SECURE ATTRIBUTION TRANSMISSION OF DATASET FOR STREAM RECORDS

A.Aarthi*, Y.Mistica Dhas

Student, Faculty of Computing, Sathyabama University, Chennai, TamilNadu, India.
Assistant Professor, Faculty of Computing, Sathyabama University, Chennai, TamilNadu, India.

Email:aarthimailhere@gmail.com, misticajimson@gmail.com

Received on 19-04-2016 Accepted on 22-05-2016

Abstract

The fast boom of the net and associated technology has presented an extraordinary capacity to get right of entry to and retransfer of text contents. In any such context, enforcing statistics possession is an important requirement, which requires well defined answers prison aspects etc. A watermark describes facts that can be used to prove the possession of facts along with the owner, starting place, or recipient of the content. Comfy embedding calls in watermarking in embedding has problem even after eliminated from watermarking information. Watermarking techniques have been evolved for digital information and also for software program and herbal language text. Watermark embedding for relational records is made possible by the truth that actual information can very frequently tolerate a small amount of errors with none vast degradation with recognize to their usability. Especially, our proposed technique is resilient to tuple deletion, alteration, and insertion assaults.

Keywords: Invisible watermarking, IPDs ,MD5.

Introduction

The speedy boom of the net and related technology has presented an unheard of ability to get admission to and redistribute virtual contents. [1]In the sort of context, imposing information possession is an crucial requirements, which needed well defined answers, posing technical, organizational, and legal elements. Even though we're still a long way from such complete solutions, inside the last years, watermarking techniques have discovered its own difficult constructing block that performs a vital position in addressing the ownership problem. Such techniques permit the owner of the records to embed an imperceptible watermark into the statistics. [3]A watermark describes statistics that may be used to prove the possession of statistics together with the owner, foundation, or recipient of the content. [2]Secure
embedding requires that the embedded watermark have to no longer be without problems tampered with, forged, or eliminated from the watermarked data. [4]-[5]Imperceptible embedding method that the presence of the watermark is unnoticeable within the records moreover, the watermark detection is blinded, that is, it neither calls for the knowledge of the unique statistics nor the watermark.[7] Watermarking techniques had been developed for video, images, audio, and textual content information and additionally for software and natural language text. [6]By using assessment, the hassle of watermarking relational records has not been given suitable attention. [8]There are, however, many application contexts for which data constitute an important asset, the ownership of which need to hence be carefully enforced. that is the case, as an instance, of climate facts, stock market information, energy consumption, purchaser behavior facts, and medical and medical records.[9][10] Watermark embedding for relational records is made viable by way of the reality that actual information can very regularly tolerate a small quantity of blunders without any enormous degradation with appreciate to their usability.[11]

As an instance, while handling climate information, changing some daily temperatures of more than one degree is a alternated which leaves the facts nevertheless usable.[12]-[13]To this point, only some procedures to the trouble of watermarking relational data have been proposed.[14] Those techniques, however, aren't very resilient to watermark assaults.[15]-[16] In this paper, we gift a watermarking method for relational statistics that is rather resilient in comparison to these strategies.[17] Particularly, our proposed approach is resilient to tuple deletion, alteration, and insertion assaults. The principle contributions of the paper are formulated as follows:. As we prepared the relational databases in watermarking in a confined optimization trouble and talk efficient techniques to address the limitations.

Related Work:

As increasing quantities of information in fast, networked surroundings, certainly one of its most important functions threatens to come to be its worst enemy: zero-value verbatim copies. Image watermarking for relational databases emerged as a candidate method to offer copyright protection of relational facts, retaining integrity of the database statistics, tamper detection, traitor tracing and many others. In this paper, we introduce the perception of continual watermarking that serves as a way to apprehend the integrity and possession evidence of the database bounded with a set of queries. It permits the evaluation of the database whilst applying the queries. We preserve the persistency of the watermark by way of exploiting two invariants of the database state with respect to the set of queries: solid Cells and
Semantics based homes of the information. We talk how we are able to enhance the prevailing techniques. As we suggest a unique continual watermarking scheme that strictly improves the set of rules proposed via Li and Deng.

In this paper, a new approach for protecting the possession of relational database is presented. Such method is implemented for protective both textual and numerical data. That is accomplished with the aid of adding most effective one hidden file with a mystery feature. For each characteristic, the price of this function depends on the facts saved in all different facts. Therefore, this method is greater effective against any attacks or adjustments inclusive of deleting or updating cell values. For example, there is no want for added garage place as required when adding additional columns especially with big databases. In addition, in case of protective information by adding columns, we want to feature a number of columns similar in range of sorted matrix has been deleted. Right here, handiest one record is sufficient to guard all varieties of statistics. Furthermore, there may be a opportunity to apply a different function for every discipline results in greater robustness. Ultimately, the proposed technique does now not have another necessities or regulations on both database design or database administration.

Proving possession rights on outsourced relational databases is a vital problem in nowadays internet-based application environments and in lots of content distribution applications. In this paper, we present a mechanism for proof of possession based on the relaxed embedding of a robust imperceptible watermark in relational information. Our technique overcomes a main weak spot in existing techniques. Watermark interpreting is based on a threshold-primarily based approach characterized by using a most appropriate threshold that minimizes the chance of deciphering mistakes. We applied a important performance in our watermarking technique and showed by means of experimental outcomes that our method is given importance to tuple deletion, alteration, and insertion attacks.

**Existing System**

Watermarking in least sizeable bits (LSB). This method embeds the watermark bits inside the least giant bits of decided on attributes of a particular subset of tuple’s. It makes use of mystery key in watermarking. For every tuple’s a at ease message, authenticated code is computed the usage of the name of the game key and tuple’s number one key. The computed MAC is used pick candidate tuple’s attributes and the LSB positions in the selected attributes. This approach does now not offer mechanism for multi bit watermarks. The watermark may be without difficulty compromised by means of very trivial attacks.
Proposed System

Overview:

Watermarking embeds ownership records in digital content. Watermark describes facts that can be used to show the ownership of relational database. Right here the embedding is hidden that the user cannot see the invisible watermarking technique. It is not resilient to watermark assaults. Gold standard threshold reduces probability of deciphering errors. It is not resilient to watermark attacks. Optimal threshold minimizes the statistics of encryption error. Additive security in watermark bit in dataset is increased in this technique.

![System Architecture](image)

**Figure 1: System Architecture**

Material and Methods

A. Data Partitioning

Information partitioning in relational information warehouse can carried out with the aid of objects partitioning of base tables, clustered etc. range partitions consult with table partitions that are defined by using a customizable variety of records. The quit user or database administrator can define the partition function with boundary values, partition scheme having record institution mappings and table which are mapped to the partition scheme. by means of using flow the information set is partitioned into several non-overlapping walls.
Original data set D is requires Hash function(MD5) in partitioning. Single bit encoding function consists of three input parameters hidden key K, Elements of partitions m and watermark W i.e. known to only copyright owner. It transforms original data into watermarked data W. Using usability matrix the single bit dataset gets alternate in data. Bit position is selected with the aid of using general quantity of partitions and watermark period which is known to most effective owner of data. When the unique price of statistics receives changed because of the watermark bit, it always tests the information usability constraints after placing the watermark within the partition, merge all walls and get the complete watermarked information. While interpreting, use majority voting algorithm to get the ideal watermark. So the retrieval of the original watermark claims the copyright of records.

B. Watermark Embedding

A watermark bit is embedded in every partition through single Bit Encoding set of rules. Watermarking is a technology for embedding various kinds of records in digital content material. In general, information for defensive copyrights and proving the validity of records is embedded as a watermark. Watermarks are added to pictures or audio information in this type of way that they're invisible or inaudible Ñ unidentifiable by way of human eye or ear. Moreover, they can be embedded in content material with a spread of file code. Watermarking is the content safety approach for the multimedia era.

C. Optimal Threshold Evaluation

The bit embedding records are used to compute the premier threshold that minimizes the possibility of decoding error. The optimization technique used in this test is sample search technique (PS). PlayStation strategies are direct seek techniques for non-linear optimization. It starts at an initial factor and samples the objective characteristic at a predetermined pattern of points targeted approximately that point with the intention of producing a new better iterate.
A. Dataset Partitioning

Via using the statistics partitioning algorithm, the facts walls are generated from watermarked dataset. The single bit encoding includes data partitioning and watermark embedding algorithms. In watermark detection algorithm we use single bit decoding. As proven in the flowchart, the first step is to partition the authentic information and assign partition variety to every and every tuple of the relation using Cryptographic Hashing feature (MD5). After that the preferred watermark has been selected, however the length of the watermark should be very less than the wide variety of partitions, so that the identical watermark have to be embedded often inside the unique data. recollect only those partitions for watermarking wherein variety of tuples in that partition ought to be extra than some threshold value. at the same time as converting the information, bit bi is chosen from each partition and that bit bi is modified using watermark W. So every time the new bit role of the attribute is selected from different walls. therefore it's far hard to find out that particular bit as encoding is relies upon on wide variety of partition of facts and secrete key Ks.

B. Threshold Based Decoding

The information of each partition are evaluated, and the embedded is decoded using a threshold based totally scheme based at the top-quality threshold. The opportunity of bit decoding errors is deni ned because the probability of an embedded bit decoded incorrectly. The deciphering threshold T is selected such that it minimizes the possibility of interpreting errors. The bit embedding stage is primarily based at the maximization or minimization of the tail count number; those optimized hiding feature values computed at some point of the encoding stage are used to compute the ideal threshold T.
Algorithm

A. Data Partitioning Algorithm

Now we present the detail of Data partition algorithm using MD5. It includes four algorithm: KeyGen, Insert, Modify, Delete.

Input: Data Set D, Secret Key Ks, Number of partitions m

1. \{S0, S1, ……Sm-1\} \[ \square \{ \} \]
2. For each tuple r \( \in D \)
3. Partition( r ) = H(Ks \( \parallel \) H( r .p \( \parallel \) Ks)) mod m.
4. Insert r into S partition(r)
5. Return S0, S1…… Sm-1

Output: Data Partitions S0, S1…… Sm-1

B. Single bit Encoding Algorithm

Now we present the detail of single bit encoding algorithm. It includes three algorithm: Insert, Modify, Partition

Input: Data Set D, Watermark W = \{ b0, b1, b2…… bl-1\} of length l, Secret key Ks, number of partition m

1. \{S0, S1, ……Sm-1\} = get_partition(D,K,m)
2. for each partition S
3. get_watermarked_bit(W) // select watermarked bit from W
3.1 i = k mod l // get I bit from selected attribute
3.2 bi = get_bi (Selected _ attribute) // get the original value of bit bi
3.3 change_bit(bi, selected _attribute) // change bit bi with watermarked bit
3.4 Check_Usability_Matrix(selected _attribute , bi)// checks the watermarked value with usability matrix
3.5 insert S k w into D w // insert partition S into watermarked data
4. Next partition // get the next partition
5. Return D w

Output: Watermarked Data Set D.
Result and Discussion

Data partitioning consume less time during transformation. Even if power fails, there is no data loss and stored in database. The system is designed to address the requirements of typical enterprise media streaming systems. While building this system we have addressed many challenges we are investigating various optimizations in the coding and streaming to improve the bandwidth utilization while minimizing the distortion experienced by the clients in wired and wireless networks. Text files can be transferred with secure and less packet loss. Audio files are encrypted and divided into packets. Video files are separated into frames then into pixels. In Proposed System, the resilient to random tuple deletion, insertion, modification attack have access lower data to the attackers. The watermark accuracy deteriorates for various percentages as 20%, 40%, 60% have given a access of different values shown in Table 1. The original number of tuples is inserted in the insertion attack and similar watermark accuracy when subject to the insertion attacks. In Modification the watermark embedding technique exploits the feasible alteration space by solving an optimization problem to enforce the competing objectives based on the bit to be inserted. Furthermore, decoding threshold is computed based on the embedding statistics to maximize the probability of decoding error.

Comparative Study

Table-1 illustrates the performance of comparative study of our proposed system with the existing system. The result shows that the proposed system consumes minimum time while partitioning compared to existing mechanism. Multiple embedding of watermark bits in the dataset increases additional security.

Table 1: Bit Error Rate for Existing System and Proposed System.

|                | Insertion |       |       | Deletion |       |       | Modification |       |       | Existing system |       |       | Proposed system |       |       | Existing system |       |       | Proposed system |
|----------------|-----------|-------|-------|----------|-------|-------|-------------|-------|-------|----------------|-------|-------|----------------|-------|-------|----------------|-------|-------|----------------|-------|-------|----------------|-------|-------|
|                | 20%       | 40%   | 60%   | 20%      | 40%   | 60%   | 20%         | 40%   | 60%   | 20%            | 40%   | 60%   | 20%            | 40%   | 60%   | 20%            | 40%   | 60%   | 20%            | 40%   | 60%   |
| Existing system| 0.0874    | 0.064 | 0.079 | 0.5      | 0.5   | 0.03  | 0.067       | 0.056 | 0.059 | 0.023          | 0.013 | 0.016 | 0.121          | 0.114 | 0.124 | 0.023          | 0.013 | 0.016 |
| Proposed system| 0.042     | 0.0296| 0.035 | 0.121    | 0.114 | 0.124 | 0.067       | 0.056 | 0.059 | 0.023          | 0.013 | 0.016 | 0.121          | 0.114 | 0.124 | 0.023          | 0.013 | 0.016 |
Conclusion

On this task, we’ve offered a resilient watermarking approach for relational information that embeds watermark bits in the data statistics. The watermarking hassle turned into formulated as a restrained optimization trouble that maximizes or minimizes a hiding feature primarily based on the bit to be embedded. GA and PS strategies were employed to clear up the proposed optimization hassle and to handle the constraints. Moreover, we presented a data partitioning technique that doesn't depend upon special marker tuples to locate the mistakes in partition and watermark synchronization. We evolved an efficient threshold-based totally method for watermark detection this is based on an ideal threshold that minimizes the possibility of interpreting error. The watermark resilience changed into improved by way of the repeated embedding of the watermark and using majority balloting approach in the watermark deciphering phase. Moreover, the watermark resilience became stepped forward by means of the use of a couple of attributes.

A proof of idea implementation of our watermarking method changed into used to conduct experiments the use of each artificial and actual-world statistics. A evaluation our watermarking technique with completed techniques shows the prevalence of our technique to deletion, alteration, and insertion assaults.

Reference


Corresponding Author:
A. Aarthi*,
Email: aarthimailhere@gmail.com