Volume 6, Issue 4, April 2016 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Survey of Ownership Protection for Relational Database based Watermarking

Arti Rajguru*

Department of computer Engineering & Savitribai Phule, Pune University,
Pune, Maharashtar, India

Abstract— Ownership protection on relational database that requires development of watermarking schemes that satisfies the following challenges: a] Robustness against different types of attacks b] Ability to conserve the knowledge in the databases resulting in an effective part of knowledge aware decision support system c] Maintaining the Balance between the inconsistent requirement of database owners and database recipients d] Ensurement of minimum Information loss in the data e] Usability restrictions for functions and recipients

Keywords— usability constraints, distortion free, database watermarking, ownership protection, robust watermarking

I. INTRODUCTION

The piracy of digital assets such as software, images, video, audio and text has long been unrest for owners of these assets security schemes are usually based upon the insertion of digital watermarks into the data. The watermarking software introduces small faults into the object being watermarked. These intentional errors are called marks and all the group of marks collectively which is called watermark. The marks are chosen so as to have an unrelated impact on the convenience of the data and are placed in such a way that a spiteful user cannot destroy them without making the data significantly less useful. Although watermarking does not prevent prohibited copying, it prevents such copying by providing a means for establishing the original ownership of a redistributed copy. The increasing use of databases in applications beyond at the back of the firewalls data processing is creating a similar need for watermarking databases. The Internet is exerting unbelievable pressure on data providers to create services that allow users to search and access databases remotely. Although this trend is a benefit to end users, it depicted the data providers to the danger of data theft. suppliers are insisting technology for identifying pirated copies[3].

Watermarking without any exception, has been used for ownership protection of a number of data arrange in images, video, audio, software, 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 sense interesting patterns that provide significant support to decision makers in making well-organized, exact, and pertinent decisions which produce the result of sharing of data between its owners and data mining experts is significantly increasing. as a result it has become pertinent to explore appropriate watermarking techniques for ownership rights protection of relational databases that should be imperceptible, robust with blind decoding. Moreover, once the owner of data embeds the watermark, the distortions in the original data are reserved within certain boundary which are defined by the usability restrictions, to protect the knowledge contained in the data.[1]

Innovative watermark decoding algorithm make sure that its decoding accuracy is independent of the usability constraints Our approach make possible Alice to define usability constraints only once for a particular database for every possible type of suggested application. it also take care that the watermark set ups the smallest amount possible distortions to the original data without losing the robustness of the inserted watermark.[1]

Attack Analysis

The watermarked database may suffer from various types of intentional and unintentional attacks which may damage or erase the watermark:

Deletion Attack

In this attack Attacker eliminated the selected tuples form watermark dataset to remove the watermark .He can also arbitrarily delete the tuples or selects them in a complicated manner on the basis of their statistical allocation of attribute values

Insertion Attack

There are two types of insertion attack 1)fixed insertion and 2) constraint reliant insertion. In the first attack, attacker inserts new n tuples by duplicating values of existing g tuples. In the second attack, he makes the Tuple values based on the mean µ and the standard deviation σ of watermarked data set.

Alteration Attack

This attack try to destroy the watermark by altering one or more bits in the watermarked data. More additional information about the marked bit position makes attack more triumphant

II. RELATED WORK

Agrawal and Kiernan [3] projected a bit-resetting algorithm in which it sets the LSB of the candidate attribute of the selected subset of tuples. The parameters selection for watermarking is based on calculating message authenticated code (MAC), where MAC is calculated using the secret key and the tuples primary key. This technique assumes unconstrained LSB manipulation during watermark embedding process .Such out-of-bound modification of data might also generate unwanted results. Although LSB-based data hiding techniques are efficient, but disadvantage produce with this technique is an attacker is easily eliminate watermark by simple Manipulation of data by using bit-resetting techniques

Sion et al [10] projected a statistical-based algorithm in which a database is divided into a maximum number of unique, nonintersecting subsets of tuples. The data partitioning concept is based on the use of special marker tuples, making it disposed to watermark synchronization errors particularly in the case of tuples insertion and deletion attacks, as the position of marker tuples is troubled by these attacks. Such errors may be reduced if marker tuples are stored during watermark embedding phase and it can be used for creating the data partitions once more during watermark decoding phase. Disadvantage produce with this technique is that using the stored marker tuples to reconstruct the partitions violates the requirement of blind decoding of watermark

Another class of watermarking techniques is distortion free techniques S. Bhattacharya, A. Cortesi Using these techniques, data is delivered to the future recipients without making any distortion in the data. The techniques reported in [10] and [11] are vulnerable to even minor malicious attacks, and therefore, cannot be used for enforcing ownership protection.

A recent survey R. Halder, S. Pal, and A. Cortesi presented a review of Ownership protection of Relational database based on Watermarking .it ensures that the decoding accuracy is independent of the usability constraints, an vendor does not need to define them for each type of future application and use. Our proposed technique not only has this feature but it is also able to gather the difference robustness requirement of the data owner and minimum information loss requirement of the intentional recipient

Content characteristics as watermark information [5] In these techniques, the extracted information from the database contents are treated as watermark and it can be embedded in the data in the same database. But since the content keeps changing under various operations in Q, the modification of the content may lead to the extracted information differently in the verification phase from that in embedding phase. We can use this method in order to take out only the invariant properties and embed it into the stable part of the database content.

III. PROPOSED WORK

In proposed work the algorithm proceeds as follows-

1. We make the partitions so that all tuples are divided into proper cluster.
2. Calculating Threshold values.
3. Select only those rows whose hash values are Even.
4. Embed the watermark
5. Decode the watermark.

1. Algorithm for Data partitioning

In this partitioning algorithm it divides the dataset D into m non overlapping partitions that is S0 to Sm-1 such that for any two partitions Si and sj is not null. This algorithm partitions the dataset into logical groups. Partitioning is based on a secret key Ks and a cryptographic hash function Message Digest (MD5)

```
for each Tuple r∈D do
    par(r) = H(Ks || H(r.PK ||Ks)) mod m
    insert r into Spar(r)
end for
return S0,....,Sm-1.........[1]
```

2. Algorithm for Data Selection Threshold

In this step a threshold calculated for each attribute. If the value of any attribute of a tuples is Above its respective calculated threshold select only those tuples for watermarking. The confidence factor c is kept secret to make it very hard for an attacker to estimate the selected tuples in which the watermark is inserted.

```
for i = 0 to m - 1 do
    for each Attribute A∈Si do
        compute μ and σ on A
        calculate Threshold
    end for
end for
return D`T ←     for  all R >T ..........................[1]
```

3. Algorithm for Hash value calculation

In this step, a cryptographic hash function MD5 is applied on the selected data set it select only those tuples whose hash values are even This step achieves two objectives a) it further improves the watermark security by hiding the identity of the watermarked tuples from an intruder b) it further reduces the number of to be watermarked tuples to bound distortions in the data set.

```
For each r∈D`T do
    Even Value(r) = H(Ks || r.PK) mod 2
    if Even Value(r)==0 then
        insert r into D"T
    else
        don't consider this tuples for watermarking
return D"T ............ [1]
```

4. Algorithm for Embed watermark

The watermarking algorithm makes the use of multiple bit at a time property and makes more powerful to any number of attributes .Watermark embedding function changes the input data D to a watermarked data DW after performing some data exploitations

```
DW = D
D"T = Get Even Hash Value Data Set(D`T ,Ks)
for each row r in D"T do
    temp = r.PK
    if b == 1 then
        compute ƌ using (15) subject to G
    else
        compute ƌ using (16) subject to G
    end if
    DWtemp =( ƌ +DWtemp)
end for
insert ƌ into Δ
Compute ٢
return Dw , ٢…………..[1]
```

5. Algorithm for Detect Watermark

The watermark decoding algorithm take outs the inserted watermark using the undisclosed parameters: Ks ,m, ٢ .watermark bits interpreted in the reverse order last bit interpreted first .

```
ones = 0
zeros = 0
SW0,....., SWm-1 = Get Partitions (DW, Ks, m)
for each partition SWi do
    D`WT= Get Data Selection Threshold(RW>T ,c)
    D"WT= Get Even Hash Value Data Set(D'WT,Ks)
    for each row r in D"WT do
        ϒ=Val-٢
        if ϒ≥ 0 then
            ones[i]= ones[i]+ 1
        else
            zeros[i]= zeros[i] + 1
        end if
    end for
    .if ones[j] > zeros[j] then
        Increment 1's counter
    else
        Increment 0's counter
    end if
end for
return WD ............[1]
```

IV. CONCLUSIONS

This paper presents survey on various Techniques used for watermarking The work proposes a robust watermarking technique for relational database that tries to meet the mentioned challenges. Watermark is embedded in the relational data in characters, integers and unsigned integer .To verify for robustness system will be tested for various attacks.

REFERENCES

[1] Kamran, M., Suhail, S.; Farooq, M., "A robust, distortion minimizing technique for watermarking relational databases using once-for-all usability constraints," IEEE Transactions on Knowledge and Data Engineering, *on* , vol.PP, no.99, pp.1-1, Dec 2013

[2] M. Shehab, E. Bertino, and A. Ghafoor, “Watermarking Relational Databases Using Optimization-Based Techniques,” IEEE Trans. Knowledge and Data Eng., vol. 20, no. 1, pp. 116-129, Jan. 2008.

[3] R. Agrawal and J. Kiernan, “Watermarking Relational Databases, Proc. 28th Int’l Conf. Very Large Data Bases, pp. 155-166, 2002.

[4] R. Agrawal, P. Haas, and J. Kiernan, “Watermarking Relational Data: Framework, Algorithms and Analysis,” The VLDB J., vol. 12, no. 2, pp. 157-169, 2003.

[5] H. Guo, Y. Li, A. Liu, and S. Jajodia, “A Fragile Watermarking Scheme for Detecting Malicious Modifications of Database Relations,” Information Sciences, vol. 176, no. 10, pp. 1350-1378, 2006.

[6] S. Bhattacharya and A. Cortesi, “A Distortion Free Watermark Framework for Relational Databases,” Proc. Fourth Int’l Conf. Software and Data Technologies (ICSOFT ’09), pp. 229-234, 2009.

[7] S. Shah, S. Xingming, H. Ali, and M. Abdul, “Query Preserving Relational Database Watermarking,” Informatica, An Int’l J. Computing and Informatics, vol. 35, no. 3, pp. 391-396, 2011.

[8] R. Halder and A. Cortesi, “A Persistent Public Watermarking of Relational Databases,” Proc. Int’l Conf. Information Systems Security, pp. 216-230, 2011.

[9] R. Halder, S. Pal, and A. Cortesi, “Watermarking Techniques for Relational Databases: Survey, Classification and Comparison,”J. Universal Computer Science, vol. 16, no. 21, pp. 3164-3190, 2010

[10] R. Sion, M. Atallah, and S. Prabhakar, “Rights Protection for Relational Data,” IEEE Trans. Knowledge and Data Eng., vol. 16, no. 6, pp. 1509-1525, Dec. 2004.

[11] Y. Li and R. Deng, “Publicly Verifiable Ownership Protection for Relational Databases,” Proc. ACM Symp. Information, Computer and Comm. Security, pp. 78-89, 2006.