INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

Robust and Reversible Relational Database Protection Using Watermarking Technique

TARUN DHAR DIWAN[1], VINAY SAHU [2]

[1]Dr.C.V.RAMAN UNIVERSITY, BILASPUR, INDIA
ASSISTANT PROFESSOR
taruncsit@gmail.com

[2]DR.C.V.RAMAN UNIVERSITY, BILASPUR, INDIA,
M Tech Scholar
vinay.sahu1111@gmail.com

Abstract: *This paper presents securing transmission of digital watermarking is applied broadly to relational database for ownership protection and information hiding. The process of watermark embedding and detecting, the databases tuples are adaptively clustered into groups according to the length of binary watermark. But vigrousness and reversibility are two main challenges due of the frequently database maintaining operators on these tuples. Digital watermarking is developed in recent years as a potential information security key technology, which can determine the legal control or originality of digital content by embedding observable or unobservable information in digital works. Database watermarking has been proposed on large database security-control. The watermarks are embedded into a relational database on the group basis under the control of a secure embedding key. The embedded watermarks form a watermark grid which can detect and localize any modifications made to the database and also be able to recover true data from modified cells. This technique promises that some bit positions of some of the attributes of some of the tuples contain specific values. The tuples attributes within a tuple, bit positions in a characterstic, and specific bit values in a robust watermarking scheme, the embedded watermark should be robust against attacks which aim at removing the watermark or making it undetectable. While in a fragile watermarking scheme, the embedded watermark should be broken to modifications so as to detect and localize or even recover the modifications. Most of the broken watermarking scheme studied in the last few years, were about multimedia watermarking. Most of them emphasis on digital images some have been extended to digital video, and audio data. It has better characteristics on security, invisibility and robustness.*

Keywords: *Security Control, Binary, Data, Transaction Number, Parser, Automatically, Hybrid method.*

1. INTRODUCTION

Digital watermarking is technology that lay machine-readable information within the content of a digital media file (image, audio, or video). The information is encoded through suitable image, audio, or video change. Much like watermarks on stationary, these changes typically would not be significant to any person viewing or listening to the content [1]. Indeed, digital watermarks often are not noticable by humans at all, but rather are designed to be detected and decoded only by machines specifically programmed to do so the general elements of a digital watermarking system are as follows.Embedding of watermark in content Each watermarking application initiates by placing a watermark into digital content. This involves changing the content using a unique algorithm.

The algorithm translates the data to be carried by the watermark into specific, suitable modifications to the content [2].Subsequent reading of watermark by device/software each watermarking application consists some ability for the embedded watermarks to be subsequently recognized. Recognizing the watermark needs knowledge of the algorithm used to embed it, because the reader device or software needs to know what changes to look for. Therefore, readers are system- or vendor specific; there is no reader capable

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

of recognizing and detecting all watermarks from all watermarking vendors. Back end database for determining meaning of watermark most watermarking applications involve maintaining a database for collecting and looking up data comprisis with specific watermarks. For example, the Information within a watermark itself might be easy serial number, while the database would enable that serial number to be interrelated with right information or a specific consumer. Similarly, the information in a watermark might consist of some type of coded message, needs accessable database to decode its meaning.

1.1 Actions triggered upon reading of watermark – In many watermarking applications, the recognition or reading of a watermark triggers or enables some type of action. Some actions may occurby there own, via appropriately programmed hardware or software that looks for watermarks and responds in known ways [3]. Other actions may depend on single decisions and responses of people to whom the information in the watermark has been convayed. Examples of actions that could be taken in response to reading a watermarkconsists of: Reporting or recording some information about how the watermarked media is being transmitted, accessed, or used. A watermark can be considered to be some kind of information that is embedded into underlying data for tamper detection, localization, ownership proof, and/or traitor tracing purposes. Database watermarking techniques complement the Database Protection Act and are becoming highly important as people realize that "the law does not provide sufficient protection to the comprehensive commercial and public oriented useful databases.

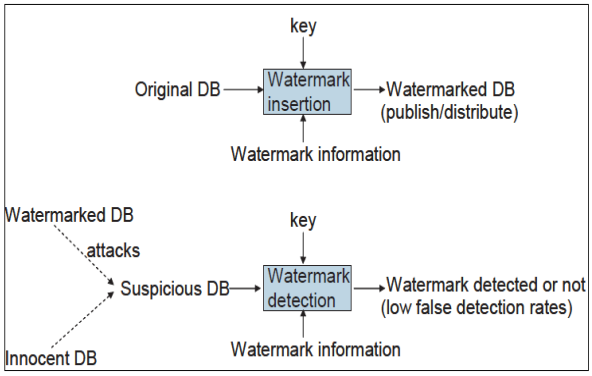

Figure 1: Basic Watermarking Process

While the basic processes in database watermarking are quite similar to those in watermarking multimedia data, the approaches developed for multimedia watermarking cannot be directly applied to databases because of the difference in data properties [4]. These differences include: A multimedia object consists of a large number of bits with considerable redundancy. Thus, the watermark has a large cover in which to hide. A database relation consists of tuples, each of which represents a separate object. The watermark needs to be spread over these separate objects. The relative spatial/temporal positioning of various pieces of a multimedia object typically does not change. Tuples of a relation, on the other hand, constitute a set, and there is no implied ordering between them. Multimedia objects typically remain intact; portions of an object cannot be dropped or replaced arbitrarily without causing perceptual changes in the object [5]. On the other hand, tuples insertions, deletions, and updates are the norm in the database setting.

2. RELATED WORK

This technique marks only numeric attributes with one-bit watermark scheme. The watermarking software represents only small errors into the object being watermarked. These known errors are called marks and all the marks together comprise the watermark. The marks must not have a significant impact on the utilitarian of the data and they should be placed in such a way that a evil user cannot destroy them without making the data less useful. Thus, watermarking does not prevent copying, but it deters illegal copying by giving a means for establishing the original ownership of a redistributed copy [6].The ascending use of databases in applications beyond "behind-the-firewalls data processing" is building a similar need for watermarking databases. For instance, in the semiconductor industry, parametric data on semiconductor parts is provided primarily by three companies: Aspect, IHS, and IC Master. They all employ a large number of people do manually; they extract part specifications from datasheets. Then they license these databases at high prices to design engineers. Companies like Acxiom have compiled large collections of consumer and business data. In the life

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

sciences industry, the primary assets of companies such as Celera are the databases of biological information [7]. The Internet is exerting tremendous pressure on these data providers to create services (often referred to as e-utilities or web services) that allow users to search and access databases easily. While this trend is a privilage to end users, it is bewraying the data providers to the threat of data theft. They are therefore impositioning capabilities for identifying pirated copies of their data.We suggest that rights management of relational data through watermarking should become an important topic for database research. Database relations that can be watermarked have authorship should be such that changes in a few values do not affect the applications [8].there real-world datasets that can resist a small amount of error without degrading their usability consider the ACARS meteorological data used in building weather prediction models. The wind vector and temperature accuracies in this data are calculated to be within 1.8 m/s. The errors introduced by watermarking can easily be managed to lie within the measurement resistance in this data. As another example, consider experimentally obtained gene expression datasets that are being analyzed using various data mining techniques [9]. The nature of some of the data sets and the analysis techniques is such that changes in a some data values will not cause any affect in the results. Similarly, the customer divisions results of a consumer goods company will not be affected if the external provider of the supplementary data adds or subtracts some amount from a some transactions. Later in the paper, we report experimental results using a forest cover dataset. It contains measurements for variables such as elevation, aspect, slope, distance to hydrology and roadways, soil type, etc. Small changes in some of the measurements do not affect the vitality of this data. Finally, consider the parametric data on semiconductor parts adverted earlier. For many parameters, errors inculcated at watermarking can be made to lie within the measurement acceptance [10]. It is noteworthy that the publishers of books of mathematical tables (e.g. logarithm tables and astronomical ephemerides) have been introducing small errors in their tables for centuries to identify pirated copies. A cryptographically secure false random sequence generator *G* predictabilly generates a sequence of numbers in which calculatively it is not possible to predict the next number in the sequence. Statistically, the numbers generated by *G* appear to be a realized sequence of independent and identically distributed random variables, in the sense that the numbers pass standard statistical tests for these properties [11].

2.1 PURPOSE of RESEARCH:

One potential solution for claiming the ownership is to use electronic stamps or also called watermarks, which are embedded in the form of images, and have secured the transmission of digital watermarking it is widely applied to relational database for ownership protection and information hiding. The process of watermark embedding and detecting, the databases tuples are adaptively grouped into groups according to the length of binary watermark.• undeletable by hackers, Identification of the rights management of relational data through watermarking as a very important and technically challenged problem for database research[3,12].

• perceptually invisible, i.e., the watermark should not render visible artifact;
• statistically undetectable;
• resistant to lossy data compression, e.g., the Joint Photographic Experts Group (JPEG) compression;
• resistant to image manipulation and processing operations, cut-and-paste, filtering, etc.

3. EXPERIMENT DESIGN SPECIFICATION

The suggested algorithm is Watermark Embedding or Insertion. In watermark embedding database is embedded into image. Here we first store binary data of image in one data structure (suppose array) and then database content in another array. Then we copy content of image and database in third array [13]. If image data size is [a*n] second database size is [9*n] and third data structure size is [(a+9)*n] where n=6 in this algorithm. This is one time process when watermark is embedding

Coding of Check Column

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

Check column: In this column checking is done because it is important to check image code and database. They must be same while combining [14, 15].

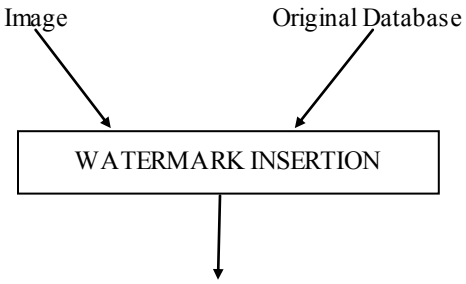

Figure 2: Watermark Embedding or Insertion

In watermarking algorithm uses main logic. [16]

1. Check column
2. Check column: In this column checking is done because it is necessary to check image code and database. They must be same while merging.

I. Extract the database values into dynamic array and change it into ASCII and after that binary.
II. Same procedure is done for image which is considered as key
III. Merging the output of step 1 and 2 into third dynamic array.
IV. After merging third array is again converted into image, this will be main data used for comparison or our watermarked image.

Figure.3: Watermarking Embedding or Insertion Algorithm [17]

Merger: First image binary code is stored in one data structure and database content is stored in other data structure. Then first image binary code is stored in third data structure and in same data structure after storing image, database content is stored.

```
/*
 * To change this template, choose Tools | Templates
* and open the template in the editor.
 */
package encrypt;
import db_processing.DataSetExtractor;
import image.CreateImage;
import java.util.LinkedList;
import java.util.List;
import java.util.ListIterator;

/**
 *
 * @author VINAY
 */
public class Merger {

    public List<String> mergeTheList(){
        final List<String> mainList = new LinkedList<String>();
        final List<String> it1 = new CreateImage().getImageBinaryIterator();
        int count = 0;
        for( count=0;count<it1.size();count++){
            System.err.println("image list:"+it1.get(count));
                mainList.add(it1.get(count));
        }
        System.err.println("image count"+count);
        final List<String> it2 = new DataSetExtractor().getStringEquivalentOfDataList();
        for(count=0;count<it1.size();count++){
            mainList.add(it2.get(count));
            System.out.println("data list:"+it2.get(count));
        }
System.err.println("database count"+count);
        return mainList;
    }
    public static void main(String[] args){
        List<String> list = new Merger().mergeTheList();
        int count=0;
        for(count=0;count<list.size();count++){
            System.err.println("merge list:"+Integer.parseInt(list.get(count)));
        }
        System.err.println("total count"+count);
```

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

}
}

4. IMPLEMENTATION AND RESULT

For testing the validity and vigrousness of the given algorithm, we perform experiments on computer running on Windows XX, or any other configuration of windows with 2.4GHz CPU and 512 MB RAM.Algorithm is implemented on Java platform and uses Java Netbeans IDE6.9.1, Apache Tomcat 6.0.26, and MySQL Database 5.0. In our experiment we consider table of m*n where m=9 and n=6. The algorithm compares watermarked data and current database. It performscomparison tuple by tuple, if matched nothing is returned if not then it return (m*n) the place where updated or changed value

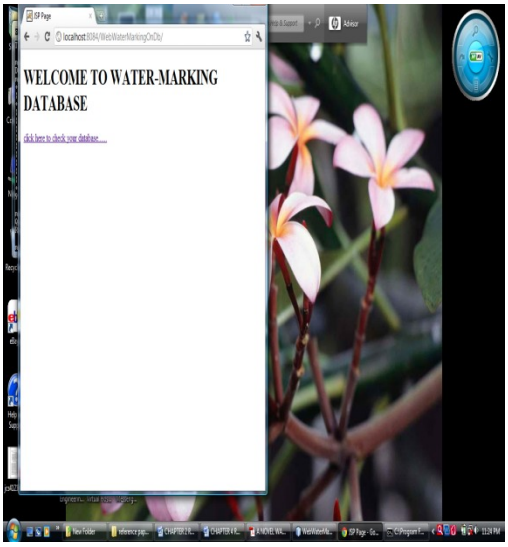

Figure.4: performed comparison tuple by tuple.

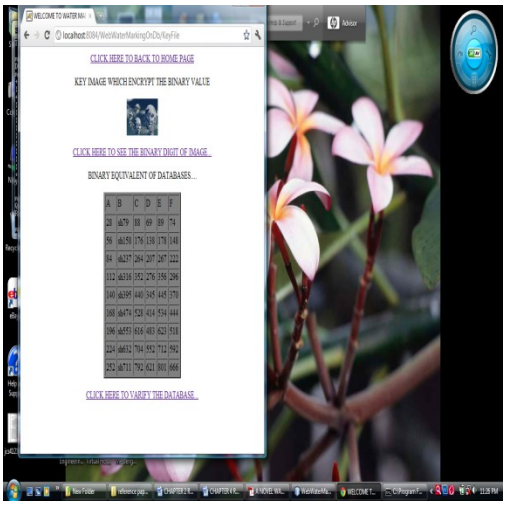

Figure.5: the place where updated or changed value.

5. CONCLUSION AND FUTURE SCOPE

In this paper we have discussed about watermarking algorithm which is used to recover database from updated value. This is a web based watermarking technique, in which if the software is stored in server then we can easily access it from client system. If this software is stored in server we can access it by giving its IP address. Database considered only accepts input of length five character either alphabet or numbers. The limitation of this is that it accepts only those data in database which have five characters. It is able to recover database that have both numeric as well as alphabetic characters a novel suitable watermarking scheme based on grouping and polar angle expansion for relational database, which first takes advantage of the disorder character among database tuples till grouping distance, and then unites with the polar angle expansion strategy to embed and drawn out watermark. The scheme shows a high robustness under blind detection for subset selection, submation and modification attack, and also can recover the original data more truly. Due to the local convergence of fast grouping and the error of restoration data, it cannot satisfy the application requirement of high-accuracy data. The next step is

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

to adopt new update strategy to speed up convergence rate and global convergence, then design a completely reversible database watermarking algorithm and prove it in theory. Distinctness of the various forms of spiteful attacks from which the watermark inserted in a relation must be conserved. First proposal of a watermarking technique specifically geared for relational data. . Extensive analysis and empirical evaluation of the robustness and effectiveness of the proposed technique to show the possibility of watermarking real life set of data.

6. FUTURE WORK

Future work will be for large amount of dataset having multiple characters for real world applications. The intent of this paper is to provide a high-level, introduction to the watermark recovery we are pursuing and document preliminary results. This work will be further detailed in futur work We are currently expanding the ideas introduced here to include automatic image recognition, image refinement in the recovery phase, and the investigation of the invariant properties between point clusters between images

REFERENCES

[1]. Rakesh Agrawal and Jerry Kiernan, (2002) Watermarking Relational Database, in proceedings of 28th VLDB Conference, Hong kong.

[2]. R Agrawal, P J Haas, J Kiernan, (2003) Watermarking relational data: framework, algorithms and analysis, VLDB, vol. 3.

[3]. Zhi-Hao Zhang, Jin, Wang, & Li, (2004) Watermarking Relational Database Using Image, in proceeding of 3rd international conference on M/C learning & Cybernetics, Shanghai.

[4]. Z. Sun, Z. Cao & Z. Hu, (2008) Multiple Watermarking Relational database s using image, International Conference on Multimedia & IT.

[5]. Li Yingjiu, (2008) Database Watermarking: A Systematic View, in hand book of Database Security.: Springer US.

[6]. Z Hu, Z Cao & J Sun, (2009) An Image Based Algorithm for Watermarking Relational Databases, in International Conference on measuring Tech. & Mechatronics Automation, 2009.

[7]. H Khataeimargagheh & H Rashidi, (2010) A novel watermarking scheme for detecting and recovering distortions in database tables, IJDMS.

[8]. R Bedi, A Thengade & V M Wadhai, (2011) A New Watermarking Approach for Non-numaric Relational Database, International Journal of Computer Applications, Vol-13, No 7.

[9]. H M Sardroudi & S Ibrahim, (2009) A new approach for Relational Database Watermarking Using Image.

[10]. M E Farfoura & S. Hong, (2010). A novel blind Reversible Method for Watermarking Relational Database, International Symposium on Parallel and Distributed Processing with Applications.

[11]. D Hanyurwimfura, Y Liu & Z Liu, (2010). Text Format Based Relational Database Watermarking for Non-numeric Data, International Conference On Computer Design And Application.

[12]. A Deshpande & J Gadge, (2009). New Watermarking Technique for Relational Databases, Second International Conference on Emerging Trends in Engineering and Technology.

[13]. Kaiyin Huang, M Yue, P Chen, Y He, X Chen, (2009). A Cluster-Based Watermarking Technique for Relational Database, First International Workshop on Database Technology and Applications.

[14]. X Dong, X Li, G Ye & L Zheng, (2009). An Algorithm Resistive to Invertibility Attach in Watermarking Relational Databases, IEEE.

[15]. A Odeh & A Al-haj, (2008). Watermarking Relational Database Systems, IEEE.

[16]. H Cui, X Cui, Mailing Meng, (2008). A Public Key Cryptography Based Algorithm for Watermarking Relational Databases, International Conference on Intelligent Hiding and Multimedia Signal Processing, IEEE.

[17]. R Halder & A Cortesi, (2010). Persistent Watermarking of Relational Databases, International Conference on Advances in Communication, Network, and Computing, IEEE.