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. 3, Issue 4, April 2015

Watermarking of Relational Databases Using Optimization Technique

Diksha Pande, Mallika Upadhyay, Shivam Pal

B.E Students, Department of Computer Engineering, Sinhgad Academy of Engineering Kondhwa , Pune, India

ABSTRACT: Watermarking of the databases is a technique where we add additional non-detectable information is added into the original data. The main aim of implementing watermarking is for ownership proof and to prevent our precious data from being tampered with by an anonymous entity. The accuracy of the item is slightly degraded but the watermark acts as a seal that henceforth identifies the owner of the software. In our paper we present an effective watermarking technique geared for watermarking of relational databases. In our watermarking technique, use of a secret key is done, which play's a very important role in protecting the owner's data from being tampered with. Only if one has access to this secret key can the watermark be detected with high probability. The watermark can be easily and efficiently be maintained using insertion, deletion and the update of the database .

KEYWORDS: Watermarking, secret key, insertion, detection.

I. INTRODUCTION

Watermarking is a technique which is used to deter data piracy and tamper-proof the data during it's transmission from one machine to another. There are realms of watermarking namely image watermarking, video watermarking and audio watermarking.
The watermarking of the relational databases is a rare and currently developing sector. But as the need of the database applications is increasing, it's exerting more pressure on the data providers to create services that allow the users to access and search their data remotely. Now this is a threat for the data providers which in turn leads to the increasing demand of the database watermarking to detect the pirated copy and protect it from being tampered with . One thing which shouldn't be assumed is that database watermarking techniques are similar to the watermarking techniques of multimedia objects. Watermarking of the database requires techniques which differ from those of the conventional ones used for multimedia watermarking purposes.

Multimedia objects cannot be dropped or replaced arbitrarily but in databases the tuple insertion, deletion and update are the main norms in database setting. Because of this, different techniques developed for multimedia data cannot be directly used for watermarking of the relational databases. Text properties and semantics can be exploited.

II. BACKGROUND

Literature Survey:

Watermark: Watermark describes information that can be used to prove the ownership of data such as the owner, origin of recipient of the content. Watermarking is the piece of data securely embedded and this is said to be imperceptible.
Ownership rights on outsourced relational databases is a very crucial issue in today's internet environment and in many content distribution applications, because the rapid growth of the internet and related technologies offer an unprecedented ability to access and redistribute digital content. In earlier existing systems the relational data would be watermarked and directly sent to the client system. In these systems, while sending relational data from server to client, an attacker can easily copy the data and create a copy of the relational data. Here, there is no security to the watermarked relational data. In our proposed system, before sending the watermarked relational data to client side, we encrypt the relational data and then send it to the client side. At client side decryption will be done to get the original watermarked data. Because of using this encryption technique even if an attacker does copy the data, he/she cannot

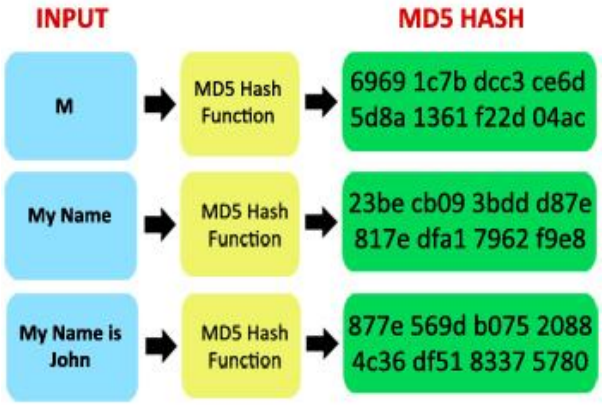
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

read the watermarked relational data.In this system, even if an attacker copies the watermarked relational data, it is not in human readable format.We are using MD5 algorithm for encrypting and decrypting the watermarked data. MD5 stands for Message Digest-5. MD5 is a cipher that encrypts and decrypts a data block of 128 bits. It uses 10, 12 or 14 rounds. The key size can be 128, 192, or 256 bits, depending on the number of rounds. So for secure watermarking of relational database we are using MD5 algorithm along with triple DES.



III. WATERMARKING PROPERTIES

Blind System-Watermark detection does not require the knowledge of the original data, only that of the secret key.
Incremental Updatability- If update of the database is required , re-computing watermark for every update is not required.
Imperceptibility- Presence of watermark is unnoticeable.
Robustness- Watermarking cannot be easily destroyed by modifying the watermarked data.
Detectability- Can be detected only with the knowledge of the secret key.

IV. WATERMARKING CHALLENGES

There are various ways in which our watermarked data can be attacked .

They are:

Bit Attacks: The simplest malicious attack attempts to destroy the watermark by altering the bits . The effectiveness of such attacks is sensitive to the number of bits altered. Example:- Suppose a hacker wants to hack a policyholder account he will alter every bit in the database , he has not only destroyed the watermark but also made the original data completely useless but if the location of the changed bits is known then the attack will prove to be very effective.

Rounding Attack: The hacker may try to loose the marks contained in the numeric attribute by rounding all the values of the attribute. For this attack to be fruitful the hacker needs to accurately guess the number of bit positions involved in the watermark. If the guess is less than required then he will not succeed, if he over guesses it he will degrade the quality of his data. Even if his guess is correct his data will be inferior to policyholder's data because his data values are comparatively less precise.

Transformations: An attack related to the rounding attack but in this the numeric values are linearly transformed. Example:-Hacker can convert the data to different unit of measurement(e.g.,kilometers to meters).The admin simply needs to convert the values back to the original system .
Subset Attack: Hacker may tak a subset of the tuples or attributes of a watermarked relation and hope that the watermark is lost.
Mix-and-match Attack: Hacker may ccreate his relatin by taking disjoint tuples from multiple relations containing similar set of information.

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

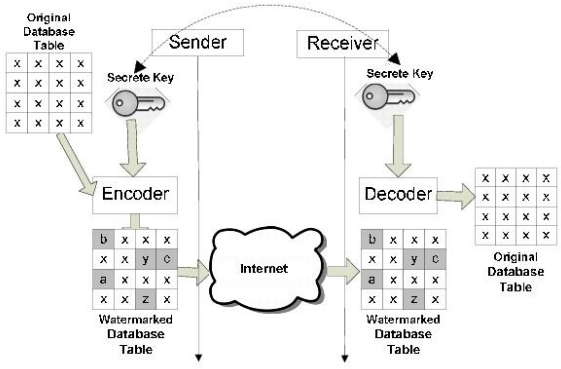
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

V. WORKING OF WATERMARKING MODEL

Let's consider an example to understand the working of our model. Consider a policy account wherein we have two actors. First, a policyholder, and second, the admin. The policyholder wants to update his account this means the policyholder wants access to the database of the bank where he has his account , for getting the access the policyholder needs to know the secret key for his name in the database provided by the admin . After he has entered his secret only he is able to access the database of his name only.



Watermarking Model

VI. ALGORITHM

We will be using partitioning algorithm for watermarking of our relational database.

A. Watermarking Insertion

Here, we determine if the tuple under consideration will be marked. Because of the use of MAC, only the owner who has the knowledge of the private key jcan easily determine which tuples have been marked. For a selected tuple, we determine the attribute that will be marked amongst the candidate attributes. Then we determine the bit position amongst _least significant bits that will be marked. Again, the results of the tests depend on the private key of the owner. For erasing a watermark, therefore, the attacker will have to guess not only the tuples, but also the marked attribute within a tuple as well as the bit position. The mark subroutine sets the selected bit to 0 or 1 depending on the hash value obtained . Thus, the result either leaves the attribute value unchanged or decrements (increments) it. Consequently, marking decrements some of the values of an attribute while it increments some others and leaves some unchanged. Databases usually allow attributes to assume null values. If a null attribute value is encountered while marking a tuple, we do not apply the mark to the null value, leaving it unchanged.

B. Watermarking Detection

Assume the admin suspects that the relation .published by the hacker has been pirated from his relation . We assume that the hacker does not drop the primary key attribute or change the value of primary keys since the primary key contains valuable information and changing it will render the database less useful from the user's point of view. The watermark detection algorithm, is probabilistic in nature. It determines if the tuple under consideration must have been marked at the time of inserting the watermark. Then it determine's the attribute and the bit position that must have been marked. The subroutine match then compares the current bit value with the value that must have been set for that bit by the watermarking algorithm.

We thus know how many tuples were tested (totalcount) and how many of them contain the expected bit value (matchcount). In a probabilistic framework, only a certain minimum number of tuples have to contain matching marked bits. The matchcount is compared with the minimum count returned by the threshold function for the test to succeed at

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. 3, Issue 4, April 2015

the chosen level of significance .. If admin finds a tuple in which he must have marked the attribute, but hacker has omitted, he simply ignores the tuple. Similarly, if a tuple is found whose attribute should have been marked, but .B. .E.has a null value, the tuple is ignored. I.e., The values of matchcount and totalcount are unaffected.

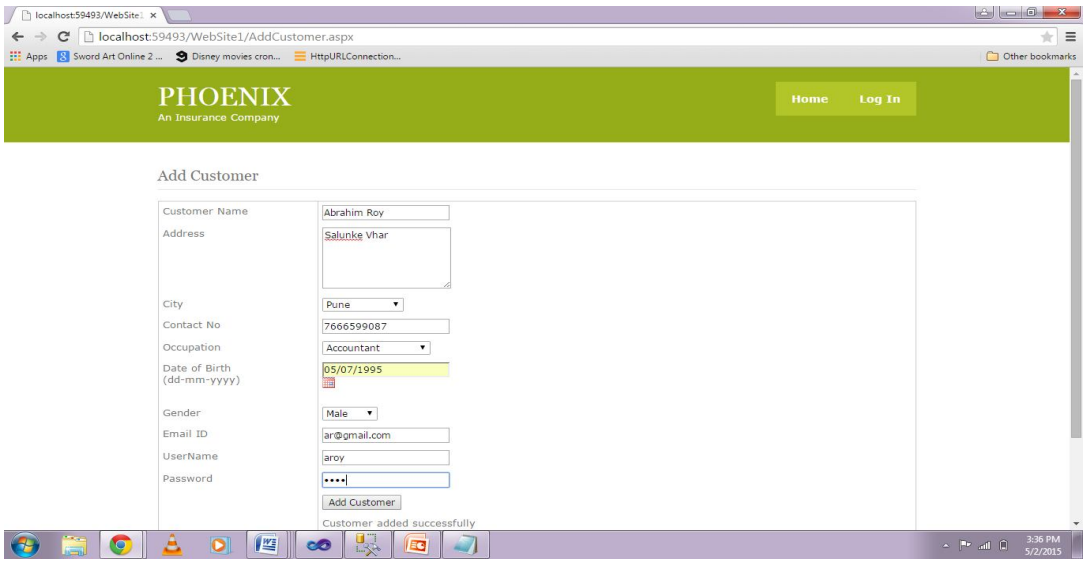
VI. MATHEMATICAL MODEL

Input : Data Set D, Secret Key Ks, Number of partitions m
1. {S0,S1……Sm-1} ß { }
2. for each tuple r € D
3. Partition(r) = H(Ks || H(r.p || Ks)) mod m.
4. insert r into Spartition(r)
5. return S0 , S1…… Sm-1
Output : Data Partitions S0,S1……Sm-1

VII. SIMULATION AND RESULTS

We will now show you the results of the working of the application.
Here at "Customer Enty", we will be entering the details to add a customer for a vehicle insurance policy system.
All validations are done.

Customer Entry



Here, in the "Watermarked Databse" section, we are showing how our data, entered into the database table(in the above image), is watermarked to ensure security of sensitive data.

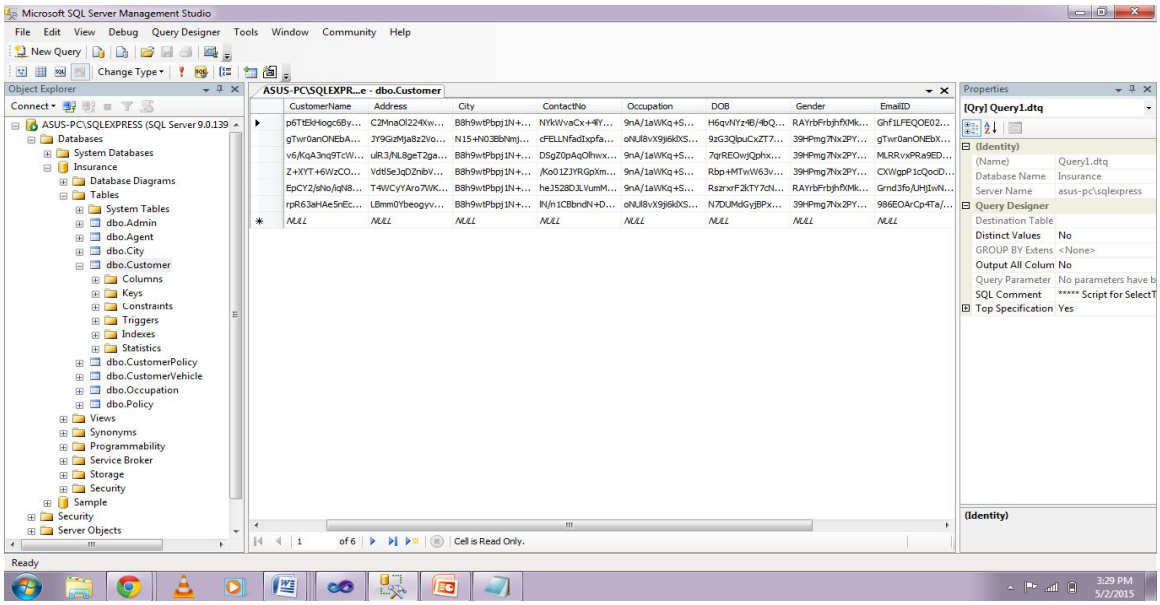
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Watermarked Database



VII. **CONCLUSION**

In this paper, we have made various contributions which include identification of the rights management of relational data through watermarking as an important and technically challenging problem for database research. We have also covered the articulation of the desirable properties of a watermarking system for relational data. Enunciation of the various forms of malicious attacks from which the watermark inserted in a relation must be protected has also been covered. Proposal of a watermarking technique specifically geared for relational data using portioning algorithm has been done.

VIII. **ACKNOWLEDGEMENT**

We wish to thank Mrs. Shalini Wankade for her experienced and valuable guidance at every step of the completion of this paper. We would also like to thank Mr. Ravi Sanap for her valuable input for the helping us in gathering the information.

REFERENCES

[1] N. R. Wagner. Fingerprinting. In IEEE Symp. on Security and Privacy, pages 18–22, Oakland, California, April 1983.
[2] B. Schneier. Applied Cryptography. John Wiley, second edition, 1996.
[3] J. Cox, M. L. Miller, “A review of watermarking and the importance of perceptual modelling,” In proc. of electronic imaging, 1997.
[4] R. Agrawal, J. Kiernan, “Watermarking relational databases,” Proceedings of the 28th International Conferenceson VLDB, pp. 155-166, 2002.
[5] R. Agrawal, J. Kiernan, “Watermarking relational data: Framework, Algorithms and Analysis,” VLDB Journal, pp. 155-166, 2003.
[6] ZHU Qin, YANG Ying, LE Jia-jin, LUO Yishu “Watermark based Copyright Protection of Outsourced Database,” IEEE, IDEAS, pp. 1-5, 2006.
[7] Sanjeev Khanna, Francis Zane, “Watermarking maps: hiding information in structured data,” Int’l Conf. SODA 2000, San Francisco, California, USA, pp. 596-605. 2000.
[8] Nagarjuna Settipalli, Prof. R. Manjula, “A new Relational Watermarking Scheme Resilient to Additive Attacks,” International Journal of Computer Application (0975-8887), Volume 10-No. 5, 2010.

BIOGRAPHY

Diksha Pande, Mallika Upadhyay and Shivam Pal are B.E students in the Computer Engineering Department of Sinhgad Academy Of Engineering, Kondhwa, Pune. Their research interests are Information Security, Watermarking Algorithms, Microsoft Visual Studio,etc.