access and secrecy of information are more often than not accomplished by method for cryptographic systems. In any case, once these components circumvent or all the more basically when the entrance is in truth, information are no more ensured. Here comes the enthusiasm for watermarking, an a posteriori assurance, that leaves access to information while keeping up them ensured in terms of respectability or traceability as case. Watermarking lies in the insertion of a message (some security qualities) or a watermark into a host report (e.g., picture or database) by marginally annoying host information. All the more unequivocally, the insertion procedure depends on the rule of controlled bending of host information.

Watermarking has been effectively connected in mixed media assurance [3]–[5], however database watermarking was just presented in 2002 by Agrawal and Kiernan [6]. Since at that point, a few techniques have been proposed [7]–[9]. Contingent upon the implanting adjustment, we can recognize "property mutilation free" techniques, that don't change properties values from "property bending based" techniques. The previous are normally in light of the balance of the request of tuples inside of a connection [10]. On the off chance that one might consider that no information irritation has been presented, such a system makes the watermark reliant on the way the database is put away, inciting

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

imperatives on the database administration framework. As an outcome, the application extend this group of strategies can be utilized for is constrained. Besides, these strategies are delicate as any reordering of tuples will dispense with the watermark.

As a rule, watermarked tuples must remain semantically cognizant so as to: i) guarantee the right translation of the data without presenting inconceivable or improbable records; ii) keep the presented annoyances imperceptible to aggressors. In reality, an "unimaginable" tuple can be measurably immaterial yet exceptionally semantically perceivable [15]. To do as such, we propose another semantic twisting control strategy which exploits a meta-physics over the database plan. As uncovered by Gomez-Perez and Benjamins [16], ontologies give a regular vocabulary of a range and characterize, with various levels of convention, the importance of the terms and the relations between them. Ontologies have been effectively connected in a few spaces from information extraction [17] to picture explanation what's more, recovery [18]. As far as anyone is concerned, they have not been yet connected to control watermarking twisting. As we will appear, one meta-physics gives semantic learning or depiction of the database that can help us to recognize the reasonable property bending in a tuple. Up to now, diverse tweaks have been considered with a specific end goal to install a message into a numerical trait in a database

II. RELATED WORK

For the inferior of techniques, it is by and large accepted that the typical understanding of information won't be annoyed if some change (e.g., adjustment of qualities [11]) is completed in the database for message insertion. By the by, with a specific end goal to consider watermark subtlety, latest "bending based" plans consider mutilation requirements. Case in point, in [7] the implanting process does not adjust numerical qualities if a few "information ease of use conditions", measured as far as the mean squared mistake, are definitely not regarded. Shehab et al. consider extra quality insights imperatives (e.g., mean, standard deviation) on quality qualities furthermore, adjust the watermark plentifulness by method for advancement systems [9].terms of numerical attributes' mean and standard deviation variations defined by the data owner and recipients. The more restrictive set of variations constitute the "once for all" constraints. They then optimize their detection based on these constraints. If a recipient has lower distortion constraints, he will receive a more distorted database leading to a more robust watermark. In [14], Lafaye *et al.* consider a query result approach and look at preserving the response to *a priori* known queries of aggregation, and modulate pairs of tuples in consequence. As exposed, the above methods focus on preserving the database statistics (of attributes [9], [13] or in-between attributes [12]) and do not take into account the full database semantics that should also be preserved. Semantics refer to the meaning of a piece of information. Although statistics may provide hints about the existence of such semantic links, as they evaluate the dependencies or the co-occurrences of values in the database, they do not allow directly identifying such a situation.

III. PROPOSED ALGORITHM

A. *Design Considerations:*

- Data Acquisition
- Hiding the data into the image
- Generating unique ID or key
- Storing data in the database in encrypted form.
- Providing authentication

B. *Description of the Proposed Algorithm:*

Aim of the proposed algorithm is to prevent access of sensitive data from unauthorized users.The proposed algorithm is consists of three main steps.

Step 1: Encrypting the data

QIM depends on the quantization of the components (tests, gathering of tests or change coefficients) of a host signal as indicated by an arrangement of quantizers in view of codebooks all together to implant the images of a message. All the more obviously, to each image si issued from a limited set .S=$\{s_u^i\}$ where u=0,…,U such that

$$Cs_u^i \cap Cs_v^i = \emptyset \; if \; u \neq v \qquad \text{eq. (1)}$$

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

the QIM partners a codebook { Cs_u^i}u=0,...,U such that CsWith a specific end goal to insert the image s_u^i into one component X of the signal, this one is supplanted by XW which compares to the closest component of X in the codebook Csiu .This procedure can be seen as:

$$X_W = Q(X, Cs_u^i) \quad (3) \qquad \text{eq. (2)}$$

where the capacity Q gives back the closest component to X in cs_u^iNotice that the watermarking twisting compares to the separation in the middle of X and XW. To epitomize this procedure, let us think of one as pixel X of a picture, which might take its qualities from a one-dimensional space [0, 255]. We isolate this scalar space into non covering cells or interims of break even with size. Every cell is then identified with one and only codebook {cs_u^i } u=0,..., U in order to fulfill (2). Subsequently, an image s_u^ihas a few representations in [0, 255] and Q compares to a scalar quantizer. In the insertion process, if X has a place with a cell that encodes the coveted image s_u^i , its watermarked adaptation XW compares to the centroid of this cell. Something else, X is supplanted by the centroid of the closest cell encoding s_u^i .In the extraction, the information of the cell to which XW has a place is sufficient to recognize the implanted image. This procedure is delineated in Fig. 5 on account of a parallel message,

i.e., $s_u^i \in 0, 1$ and two codebooks C0 and C1 for which the cells are characterized by uniform scalar quantization of quantization step ρ. In this illustration, X will be quantized to the closest square or hover keeping in mind the end goal to encode s_u^i.An augmentation of this approach, whose reason for existing is to diminish the contortion, is the Compensated QIM [22] where a part of the quantization blunder is added back to the quantized esteem in order to better deal with the watermark heartiness/in distinctness trade off.

Step 2: Hiding the data

A database DB is usually characterized as a limited arrangement of relations { R_i }i=1,...,NR . In this work, for purpose of effortlessness, we consider a DB with one single connection constituted of N unordered tuples { t_u}u=1,...,N , each of M qualities {A1, A2, . . . , AM}. A quality A takes its qualities inside a trait area and t_u.An alludes to the estimation of the nth trait of the uth tuple. Each tuple is particularly recognized by either one property or an arrangement of properties, we call its essential keyt_u.PK. Two major stages are considered in the greater part of database watermarking plans: message implanting and message identification/extraction. As portrayed in Fig. 1, the installing stage incorporates a pre processing process, the target of which is to make the watermark insertion/perusing free of the way database is put away. It normally comprises in the development of gatherings of tuples, making an arrangement of Ng non-converging gatherings of tuples {G_i} i=1,..., N_g. Ordinarily, the gathering number for one tuple n_u is gotten from the consequence of a cryptographic hash capacity connected to its essential key t_u.PK, connected with a mystery watermarking key, K_s for example, [9]:

$$n_u = H(K_s|H(K_s|t_u.PK)) \bmod N_g \qquad \text{eq. (3)}$$

where "|" speaks to the connection administrator and Ng is the quantity of gatherings to construct. The utilization of a cryptographic hash capacity, e.g., Secure Hash Algorithm (SHA), guarantees the protected and equivalent conveyance of tuples into gatherings.

Step 3:Retrieving or accessing the data

It is necessary to access data in the easier manner.And also it is in need to provide the authorized access.To minimize the time complexity,Index based hashing technique can be used.In this technique,assign the primary key as index.For that,The similarity of grouping must be carried out.It can be possible with the help of metric preserving transformation.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

X and Y be metric spaces , d_X and d_Y are named as metrics. A map function f such that $f : X \rightarrow Y$ is called an isometry or distance preserving if for any $a,b \in X$ one has

$$d_y(f(a),f(b)) = d_x(a,b) \qquad \text{eq. (4)}$$

IV. PSEUDO CODE

Step 1: Collect the data.
Step 2: Associating the finite sets of symbols with codebook using eq. (1).
Step 3: Replace the element X by X_W using eq. (2)
Step 4: Embed the message .
Step 5: Assign the primary key as t_u.PK which combined with secret watermarking key K_s and then evaluate it using eq. (3).
Step 6: Retrieve the data by similarity grouping
Step 7: if request == accept
Access the content
else
Authentication error
Step 8: End.

V. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a new method to secure the data from unauthorized access by using robust database watermarking which based on semantic distortion control method and QIM with index based hashing technique that can be clustered by metric preserving transformation. As we have shown, semantic distortion control aims at two main objectives: i) ensure the correct interpretation. ii) Protect from hacker. Also the main drawbacks of the existing system are time complexity and space complexity. To avoid such complexities, creating key as index and searching it with that key. Hence consumption of time can be minimised. Thereafter without annoying the server, creating a separate cache to store data and also the data can be secured.

REFERENCES

[1] S. Rogers. *WikiLeaks Embassy Cables: Download the Key Data and See How it Breaks Down*. [Online]. Available http://www.theguardian.com/news/datablog/2010/nov/29/wikileaks-cables-data,accessed Nov. 21, 2013.
[2] M. McNickle. *Top 10 Data Security Breaches in 2012*. Healthcare Finance News. [Online]. Available: http://www.healthcarefinancenews.com/news/top-10-data-security-breaches-2012, accessed Nov. 21, 2013.
[3] W.-H. Lin, Y.-R. Wang, and S.-J. Horng, "A wavelet-tree-based watermarking method using distance vector of binary cluster," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9869–9878, 2009.
[4] D. Rosiyadi, S.-J. Horng, P. Fan, X. Wang, M. K. Khan, and Y. Pan, "Copyright protection for e-government document images," *IEEE Multimedia*, vol. 19, no. 3, pp. 62–73, Jul./Sep. 2012.
[5] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2nd ed. San Mateo, CA, USA: Morgan Kaufmann, 2008.
[6] R. Agrawal and J. Kiernan, "Watermarking relational databases," in *Proc. 28th Int. Conf. Very Large Data Bases*, 2002, pp. 155–166.
[7] R. Sion, M. J. Atallah, and S. Prabhakar, "Rights protection for relational data," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 12, pp. 1509–1525,Dec. 2004.
[8] D. Gross-Amblard, "Query-preserving watermarking of relational databases and XML documents," *ACM Trans. Database Syst.*, vol. 36, no. 1, 2011, Art. ID 3.
[9] M. Shehab, E. Bertino, and A. Ghafoor, "Watermarking relational databases using optimization-based techniques," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 1, pp. 116–129, Jan. 2008.
[10] Y. Li, H. Guo, and S. Jajodia, "Tamper detection and localization for categorical data using fragile watermarks," in *Proc. 4th ACM Workshop DRM*, 2004, pp. 73–82.
[11] F. Guo, J. Wang, and D. Li, "Fingerprinting relational databases," in*Proc. ACM Symp. Appl. Comput.*, 2006, pp. 487–492.
[12] M. Kamran and M. Farooq, "A formal usability constraints model for watermarking of outsourced datasets," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 1061–1072, Jun. 2013.
[13] M. Kamran, S. Suhail, and M. Farooq, "A robust, distortion minimizing technique for watermarking relational databases using once-for-all usability constraints," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 12, pp. 2694–2707, Dec. 2013.
[14] J. Lafaye, D. Gross-Amblard, C. Constantin, and M. Guerrouani, "Watermill: An optimized fingerprinting system for databases under constraints," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 4, pp. 532–546, Apr. 2008.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

[15] J. Franco-Contreras *et al.*, "Data quality evaluation in medical database watermarking," *Studies Health Technol. Inform.*, vol. 210, pp. 276–280,May 2015.
[16] A. Gomez-Perez and V. R. Benjamins, "Applications of ontologies and problem-solving methods," *AI Mag.*, vol. 20, no. 1, pp. 119–123, 1999.
[17] W. Su, J. Wang, and F. H. Lochovsky, "ODE: Ontology-assisted data extraction," *ACM Trans. Database Syst.*, vol. 34, no. 2, 2009, Art. ID 12.
[18] L. Hollink, G. Schreiber, J. Wielemaker, and B. Wielinga, "Semantic annotation of image collections," in *Proc. Workshop Knowl. Markup Semantic Annotation (KCAP)*, 2003, pp. 41–48.

BIOGRAPHY

Mr T Jaison Vimalraj is a Teaching fellow in the Information Technology Department, University College of Engineering, BIT – campus, Tiruchirappalli, Tamilnadu, India.

S Praveena is a student in the Information Technology department, University College of Engineering, BIT – campus, Tiruchirappalli, Tamilnadu, India.
.
S Vishnupriya is a student in the Information Technology department, University College of Engineering, BIT – campus, Tiruchirappalli, Tamilnadu, India.
.