Volume 5, Issue 5, MAY 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Database Tamper Detection

Chanda Kataria
Scholar (M.Tech, CSE)
Department of Computer Science and Applications
Kurukshetra University, Kurukshetra,
Haryana, India

Dr. Kanwal Garg
Supervisor (Assistant Professor)
Department of Computer Science and Applications
Kurukshetra University, Kurukshetra,
Haryana, India

Abstract: Database is a major component of each and every organization. But to store data in database is not sufficient for any organization, since they have to deal with all issues related to database, from which one of the main issue is database security. This paper deals with the basic approach that determines whether data stored in database is tampered or not. To deal with this problem, the concept of hashing is used and the result is the tampered table or row. Oracle 10g is used to deal with this problem.

Keywords: Database, Tampering, Hashing, Forensic analysis algorithms.

I. INTRODUCTION

Technology has spread its range upto highest peak and will spread it more in the upcoming time. But with its increasing height, it has to face many ups and downs which include security breaching, data corruption, system destroy and many more. Database has gradually reached in almost every field so its security is considered a major issue. Database is a field designed and developed for data and its storage. It is possible to change the stored data but if data is modified intentionally only for revenge purposes or for someone's own sake then that change in data is the breach of security and its laws. Such type of data change and data corruption is known as tampering. Database architecture is divided into three schemas: external schema, conceptual schema and internal schema. Internal schema refers to physical level that contains all physical files such as data files, redo log files, backup files, archive files etc. Any change in these files refers to database destroy. Conceptual level refers to logical/conceptual schema which represents metadata and relationship among objects and any tampering on this level is known as data corruption. External schema refers to application level which display data to users and security is more needed at this level to detect whether user is authorized or unauthorized. Tampering at this level refers to changed database [3]. It is concluded that each level of database architecture require security.

II. RELATED WORK

Research in this field has started since the era of 80's. Many researchers have come forward to solve the issue of database tampering, especially, to detect tampering in temporal databases during transactions. Richard T. Snodgrass, Shilong Stanley Yao and Christian Collberg, 2004 has provided his method to maintain and check integrity in audit logs [1]. In 2006, KyriacosPavlou, Richard T. Snodgrass explain the basic concept of tamper detection in database by providing tamper detection approach and also explain basic alforensic analysis algorithms to detect tampered locations [2]. KyriacosPavlou, Richard T. Snodgrass in 2008 explain more reliable and accurate forensic algorithms and characterize the 'forensic cost' under worst-case, best-case and average-case assumptions on the distribution of corruption sites of different algorithms [4]. Dr. B.B.MeshramShwetaTripathi, Sindhu K.K in 2012 explains forensic investigation procedures using a WinHex tool. How to recover deleted files and windows recycle bin analysis is explained by using WinHex tool [5].Pallavi D Abhonkar, Ashok Kanthe, 2012 proposed a novel solution to overcome the problem of tamper detection by notarizing the original data [6]. Arafat Mohammed Rashad Al- Dhaqm, SitiHajar Othman, ShukorAbdRazak and AsriNgadi in 2014 explain different dimaenions of database architecture possesses different changes according to the files they have at that level and also explain some of the basic problems faced in detecting tamper in database [3].

III. TAMPER DETECTION

Tamper in database can be detected by using either by using any tool or by using oracle or any SQL server. Tool such as winhex [7] is a forensic analysis hex editor tool that calculates hash value of any file and after rehashing can find whether file has been modified or not. It also provides many more functions which are helpful in forensic analysis such as data cloning, provide encryption, calculate checksum of any file etc. Except tool, tampering can be detected in database by applying queries. Here, Oracle 10g is used. Now, how one knows whether data is compromised. Solution is to apply validation events after a fixed interval of time that checks data integrity. Data integrity can be checked by considering any parameter and here that parameter is hash value. After detection of tampering, possible corrupted data is send to forensic analyzer whose job is to find the possible table or row which is corrupted and by whom. Forensic analyzer with the help of various forensic analysis algorithms computes the affected region and by checking log files he/she can determine the name of culprit. The hash algorithm which is used is strong one-way hash function.

A. Hashing Technique

Hash value acts as the basic parameter to determine tampering in database. [1] There are two techniques for hashing. One is Opportunistic hashing and other is Linked hashing. In opportunistic hashing, each tuple of a table is hashed at the time of insert and update query [1]. In linked hashing, when database is created, hash value of one tuple is calculated by combining its timestamp and schema, when second tuple is inserted its hash value is calculated by combining its timestamp and previous tuple's hash value, for the third tuple also same procedure is followed and so on [1].Here, opportunistic hashing technique is used. Hash function is known as strong one-way hash algorithm or function for reason that for a particular tuple its hash value is calculated but from that hash value to determine its tuple is difficult.

B.Tamper Detection

To detect tampering in database, create a table first, then modify the table and then show whether tampering has occurred or not. First of all, run SQL command line, type connect, provide user-name and password. It will get connected. Now create a table by typing

SQL> CREATE TABLE EMPLOYEE (emp_namevarchar(20), emp_id number(10), emp_salary number(10), emp_check_in timestamp);

After table creation, insert values in tuples with the help of insert query as

SQL> INSERT INTO EMPLOYEE VALUES('&emp_name', '&emp_id', '&emp_salary', current_timestamp);

This query will insert as many tuples as one wants to insert in the table. Inserted values of table are shown in Fig. 1. Function called 'ORA_HASH()' is used which calculate unique hash value for each tuple. This function is defined in the versions of Oracle. If any value in any tuple gets updated, the computed hash value automatic gets change for that particular tuple. Ones computed hashed value should be stored in any other file for future use at the time of validation. After update, hash values are again computed which are then matched with the previous saved hash values. If it matches for each tuple it means data is secured but if it does not match then data is corrupted.

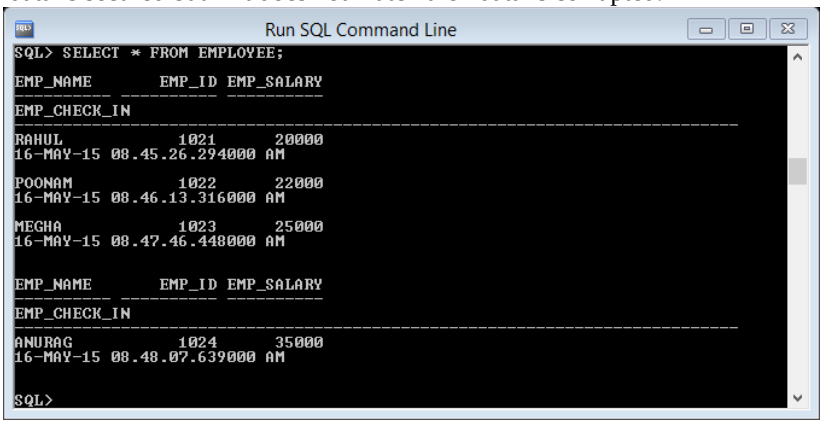

Fig. 1 Table "Employee''

*1) ORA_HASH Function:*ORA_HASH [8] is a function that computes a hash value for a given expression. Here, this function is used to find unique value for every tuple so that tampering can be detected.

Syntax: ORA_HASH(expression, max_bucket, seed_value)

The argument 'expression' determines the data for which hash value is computed. Variable length of data can be used. No user-defined data-type is used for ORA_HASH. The argument max_bucket determines maximum bucket value returned by the hash function. Its value range is 0-4294967295. This is an optional argument. At last, argument seed_value is used to produce many different results for same set of data. This is also an optional field. Return type of ORA_HASH function is a number value.SQL query for this function is

SQL> SELECT ORA_HASH(emp_name||'~'||emp_id||'~'||emp_salary) from EMPLOYEE

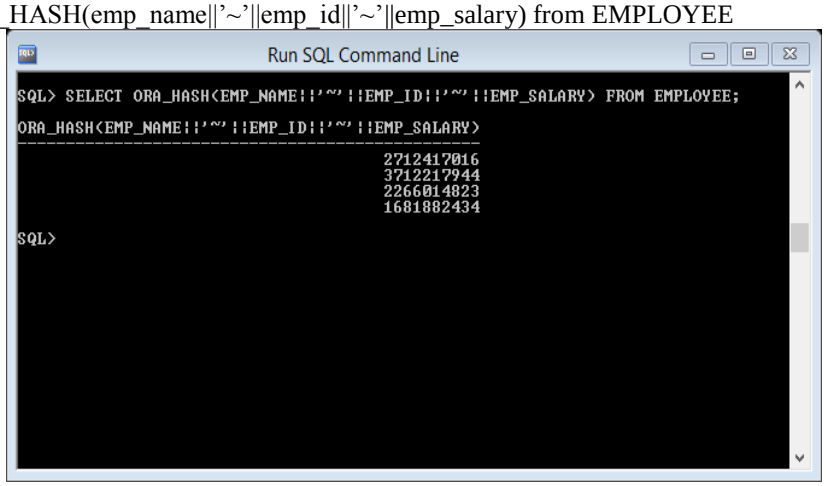

Fig. 2 Computed Hash Values

Suppose, four tuples are inserted in table EMPLOYEE. The hash values for the inserted tuples are shown in Fig. 2. These values are to be stored in a separate file so that rehashed values can be easily matched and can check whether data has been modified or not. If it is needful to know the timestamp when validator detected whether database is tampered or not, timestamp with hash value can be calculated by query:

SQL> SELECT (CONCAT(ORA_HASH(EMP_NAME||'~'||EMP_ID||'~'||EMP_SALARY),
CURRENT_TIMESTAMP)) FROM EMPLOYEE;

This will provide hash value with current time of the system and when rehashing will be done it will provide new timestamp.Later on when it is needed to validate data integrity, again hash values are generated for same tuple and match with the previous hashed values. If both hash values does match, it means no tampering is done but if both hash values does not match, it means tampering is done which has been detected. Further, now the job is of forensic analyzer to analyze where data tampering is done. For this analysis of data and to detect exact tampered location in database, many forensic analysis algorithms have been defined by the database forensic science. All algorithms works on the concept of chains which are known as hash chains. Some of the algorithms are:

*Monochromatic Algorithm:*It uses only single chain to identify tampering. This chain is black chain. This chain is produced for every validation event, i.e., it is produced every time when validation is occurred [4].

*RGBY Algorithm:*It uses four more chains with the black chain which are red, green, blue and yellow in color. Black chain is produced with every validation along with for every odd order of validation red and yellow chains are produced and for every even order of validation blue and green chains are produced. All colored chains are partial chains [4].

*Tiled-Bitmap Algorithm:*It produces multiple partial chains for every validation is stored in the form of tiles. It can handle multiple corruptions but at once but it can overestimate the degree of corruption by resulting in false positives [4].

*A3D Algorithm:*There is a slowly increase in the number of chains at each validation. Concept of fixed pattern chain is not repeated in this algorithm which helps it to locate only positive corrupted regions while handling multiple corruptions. It overcomes the limitation of tiled-bitmap algorithm. It is an improved version of all forensic analysis algorithms in case of database security [4].

These algorithms in all provide the corrupted region by inputting the value of timestamp when tampering is detected.

IV. RESULT

After applying the stated queries, tampered tuple is displayed that shows the tampered table.Suppose third tuple is modified as shown in Fig. 3 wherein EMP_SALARY is changed to 31000 from 25000. After rehashing same tuples, all hash values are same except third (tuple which was modified) one as shown in Fig. 4. This result can also be shown along with the timestamp. After getting the result one can conclude which tuple has been tampered by the intruder, either intentionally or unintentionally.

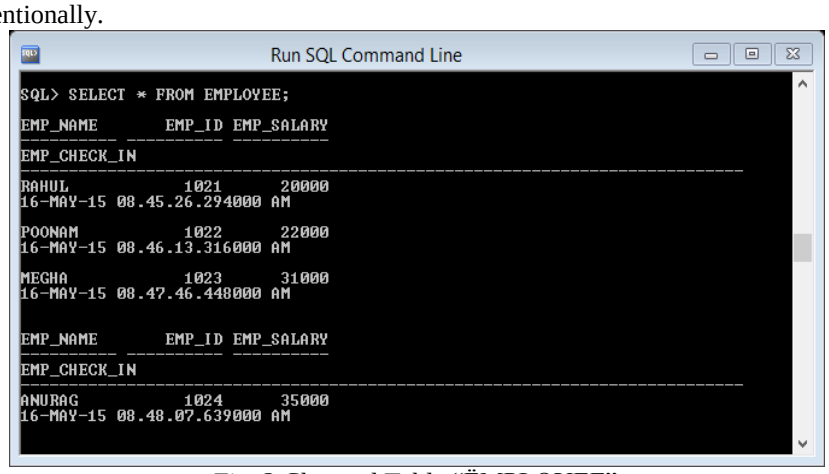

Fig. 3 Changed Table "ËMPLOYEE"

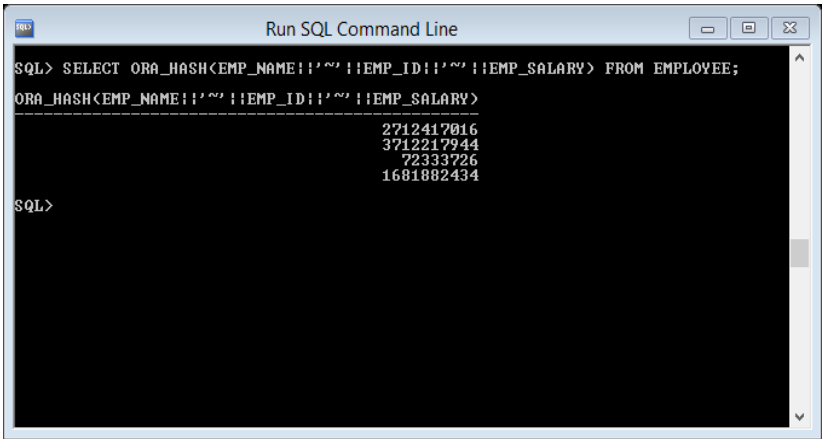

Fig. 4Rehashed Values

V. CONCLUSION

Security is the major concern in every field so as in the field of database. Many researches have been done in this field which has provided many techniques to overcome this problem. The approach used in this paper is useful for short database. Forensic tools are available to deal with large databases and database forensic field provides algorithms to find who and when tampered the data.

REFERENCES

[1] Richard T. Snodgrass, Shilong Stanley Yao and Christian Collberg, "Tamper Detection in Audit Logs", Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004.

[2] KyriacosPavlou, Richard T. Snodgrass, "Forensic Analysis of Database Tampering", SIGMOD'06, June 27-29, 2006, Chicago, Illinois, USA.ACM 1-59593-256-9/06/0006.

[3] Arafat Mohammed Rashad Al- Dhaqm, SitiHajar Othman, ShukorAbdRazak and AsriNgadi, "Towards adapting Metamodelling technique for Database Forensics Investigation Domain", International Symposium on Biometrics and security Technologies (ISBAST), 2014, 978-1-4799-4/14.

[4] KyriacosPavlou, Richard T. Snodgrass 2008, "Forensic Analysis of Database Tampering", ACM Trans. Datab. Syst. 33, 4, Article 30 (November 2008).

[5] Dr. B.B. Meshram, ShwetaTripathi, Sindhu K.K, "Digital Forensic Investigation on File System and Database Tampering", IOSR Journal of Engineering (IOSRJEN), Volume 2, Issue 2,February 2012, pp.214-221.

[6] Pallavi D Abhonkar and Ashok Kanthe, "Enriching Forensic analysis Process for Tampered Data in Database", International Journal of Computer Science and Information Technologies, Volume 3 (5), 2012, 5078-5085. URLs

[7] http://www.x-ways.net/winhex/ [retrieved on 15/5/2015].

[8] http://docs.oracle.com/cd/B12037_01/server.101/b10759/functions097.htm [retrieved on 15/5/2015].

