International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 2 Issue 9‖ September 2013 ‖ PP.84-89

Research of Applications in Relational Database on Digital Watermarking Technology

Zhanfei Lei[1], Rong Li[2]

[1]*(School of Information Sciences, University of Pittsburgh, America,)*

[2]*(School of Information and Control Engineering, Xi'an University of Architecture and Technology, China)*

ABSTRACT :*Digital watermarking technology is the technology that is often used in the field of digital multimedia copyright protection. It has also been gradually applied in the ownership management of relational database in order to protect database ownership through embedding digital watermarking technology containing practical significances in relational database. Firstly, this paper compares the differences of digital watermarking technologies in multimedia area and relational database and describes the development respectively; Secondly, combines the current watermarking technologies and detection models in relational database with the existing algorithms of database digital watermarking technology; finally, discusses the latent menaces of watermarking technology in database.*

KEYWORDS: *Relational Database, Digital Watermarking Technology, Database Copyright Protection*

I. INTRODUCTION

With the development of information technology, multimedia digital information is turning into indispensible parts in human's working and living. Digital information is being disseminated broadly on the Internet via more convenient methods, so that it brings a great deal of attention to the problems such as data theft, illegal diffusion, etc[1].Digital watermarking technology implements data copyright protection according to specific way embedding copyright information in digital objects. At the moment, the technology is taken as a new security technology to protect copyright within not only multimedia digital area but also many fields. The more quickly database technology develops with larger data volume, the more important security issue in relational database becomes. In that case, it is so important to solve database security problems that digital watermarking technology needs to be applied in database copyright protection by embedding watermarking information of copyright in relational database based on present research of watermarking technologies in multimedia fields.

II. DATABASE DIGITAL WATERMARKING TECHNOLOGY AND ITS DEVELOPMENT

The working of database digital watermarking technology is that to achieve database copyright protection by means of information processing which embeds copyright information with particular format in relational database. The precondition of it is to ensure what is contained in the database can surly be manipulated correctly[2]. Data models in relational database are flat, two-dimensional. Being similar with multimedia digital watermark, database watermark possesses characteristics of robustness, imperceptibility, detection-resistant and security. For now, compared with traditional multimedia digital watermarking technology, the specialty of relational database determines

the situation that database watermark will face significant challenges. First of all, redundancy in relational database is relatively lesser, therefore bandwidth of noise region available for embedding watermark information is narrow which makes it more difficult for watermark than multimedia data to embed watermark in database. In addition, algorithms of database watermark demand high standards of robustness. Low level of robustness in watermark algorithm means that normal update manipulations or malicious attacks to database may result in loss of watermarking information embedding in database leading to lose meanings of digital watermark protection. The effective way researchers in this field concentrating on to explore actively is to make compromises of how to increase database redundancy, how to improve robustness of digital watermark and how to embed digital watermarking information in relational database, on the basis of ensuring regular use of database and demands, such as digital watermarking invisibility and etc[3]. On account of particularity of relational database, to apply existing multimedia watermarking technology into relational database there are difficulties. Right now researches of the aspect lie in development. The direction of research in database watermarking technology at present will be the study of database signal capacity and embedding watermark in channels. To protect database ownership through digital watermarking technology possesses vital theory significances and a wide range of potential applications.

III. DIGITAL WATERMARKING TECHNOLOGY IN RELATIONAL DATABASE

3.1.Differences between relational database data and multimedia data : It is reasonable to apply the basic ideas of multimedia digital watermarking algorithms into database area in the study of database watermarking technology. However, it is much more complex to deal with database watermarking technology than multimedia digital watermark. Compared with multimedia data, data in relational database has many different aspects[4]:

(1) Multimedia data objects usually consist of large numbers of bits with redundancy in which plenty of digital watermarking information can be stored. In relational database, a great many independent unique tuples constitute data. Since every space of tuple stores identified value leading to low redundancy, it is not easy to store watermarking information in separate tuples dispersedly.

(2) Each point in multimedia data objects has constant relative space and time positions, while in relational database tuples and fields to form tuples are not displayed in specific sequence.

(3) If there are some changes or missing data in certain parts of multimedia data objects, it is palpable to detect. Yet things are different in relational database, even partial tuples are replaced or deleted generating changes of database they are hard to be discernible.

(4) Multimedia data stay unchanged in most conditions, where as it is necessary to update and maintain data in relational database.

3.2 Models in database digital watermarking systems

Typical models of relational database digital watermarking systems include two aspects: watermark embedding and watermark detection. The elements in the process of watermark embedding are original database, watermarking information, key and algorithm of watermarking embedding, simultaneously elements in the process of watermarking detection include database embedded with watermark, key and algorithm of watermarking extraction. These two processes are converse to each other. The inner workings are shown in Figure 1 and Figure 2 respectively[5].

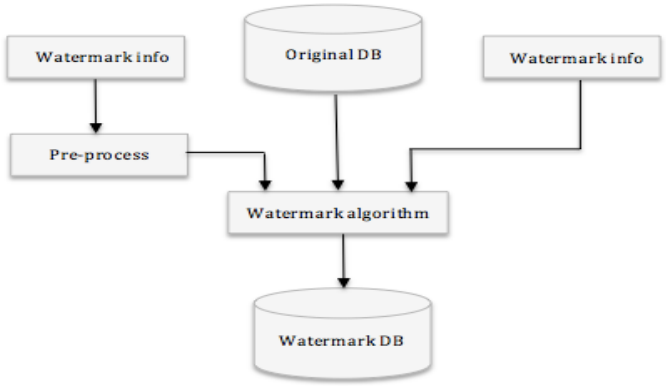

Figure 1 Model of relational database watermarking embedding

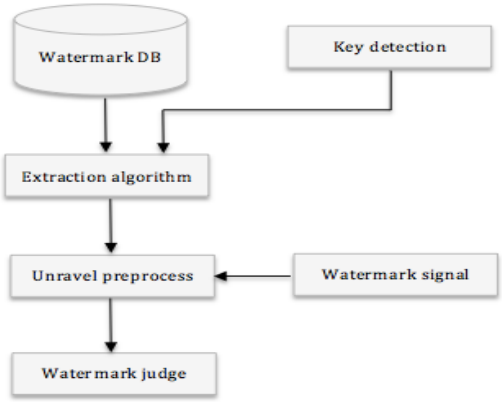

Figure 2 Model of relational database watermarking extraction

In the models above, digital watermarking signal is binary bit sequence. The embedment and detection of watermarking signals are controlled by key. The course of embedding database watermark is on the basis of preprocessing specific watermark information to embed the preprocessed digital watermark information into the original database applying key and watermark embedding algorithm, on the purpose of forming database containing watermark. In the process of extracting relational database watermark, with the help of key and algorithm of watermark extraction the aim is to detect watermarking information in the database and further to identify the embedding watermarking signal with the participation of watermarking information after unraveling the preprocess. The embedding algorithm of digital watermark in this model is public so the system security is dependent on key.

IV. THREE TYPES OF DIGITAL WATERMARKING ALGORITHMS IN RELATIONAL DATABASE

4.1 R. Agrawal's digital watermarking technology in relational database : R. Agrawal proposes a method to mark numeric attribute values in database, and the basic marking idea is to allow errors in numeric attribute values to some extend in presumptive relational database. Under the condition of guaranteeing normal usage of database within the error range, there exist two steps: firstly, according to original key from user settings, primary key value of tuple and percentage of tuples need to be marked, determine to mark which elements by the usage of one-way hash functions in encryption algorithm, such as SHA function and MDS function; secondly, identify attributes and

positions of bit need to be marked based on the number of attributes and bits that could be marked and then mark part of certain attributes and bit positions which meet eligibility requirements as 0 or 1 taking for markers in the relational database. The bit pattern combined with multiple bit markers is the digital watermarking information embedded in relational database. Throughout the entire process, tuples that need markers, the attributes of tuple, and the positions of bits chosen in attributes and the specific settings of bit values are all confirmed by key, primary key value of tuple and the algorithms controlled by percentage of tuples that need to be marked. Especially, the key, percentage of tuples marked, dimensions of attributes can be marked and the number of bits can be marked are only known by owner of the relational database. In that case, the security of database digital watermark algorithm has been largely improved, because on the premise that attacker for malicious purposes didn't modify the attribute values of tuples, watermarking information embedded in the database could be extracted only by getting the key.

4.2 R. Sion's digital watermarking technology in relational database

R. Sion and etc. further study database watermark also coming up with an algorithm to mark numeric attribute values. The core of the strategy is to preset a numeric item set $S = \{s_i, s_i + 1, \dots, s_n\} \in R$ and an ordered key *K* ,and then implement the strategy by two steps: first, sort the items in the set according to the values of hash function which correspond to key value of bits containing the most information among normalized items, for example, $index(s_i) = H(K, msb(s_i), K)$; second, construct subset s_i to embed watermarking marker of bits into it and suppose that the length of information is m bits and that each bit be embedded into S_i, then this kind of method improves the capability to resist different attacks like "subset selection" attack and "subset increase" attack. Suppose that attacker for malicious purposes took out 5% data, it would reduce 5% to each subset S_i. If the subset S_i were small enough, the watermarking algorithm would appear robustness to such kind of decrease. To embed a bit into each subset, it is necessary to acquire the length of whole marking code, that is the size of subset s_i, $|S_i|$, according to which the size of bandwidth that can be used to mark is $s_i/|S_i|$ bits. Usually speaking, it is implementable to embed the original watermark for several times as long as *S* is large enough. The method works well to prevent subset shearing attack. It needs to ensure the times of embedding watermark would not be over than $s_i/(|S_i| * m)$. When extracting watermarking information, taking the approach of plurality election in all recovery bits after given data recovered to generate watermark backup, makes sense to identify the most probable original digital watermarking bits.

4.3 N. Xiamu's digital watermarking technology in relational database

N. Xiamu and etc. put forward the idea that to add watermarking technology with less practical meaning to relational database[6]. In this algorithm, with the application of current digital watermarking algorithms in relational database, the key point is to implant for a matching relation to certain bit in the value of chosen attribute. That is, set the value of the least significant bit as 1 when hash value corresponding to the serial number of attribute value is even and set the value of the least significant bit as 0 when hash value is odd. Such algorithm is one kind of verification algorithms that can only testify whether watermarking information has been added to the database, but fails to embody watermarking information with practical meanings. In that case, it is impossible to attest the copyright information of database by the result of this verification. The useful value of digital watermark could be enhanced greatly if significant bit value with practical meanings added in relational database. Actually, among the bits with setting value of 1 in above verification algorithm, reset the bit value as 1 if there is watermark and reset 0 if not. In other word, watermarking algorithm here is to add 1 to appropriate

bits in database. At this point, the algorithm just involves a small subset of relational database with the serial number divisible by λ. Suppose that there are *m* subsets available in relational database $(m < \lambda)$ and that each subset be added with bit value of 1, then the whole database can be embedded with bit string with length of *m*. First of all, number each attribute value A_i that can be implanted digital watermarking, set the serial number as $index(A_i)$. Then group the attribute values by setting $j = index(A_i) mod\ \lambda + 1$, and the result of grouping is that to assign attribute value A_i to subset S_j. Suppose that each bit of watermark is *w*, the length of each bit is $m(m < \lambda)$ and every bit value of watermark is $w_k(1 \le w_k \le m)$, then the determinant of implanting watermark to k_{th} subset in database is whether the value of w_k is 1. That is if $w_k = 1$, implant digital watermark into each attribute value of subset S_k according to above watermark algorithm.

V. SECURITY MENACES IN RELATIONAL DATABASE WATERMARK

Because of the specialty of relational database, the copyright protection of watermarking information would be loss of meanings if watermark embedded in database lost for updating manipulation in database. What's more, relational database might face various kinds of malicious attacks. Common attacks are as following[5,6,7]:

(1) Subset attack. Avoiding of selecting all attributes and tuples in watermarking signal database, attackers only use subset of attributes and tuples to delete watermarking signal of data.

(2) Mixing and matching attack. Data thieves steal discrete tuples from multiple databases containing similar information to create their own relational database.

(3) Addition attack. Data thieves implant their own simple watermarking signal into the stolen relational databases that have been embedded with watermark, and then acclaim their ownership of the relational databases.

(4) Reversible attack. If data thieves find the existence of illusive watermark in stolen relational database, they would take reversible attack acclaiming their ownership of the database. In fact, the watermark they acclaimed is just randomly generating watermark in database.

VI. CONCLUSION

In the field of information security study, especially in database security area, it is a vital research trend of applying digital watermarking technology to realize copyright protection of database by embedding watermarking information with practical meanings in relational database. The paper mainly analyze the differences between data in relational database and multimedia data, characteristics of database watermark and basic theory and watermark algorithms of current main watermarking technology, security threatens exist in relational database and common types of malicious attacks. Based on assurance of normally usage of database and invisibility of digital watermark, to improve robustness of digital watermark, to embed digital watermarking information into database and to enhance assessing standards of database watermarking technology should all be focused in research for now. In sum, to study copyright protection of database through digital watermarking technology is of great theoretical meanings and application prospects.

REFERENCES

[1] Guifang Z. Research of Digital Watermarking for Databases[D]. Changsha: University of Hunan, 2006.

[2] Qin Z, Shoujian Y, Jiajin L. Research Work and Progress in Database Watermarking[J]. Computer Engineering and Applications, 2006, 42(29): 198-201.

[3] Xiangrong X, Mingxing S. On Watermarking-Based Database Security Control[J]. Computer Engineering and Applications, 2005(6): 175-178.

[4] AgrawalR, Kiernan J. Watermarking Relational Databases[R].Hongkong: Proceeding of the 28th VLDB Conference, 2002.

[5] Sion R, Atallah M, Prabhakar S. Ownership Proofs forCategorical Data[J]. IEEE Journal of Transactions on Knowledgeand Data Engineering, 2005: 17(7).

[6] Xiamu N, Liang Z, Wenjun H & etc. Watermarking Relational Databases for Ownership Protectoin[J]. Chinese of Journal Electronics, 2003, 31(12A): 2050-2053.

[7] Sion R, Atallah M, Prabhakar S. Watermarking relationaldatabases[R]. Indiana: the Center for Education and Research inInformation Assurance and Security of Purdue University, 2002.