Incorporation of Clustering and Polar Angle Expansion with Hashing and DES Cryptographic Algorithm in Relational Database Watermarking

Linu Paul

Research Scholar, Department of Computer Science, Sree Narayana Guru College,
K.G.Chavadi P.O., Coimbatore - 641 105.

Mr.J.Santhosh, MSc., MCA., MPhil., ME., (Ph.D)

Assistant Professor, Department of Computer Science, Sree Narayana Guru College,
K.G.Chavadi P.O., Coimbatore - 641 105.

Abstract

The mode of enciphering in digital watermarking is the impending feature in relational database by the methods of encipherment (privacy preservation) and watermarking (authentication of data integrity). The method of digital watermarking is broadly utilized in the relational database for data hiding and protection of data. The method projected a simple relational database scheme of watermarking which is comprised of flexible clustering technique implemented on the tuples of the database. The Mahalanobis distance is employed to measure the similarity and this distance metric is considered for the purpose of the major challenges like robustness and reversibility due to the repeated maintenance of operators on the database tuples. The tuples of the database are clustered into groups adaptively before proceeding the phases of watermark identification and entrenching. The clustered groups are based on the binary watermarking length. Likewise, the fragments of the watermark are entrenched into or identified from the groups respectively based on the numerical data of both the Lowest Significant Bit (LSB) and polar angle expansion. The popular assessment method is employed to examine the bit value of the watermark in the blind identification strategy. These databases can be protected using cryptographic methods. The algorithms are required to entrench it in the data; video, imperceptible data connected to the data owner and existing users of the group. Hence, the watermarked data can be protected by generating a secured cryptographic noise signal. Based on the source data, a watermark signal can be generated and refuse to go along with some planned and unplanned attacks. The security can be enhanced and the attacks can be resolved by hashing and DES algorithms. The first step is

to encipher the database with the help of hash functions and DES algorithm. The receiver matches the hash values and then decrypts the original data. In the receiving end, the intruder tries to obtain the original by directly performing the decryption on cipher text of the watermarked data. The result shows that the enciphering and hashing techniques effectively identifies the attackers and safeguard the content using DES.

Keywords: Digital watermarking, Hashing, DES, Relational databases

1. Introduction

More number of users is outsourcing their information into the cloud for the efficient storage and the service computation. The data are uploaded because of the huge amount of databases which requires high performance computing and space. The data that is uploaded can be managed by entrenching the extra information for the content information, data integrity and authentication and this is handled by a cloud provider. To maintain the security and protection, the information is generally enciphered before loading and transmitting. The method of encipherment is the most authoritative in protecting the information which can transfer the source data into the inarticulate cipher text information. This type of method is useful in some criteria's that the owner of the information is not willing to reveal their confidential information into the cloud databases. At the side of the bureaucrat, the extra information like keywords are related to the content, characteristics of the data or the data authentication that are preferred to be entrenched straightly into the enciphered information for affording data security and management in the cloud. Still, the owner of the information anticipates that the source data can be absolutely reinstated because the misinterpretation is not endured in most of the applications. Then, the data hiding in enciphered feature integrated the returns of both encipherment and data hiding. The information hiding in the enciphering should not expose the privacy of the information which is very functional in cloud computing areas which turns out to be a hot spot in the community of data hiding in recent times.

Due to the extensive application in the area of cloud storage, the huge number of databases comprising of enormous amount of confidential information and the databases correlated to privacy reserving and attainment of the commercial value have been accumulated and provided across the network. To enhance this type of applications, the issues like security

and privacy of the information are to be considered. The cost coverage can be developed by initiating the feature of database as a service (DaaS), which affords information management overhaul and the local database also termed as database outsourcing. The vendor of the database entrust his management jobs including the database maintenance, access and management to the other party then the end user can obtain the preferred information with the help of personal privilege access.

During the security mode of the database, the problem of security may occur while uploading the database this new database service mode, there would be some security problems of an untrusted service provider of the database when a database is being uploaded into it i.e. information can be easily exposed or maliciously corrupted in the cloud. Generally, the information security is essentially safeguarded by the authentication and privilege management methods of the user in order to assure the data consistency, integrity and the quality of the information. The feature of access control for information security is not sufficient in most of the applications. The access control is splintered while revealing the confidential information and employed illegally. The effectual metric to this difficulty is to integrate the database encipherment and the watermarking.

At present, the cryptography is the imperative method in protecting the information digitally. The technique of digital watermarking technique is consecutively applied to safeguard the databases. The database watermarking has been projected in providing the security control of the huge amount of the databases. The first thing is, the relational database is comprised of more number of tuples and attributes without the definite arrangement among the tables consisting of tuples or attributes. The second thing is, the operators maintained in the database can repeatedly modify the tuples. Additionally, the tuples of the database dispensation always depends upon the language of the logical set operational like SQL. Hence, the database watermarking must have the capability of updating the real-time applications as well as the blind identification and it doesn't not have the possibility of implementing the multimedia watermarking method directly. The difficulty is to assure the reversibility and toughness of database watermarking.

The technique of digital watermarking that entrenches into the digital multimedia data like audio, video, image copy of control information combined with the identification of the data owner to thwart the distribution of illegal access and the data forgery. The watermark

information has to be extended over the source data in the manner that it retains imperceptible to the human visual method. The high quality scattering can be accomplished by the requirement of protected pseudo-noise signal. The watermark entrenching method is made open and the level of security is fully dependent upon the key utilized in the process of pseudo-noise generation. The sequence of the pseudo noise must have the subsequent requirements:

- It must be secured cryptographically.
- It must have the high level of correlation properties.

Hence the procedure of pseudo-noise creation is performed by utilizing the hashing algorithm called the MD5 hashing algorithm and the encryption algorithm called the DES cryptographic algorithm. A comprehensive depiction of and the DES algorithm and the hash functions is obtained in this projected method and the model of using the key features in digital watermarking is illustrated in figure 1.

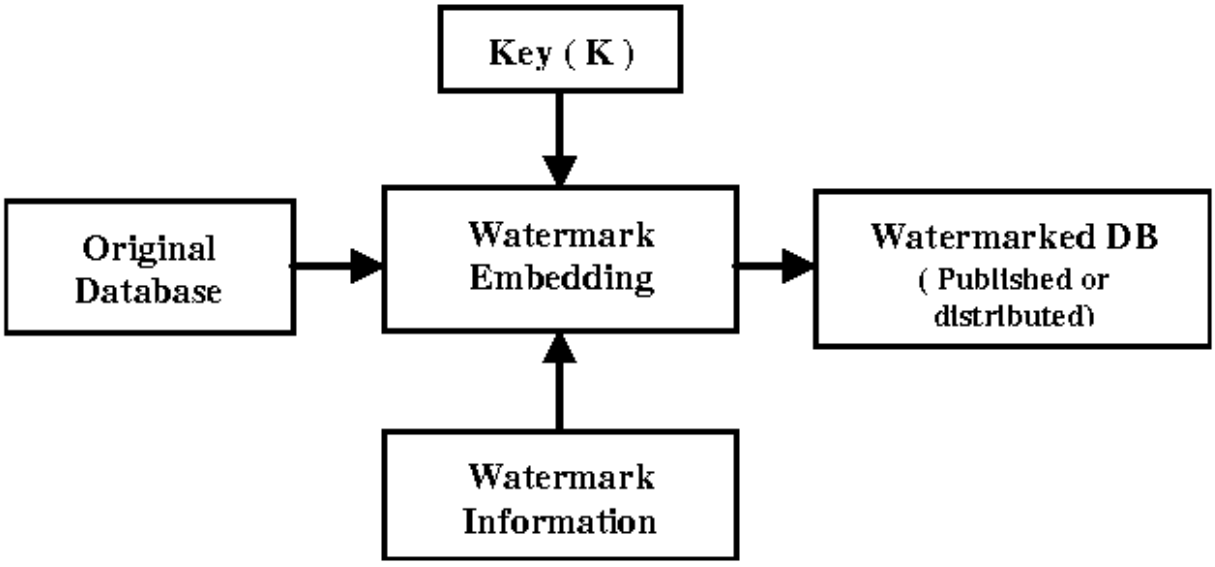

Figure No: 1 Basic Process of Watermark Embedding

2. Related Work

Agrawal et al. [1] proposed the queries regarding the range and equality. Additionally the queries like MIN, MAX and COUNT can be progressed straightly over the enciphered information. On

the other hand, the security is not delineated by Agrawal et al. and he did not state the exact security measure. The assurance needs to be afforded that the opponent cannot obtain any source information by utilizing the cipher texts. Boldyreva et al. [2] developed an effectual OPE method combined with high level of security. Furthermore, the random order-preserving function (ROPF) can be integrated with the hyper geometric distribution (HGD) and this methodology would be applicable for enhancing the safeguarding the information protection and by attaining the extrapolative effect in the cloud. Likewise, Boldyreva et al. [3] designed a modular OPE (MOPE) technique then increases the security measures of the OPE schemes. The result of this method is that the order preserving is not performed but this can be accomplished by the queries related to range query.

Agrawal et al. [4], developed the entrenching of the watermarking in the candidate attribute into least significant bit (LSB) by the subset of the tuples that is chosen. The method of watermark can be simply cooperated by few of the attacks regarding changing the LSB. Sion et al. [5] generated a method of watermarking that helps to entrench the bits of the watermark in to the information statistics by modifying the distributed characters of the information. The data detachment method is employed to utilize the special kind of tuples by creating it as very susceptible to organization errors of the watermark specifically in the deletion and insertion of tuples. Next thing is, the research is paying attention to errors of the synchronization. The methods that is stated above utilized the common property which is having the information with negligible modifications and that will not have impact on the database utilization and these type of data can be created by consuming the intricate search strategy. Moreover, the entrenching of the watermark in the domain of the transformations like DCT, DFT, and DWT can accomplish improved performance because of the orthogonal property.

Y. J. Li [6] hoisted the scheme of embedding the watermark by the order modification of the indexing of the relational data and the physical location is not modified or damaging the usage of the data value. Here, the extra information is afforded called as index that is present exterior to the information content of the relational database and the information of the watermarking is entirely lost if the indexing feature of the relational database table is established again or deleted. Y. Zhang organized the technique that helps to transfer the content of the image into the cloud droplets of the watermark based on D.Y. Li's idea related to the cloud and then

entrenched it into relational data [7]. The data is mined and the droplets of the cloud have to be compared with the source copyright data. Furthermore, Y. Zhang have come with the model of reversible watermarking for relational database [8] where the differences are placed at the relational data base end and it can be stretched by utilizing the wavelet information combined with the entrenched information of the watermark.

D.Boneh enhanced the method of disparity expansion and Lowest-Effective-Bit on integers to attain the entrenching and blind identification of the watermark but this kind of technique is employed only for the type of integer data which is not widespread [9]. Many types of watermarking developers have provided more number of efforts to endorse the progression of watermarking of the database. Still the method is filled with drawbacks in this projected method and it is involved in two features: First type is sturdiness feature of the watermark that is feeble to defy diverse level of database operations and unauthorized attacks of the watermark like assortment, insertion, alteration, etc and the second type is the source relation of the database cannot be renovated from the relation of the watermarking. The result demonstrates that the reversibility and the heftiness of the database watermarking turn to be very hard.

The rest of the sections are organized as follows: section 3 describes the overview of digital watermarking, section 4 describes the materials and methods, section 5 describes the results and discussion, section 6 describes the conclusion and section 7 describes the references.

3. Overview of Digital Watermarking

The purpose of the digital watermarking is the exclusive rights fortification to avert the illegal replication of the relational database. An additional resolution for safeguarding the information is comprised of cryptography and this method has more drawbacks. This type of relational database can be confined by encipherment at the time of broadcasting and it will restored in the source form i.e. plaintext that allows any type of attacker to achieve the access. In the given terms of digital watermarking the watermark can be entrenched and it will retain enduringly in the data and this type of watermark may be perceptible or imperceptible. The imperceptible watermark is more effectual because it scattered over the complete data and not the particular part and hence it is more difficult to remove [10]. This looks like an label which is comprised of information

about the data owner, the end user and the total number of copies. The entrenching of the watermarking contains the necessary requirements:

- Imperceptibility – the watermark can be entrenched and that remains invisible to the human.
- Protection – the mining should be impractical for any type of illegal access even when the embedded algorithm seems to be open.
- Sturdiness – the deliberate or not deliberate watermark elimination should be impracticable without any loss of the source data.

The above mentioned requirements can be satisfied by generating a pseudo noise key. The level of security or protection and the heftiness can be improved by utilizing the non ignorant watermarking methods [3]. The watermark is always based on the source signal in this method and it will be unworkable to accomplish a forgery since there is no admittance to the unharmed data and that is reserved as secret. The technique of watermarking can be implemented in the domain called as spatial or the transform i.e. wavelet, DCT and fractal. The projected method is DCT domain. The imperceptible watermarking can be achieved by the process of entrenching. Figure 2 represents the overall structure of the watermark insertion procedure.

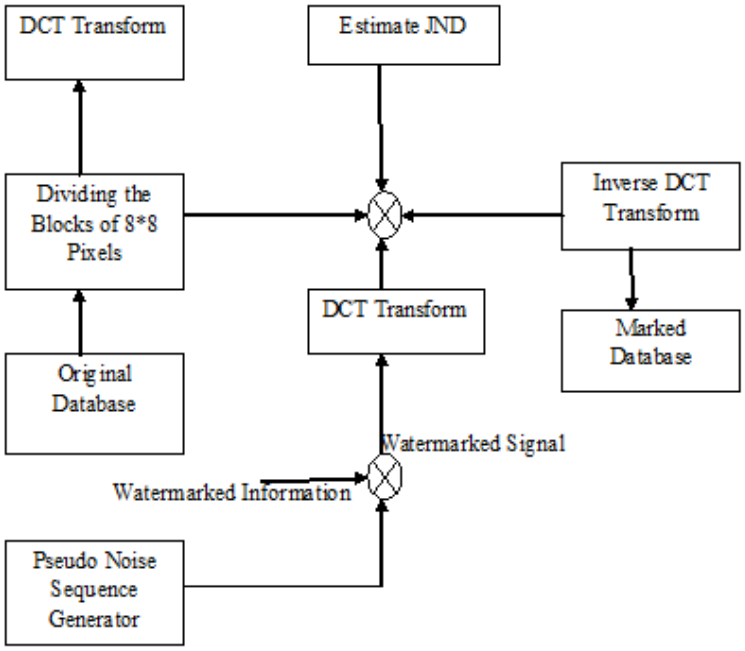

Figure No: 2 Overall Structure of Watermark Insertion Process

The signal of watermarking can be attained by broadening the bit level and by utilizing the definite amount of chip rate medium and the outcomes are transformed by the pseudo noise sequence generator. The computed watermark can be entrenched by evaluating the Just Noticeable Difference (JND) for the source database and its indistinct description to carry out the imperceptible obligation and the output is the marked database.

4. Materials and Methods

4.1 Clustering and Similarity Measurement

In spite of the disruption of the tuples and attributes of the database, inadequate surplus database space combined with the weak heftiness of the broad database watermarking method it is feasible to comprehend the entrenching the database watermarking and vigorous identification with the steady, more effectual and a huge amount of database with tuples clustering concept, which is considered as the foundation of database watermarking technique in this projected method [15]. Hence the similarity measures between the tuples of the database are computed by Mahalanobis distance because it effectually abolishes the persuade of dimension and interference of the correlation. In the same time, there are repeatedly handling database operators available on the attributes and tuples of the database that may seriously affect the heftiness of the database watermarking and the greater part of the decision theory focuses on solving the difficulty when the watermarking is mined. Depending on the clustering of the tuples and the greater part of the decision method a vigorous watermarking structure is shown in figure 3a and 3b.

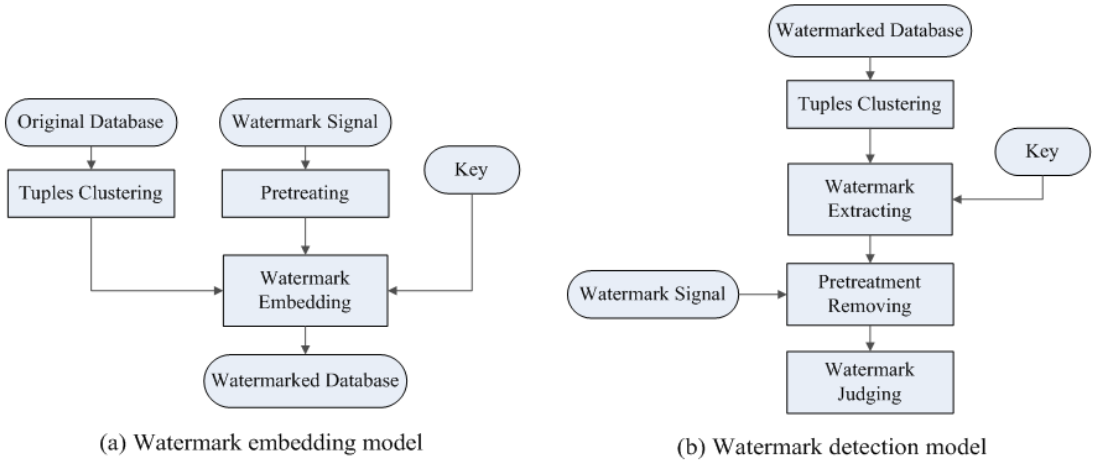

Figure No: 3 Vigorous Watermarking Structure – Tuple Clustering & Decision Oriented

The database information is mostly progressed by the connected application for a extremely accessible database where the source data is reinstated accurately after mining the entrenched watermark and this is termed that the watermark not only contains heftiness and also the reversibility [11]. The blind database detection and the reversibility concept is already studied depending upon the polar angle expression that utilizes the concept of key as a sow to generate the pseudo-random number to choose the position for entrenching the watermark and then mapping the attributes to polar coordinates one by one, and inserts the watermark into the points broadening the polar angle. The type of logistic chaotic method is applied for enciphering the watermark to enhance the security measures of the watermark before entrenching the watermark. The method of LSB is utilized to mine the watermark to attain the blind identification.

By utilizing the above described structure the reversible and vigorous relational database watermarking algorithm is described in detail. The tuples are categorized based on the given number and the definite meaning is represented by the definite category [14]. Then the key feature pseudo random number sow to choose the location of the watermark entrenching in every group and these attributes are connected to polar attributes and the points are stretched by the polar angle and the final step is entrench the watermark by the LSB method.

4.1.1 Procedure for Fast Clustering

Step 1: The set L includes the k number of initial cluster pints.

Step 2: Attainment of the initial classification rule.

Step 3: Compute new cluster points.

Step 4: The fast clustering method is measured by Mahalanobis distance to categorize the database tuples into the preferred categories.

Algorithm

G = fastclustering (k,ε,R)

Input: k→ no. of clustering

ε → Convergence threshold

R → Database

Output: G → Clustering results

- Begin:
- L = randSelect(R);
- d = minDistance(L);
- G = clustering(R,L);
- $d^m = \infty$;
- while ($d^m > \varepsilon d$) then
- L = ClusterPoint(G);
- G = clustering(R,L);
- d^m = maxchangedDistance(L);
- end

The method affords a new type of watermarking procedure depending on fast clustering and the polar angle broadening of relational database where the character confusion between the tuples of the database are clustered by the measure of Mahalanobis distance and it is integrated with the strategy of polar angle expression to entrench and mine the watermark. The method demonstrates the high level of sturdiness under the blind identification for division of selection criteria, insertion and alteration attack and this can reinstate the source data more really.

4.2 Hash Functions, Minimal Pseudo Noise Sequence Generator and DES Algorithm

In this method, the procedure is achieved by a watermark signal based on the sequence of the unharmed database to enhance the security level of heftiness and security [11]. This technique utilizes the hash functions called MD5 algorithm, a Minimal Standard pseudo-noise sequence generator and the cryptographic algorithm called DES (Data Encryption Standard). The watermarking procedures are dependent upon the principle of spread spectrum [13]. The first thing is to embedding that comprises of a pseudo-noise sequence by cryptographically secure key K. The security features and the pseudo-noise sequence producing scheme is listed in the steps given below:

Step 1: The original database is performed by MD5 hashing algorithm and the sequence of 128 bits is achieved;

Step 2: The sequence is segregated into 64 bits each where the odd bits of the initial sequence are copied in the starting of the subsequence and the even sequences will be in the second.

Step 3: The first of the two subsequences is enciphered using DES using the second sequence as the key;

Step 4: The last 32 bits of the resulted sequence will be discarded and a 32-bits number will be obtained;

Step 5: The pseudo-random sequence will be created using a Minimal Standard generator, which will have as a seed the 32-bit number from the step 4. The representation is described in figure 4.

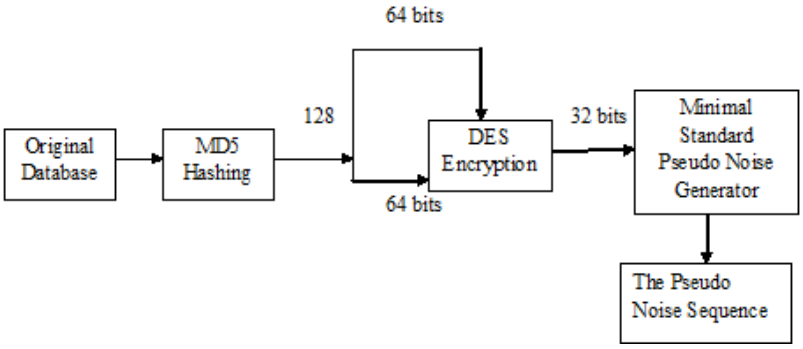

Figure No: 4 Representation of the Projected Method

The type of digital watermarking is incapable in safeguarding the information rights and the opponent is necessary to decide the ownership. Hence in this projected method, the attacker may not be able to generate and mine the false watermarking without providing the information access to the source database which is termed to be secret.

5. Results and Discussion

The mode of enciphering in digital watermarking is the impending feature in relational database by the methods of encipherment (privacy preservation) and watermarking (authentication of data integrity).

Size of the Database	Encryption Time	Embedding Time	Extraction Time	Decryption Time
128*128	7.723	2.977	0.298	3.406
256*256	11.449	6.647	0.416	7.802
512*512	30.784	20.959	0.623	25.574

Table No: 1 Extraction and Cryptographic Functions in Digital Watermarking

The method of digital watermarking is broadly utilized in the relational database for data hiding and protection of data. The method projected a simple relational database scheme of watermarking which is comprised of flexible clustering technique implemented on the tuples of the database. The extraction and cryptographic algorithms of digital watermarking is illustrated in figure 5 and table 1.

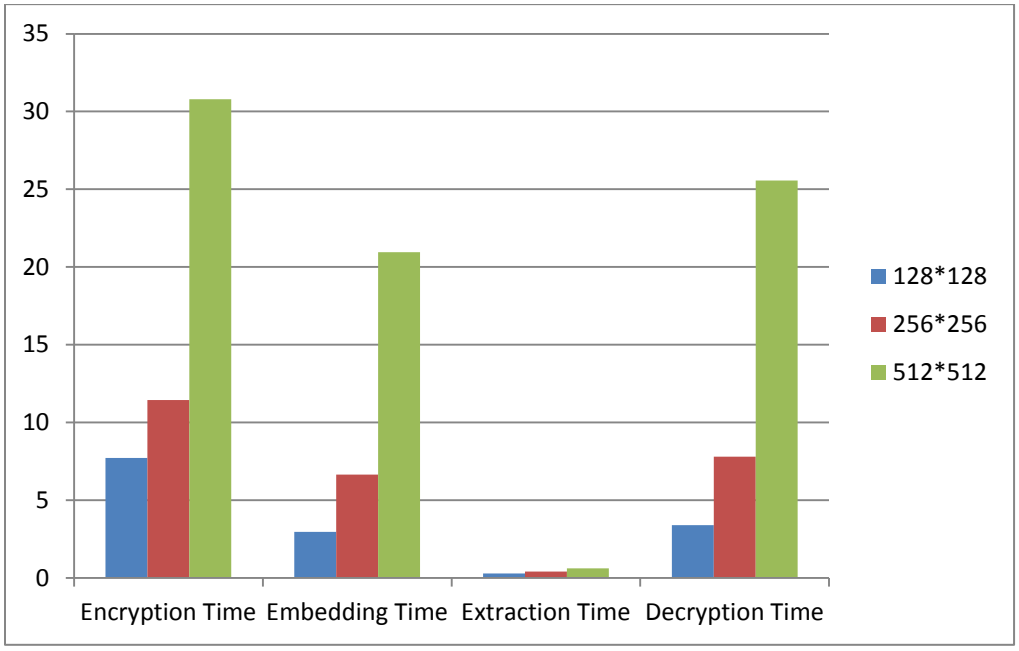

Figure No: 5 Extraction and Cryptographic Functions in Digital Watermarking

The type of digital watermarking is incapable in safeguarding the information rights and the opponent is necessary to decide the ownership. Hence in this projected method, the attacker may not be able to generate and mine the false watermarking without providing the information access to the source database which is termed to be secret.

6. Conclusion

The tuples of the database are clustered into groups adaptively before proceeding the phases of watermark identification and entrenching. The clustered groups are based on the binary watermarking length. Likewise, the fragments of the watermark are entrenched into or identified from the groups respectively based on the numerical data of both the Lowest Significant Bit (LSB) and polar angle expansion. The popular assessment method is employed to examine the bit value of the watermark in the blind identification strategy. These databases can be protected using cryptographic methods. The algorithms are required to entrench it in the data; video, imperceptible data connected to the data owner and existing users of the group. Hence, the watermarked data can be protected by generating a secured cryptographic noise signal. Based on the source data, a watermark signal can be generated and refuse to go along with some planned and unplanned attacks. The security can be enhanced and the attacks can be resolved by hashing and DES algorithms. The first step is to encipher the database with the help of hash functions and DES algorithm. The receiver matches the hash values and then decrypts the original data. In the receiving end, the intruder tries to obtain the original by directly performing the decryption on cipher text of the watermarked data. The result shows that the enciphering and hashing techniques effectively identifies the attackers and safeguard the content using DES.

7. References

[1] Agrawal, A., Kiernan, J., Srikant, R., et al.: 'Order preserving encryption for numeric data'. Proc. 2004 ACM SIGMOD Int. Conf. Management of Data, Paris, France, June 2004, pp. 563–574.

[2] Boldyreva, A., Chenette, N., Lee, Y., et al.: 'Order-preserving symmetric encryption'. Proc. 28th Annual Int. Conf. Advances in Cryptology, Cologen, Germany, April 2009, pp. 224–241.

[3] Boldyreva, A., Chenette, N., ONeill, A.: 'Order-preserving encryption revisited: improved security analysis and alternative solutions'. Proc. 31st Annual Int. Conf. Advances in Cryptology, Santa Barbara, USA, August 2011, pp. 578–595.

[4] Agrawal, R., Kiernan, J.: 'Watermarking relational databases'. Proc. 28th Int. Conf. Very Large Databases, Hong Kong, China, August 2002, pp. 155–166.

[5] Sion, R., Atallah, M., Prabhakar, S.: 'Rights protection for relational data', IEEE Trans. Knowl. Data Eng., 2004, 16, (12), pp. 1509–1525.

[6] Y. J. Li, V. Swarup and S. Jajodia, "Fingerprinting Relational Databases: Schemes and Specialties", IEEE Transactions on Dependable Secure Computing, vol. 2, no. 1, (2005), pp. 34-45.

[7] Y. Zhang, X. M. Niu and D. N. Zhao, "A Method of Protecting Relational Databases Copyright with Cloud Watermark", Proc. World Academy of Science, Engineering and Technology, vol. 3, (2005), pp. 68-72.

[8] Y. Zhang, B. Yang and X. M. Niu, "Reversible Watermarking for Relational Database Authentication", Journal of Computers, vol. 17, no. 2, (2006), pp. 59-65.

[9] D. Boneh, J. Shaw, "Collusion-Secure Fingerprinting for Digital Data, Advances in Cryptology", CRYPTO′95, Springer Verlag, pp. 452-465, 1995.

[10] R. G. van Schyndel, A. Z. Tirkel and C. F. Osborne, "A Digital Watermark", Proc. ICIP'94, vol. 2, (1994), pp. 86-90.

[11] F. Hartung, J.K. Su, B. Girod, "Spread Spectrum Watermarking: Malicious Attacks and Counterattacks", Proc. SPIE, Jan., Vol. 3957, 1999.

[12] N. Nikolaidis, I. Pitas, "Robust image watermarking in the spatial domain", Signal Processing, Vol. 66, pp.385-403, 1998.

[13] B. Schneier, "Applied Cryptography", John Wiley and Sons, 1996.

[14] H. P. Guo, et al., "A Fragile Watermarking Scheme for Detecting Malicious Modifications of Database Relations", Information Sciences, vol. 176, no. 10, (2006), pp. 1350-1378.

[15] S. Bhattacharya and A. Cortesi, "A Generic Distortion Free Watermarking Technique for Relational Databases", Proc. 5th International Conference on Information Systems Security, (2009), pp. 252-264.