Scientific Journal Impact Factor (SJIF): 1.711

International Journal of Modern Trends in Engineering and Research

ISSN (ONLINE) : 2349 - 9745

ISSN (PRINT) : 2393 - 8161

AN EFFICIENT DATA TRANSFER USING FAST ALGORITHM WITH EMBEDDED DIGITAL WATERMARKING

Ms. A. Mythili[1], Mrs. R. Sasiregha[2]

[1]M.Sc., MPhil., Research Scholar,M.Sc., [2]MPhil., Ph.D., Assistant Professor,
Department of Computer Science,
SSM College of Arts and Science, Komarapalayam, Tamilnadu, India

Abstract - The collecting and storing data of all kinds have far outpaced our abilities to analyze, summarize and extract "knowledge". Traditional data analysis methods are no longer efficient to handle voluminous data sets. Thus, the way to extract the knowledge in a comprehensible form for the huge amount of data is the primary concern. Data mining refers to extracting or "mining" knowledge from databases that can contain large amount of data describing decisions, performance and operations. The algorithms also preserve important properties of the dataset are important for mining operations and so guarantee both right protection and utility preservation. This considers a right-protection scheme based on watermarking. Watermarking may distort the original distance graph in the watermarking.

This research work, fast clustering variants drastically reduce the number of calculations of quadratic coefficients as well as the solutions of quadratic inequalities. Prove fundamental lower and upper bounds on the distance between objects post-watermarking. Tight bounds on the contraction/expansion of the original distances. The watermark aims to providing watermarking methodology preserves the FAST property of each object of the original numerical dataset of the image observation. This leads to protection of any mining related process that depends on the ordering of distances among the objects of the dataset, searching and categorization and visualizing techniques as well.

Finally experimental system takes the image observation dataset and applied the FAST classification before embedded watermarking. Then the classified data is supplied for watermarking with image and the watermarked data is transfer. This system shows that the results of the efficiency of right protection of transfer with provable distance based mining.

Index Terms: Digital Watermarking, Fast Algorithm, Privacy Preservation, Nearest Neighbor Classification, Data Security

I. INTRODUCTION

Digital watermarking is the act of hiding a message related to a digital signal within the signal itself. It is a concept closely related to steganography, in that they both hide a message inside a digital signal. What separates them is their goal. Watermarking tries to hide a message related to the actual content of the digital signal, while in steganography the digital signal has no relation to the message and it is merely used as a cover to hide its existence. Watermarking has been around for several centuries, in the form of watermarks found initially in plain paper and subsequently in paper bills. The field of digital watermarking was only developed during the last 15 years and it is now being used for many different applications.

The increasing amount of research on watermarking over the past decade has been largely driven by its important applications in digital copyrights management and protection. One of the first applications for watermarking was broadcast monitoring [10]. It is often crucially important that are able to track when a specific video is being broadcast by a TV station. This is important to advertising agencies that want to ensure that their commercials are getting the air time they paid for. Watermarking can be used for this purpose. Information used to identify individual videos could be embedded in the videos themselves using watermarking, making broadcast monitoring easier. Another very important application is owner identification. Being able to identify the owner of a specific digital work of art, such as a video or image can be quite difficult.

In this case the watermark embedded in a digital work can be used to record one or more transactions taking place in the history of a copy of this work. For example, watermarking could be used to record the recipient of every legal copy of a movie by embedding a different watermark in each copy. If the movie is then leaked to the Internet, the movie producers could identify which recipient of the movie was the source of the leak. Finally, copy control is a very promising application for watermarking. In this application, watermarking can be used to prevent the illegal copying of songs, images of movies, by embedding a watermark in them that would instruct a watermarking compatible DVD or CD writer to not write the song or movie because it is an illegal copy.

II. RELATED WORK

The first irreversible watermarking technique for relational databases was proposed by Agrawal and Kiernan. Similarly, the first reversible watermarking scheme for relational databases was proposed. Difference expansion watermarking techniques (DEW), [5] exploit methods of arithmetic operations on numeric features and perform transformations. The watermark information is normally embedded in the LSB of features of relational databases to minimize distortions. Whereas, in RRW, a GA based optimum value is embedded in the selected feature of the dataset with the objective of preserving the data quality while minimizing the data distortions as a result of watermark embedding. Another reversible watermarking technique proposed in [6] is based on difference expansion and support vector regression (SVR) prediction to protect the database from being tampered. Genetic algorithm based on difference expansion watermarking (GADEW) technique is used in a proposed robust and reversible solution for relational databases [7]. GADEW improves upon the drawbacks mentioned above by minimizing distortions in the data, increasing watermark capacity and lowering false positive rate. To this end, a GA is employed to increase watermark capacity and minimize introduced distortion.

Prediction-error expansion watermarking techniques (PEEW) like [8], [1], [2], [3] incorporate a predictor as apposed to a difference operator to select candidate pixels or features for embedding of watermark information. The PEEW proposed technique by Farfoura and Horng is fragile against malicious attacks as the watermark information is embedded in the fractional part of numeric features only. In [4], the authors proposed a robust, blind, resilient and reversible, image based watermarking scheme for large scale databases. The bit string of an image is used as a watermark where one bit from the bit string is embedded in all tuples of a single partition and the same process is repeated for the rest of the partitions. Gupta and Pieprzyks' [9], proposed reversible watermarking technique introduces distortions as a result of the embedding process. Changes in the data are controlled by placing certain bounds on LSB. On the contrary, to limit the distortions, the data outside the limited bounds is left unwatermarked. As a result, the watermark robustness gets compromised. However, RRW has no such limitations.

The reversible watermarking techniques DEW, GADEW and PEEW, proposed in [7], [8] are not robust and reversible against heavy attacks. Features are selected in these techniques for watermarking without considering their importance in knowledge discovery. RRW is robust and reversible and copes with the above mentioned problems and data quality is preserved by taking into account the importance of the features in knowledge discovery.

III. RIGHT PROTECTION THROUGH WATERMARKING

Describe first how watermarking mechanisms can embed a secret key on a collection of objects. We demonstrate the techniques for 2D sequence data. The later demonstrate how to detect the watermark using a correlation filter.

3.1 Watermark Embedding

Assume an object represented as a vector of complex numbers x = {x1, . . . , xn}, where $x_k = a_k + b_{ki}$ (i is the imaginary unit, i 2 = −1) and where the real and imaginary parts, a_k and b_k respectively, describe the coordinates of the k-th point of object x on the imaginary plain. Such a model can describe data trajectories or even image contour data which capture coordinates of a shape perimeter. The adapt a spread-spectrum approach. This embeds the watermark across multiple frequencies of each object and across multiple objects of the dataset. The robustness of the watermark embedding depends on the choice of coefficients. We embed the watermark in the coefficients that exhibit, on average over the dataset, the largest Fourier magnitudes. This makes the removal of the watermark difficult; masking it out would mean that important frequencies of the dataset will be distorted. This would diminish the dataset utility.

3.2 Watermark Choice

Given dataset D and an even integer 2 ≤ l ≤ n, The focus on the following class of watermarks: The class of watermarks with l non-zero elements, compatible with dataset D, denoted by Wl(D), is the set of all W ∈ {−1, 0, +1} n that satisfy. The DC component captures the center of mass of object x and is therefore highly susceptible to translational attacks. For example, if a part of the watermark were embedded on the DC component of an object then a simple translation would shift the center of mass of the object, thus rendering this part of the watermark useless without affecting the general shape of the object at all. In summary, the embed the watermark in the magnitudes of the Fourier descriptors and leave the phases unchanged; we leave the DC component intact and watermark the Fourier descriptors with the largest average magnitudes.

3.3 Resilience to Transformations

Right-protection mechanism provides resilience to geometric data transformations, such as rotation, translation and scaling. Global object rotation is an intelligent attack because it distorts all coordinates of the objects, but pair wise distances remain the same. Rotation in the frequency domain affects only the phases but not the magnitudes. The watermark being embedded in the magnitude space will remain unaffected. Global translation of all objects only distorts the DC component, in which no part of the watermark was embedded. Scaling attacks can be addressed simply by normalizing all objects/sequences appropriately before watermark detection.

3.4 Watermark Detection

Measure the probability of existence of a watermark by evaluating the correlation between a tested watermark and the right-protected dataset. Measuring directly the correlation between the watermark and the magnitudes of Fourier descriptors may prove ineffective. The reason being that the original level of the average of magnitudes acts as background noise, masking the embedded watermark seek to detect. Address this issue by explicitly recording the bias of average magnitudes before embedding the watermark and removing it before the detection. They also record this bias vector along with the watermark W and both are used jointly as the key. For a dataset D = {x (1), . . . , x(|D|) }, denote as δ x (i) j the magnitude of the jth Fourier descriptor of object x (i) before watermarking. The average of the magnitudes of the jth Fourier descriptor across all objects in D, denoted as µj (D), is given by µj (D) := 1 |D| X |D| i=1 δ x (i) j . (2). Measure the correlation between a watermarked dataset Db and watermark W as follows:

χ(W, Db) := µ(Db) – µ(D) µ(D) !T W,

Where the division is element-wise, excluding elements where μj (D) = 0. In other pixel, remove the bias of average magnitudes before computing the correlation. The scheme enables a very effective detection of the watermark.

IV. FAST NN-PRESERVATION

V.

The state and prove a sufficient condition for preservation of the Nearest Neighbor of an object *x*. The show that if the ratio of the Euclidean distance between *x* and some other object *y* in the original dataset over the distance *D* (*x*,*NN*(*x*)) is greater than or equal to a threshold depending solely on *p* max, then *y* does not violate the NN of *x* after the water- mark embedding, regardless of the details of the dataset or the watermark embedding. Proposed system provides two variants. One that preserves Nearest Neighbors and another that preserves the Minimum Spanning Tree. Therefore, the output of any algorithm based on these two properties will be preserved after right protection. To guarantee this, they study the critical watermark intensity to protect the dataset as well as ensure that, important parts of the object distance graph are not distorted.

It is essential to discover the highest watermark concentration for right protection. This provides guarantee of better detectability and hence enhanced security for the right protection scheme. Here they first study how distances between the objects are imprecise as a function of the watermark embedding strength. This gives imminent on how to design fast variants of algorithms that still guarantee preservation of the Nearest-Neighbors and the Minimum Spanning Tree, but it operate significantly faster than the exhaustive algorithms.

VI. DATA TRANSFER USING WATERMARK EMBEDDING TECHNIQUE

The state and prove a sufficient condition for preservation of the Nearest Neighbor of an object *x*. The show that if the ratio of the Euclidean distance between *x* and some other object *y* in the original dataset over the distance *D* (*x*,*NN*(*x*)) is greater than or equal to a threshold depending solely on *p* max, then *y* does not violate the NN of *x* after the watermark embedding, regardless of the details of the dataset or the watermark embedding. The system provides two variants. One that preserves Nearest Neighbors and another that preserves the Minimum Spanning Tree. Therefore, the output of any algorithm based on these two properties will be preserved after right protection. To guarantee this, they study the critical watermark intensity to protect the dataset as well as ensure that, important parts of the object distance graph are not distorted.

It is essential to discover the highest watermark concentration for right protection. This provides guarantee of better detectability and hence enhanced security for the right protection scheme. Here they first study how distances between the objects are imprecise as a function of the watermark embedding strength. This gives imminent on how to design fast variants of algorithms that still guarantee preservation of the Nearest-Neighbors and the Minimum Spanning Tree, but it operate significantly faster than the exhaustive algorithms.

ALGORITHM: 1

Fast NN-Preservation Algorithm For Image Embed Process

Input: Image, Watermark Data

Output: Embedded Image, Extracted Watermarked Data

1. Select the Image (I).
2. Enter watermark content (W).
3. Find the Pixels (PG-Pixel Group) where Red component values falls from 137 to 147, 157 to 167 and 187 to 207.
4. Convert the watermark data to bytes and find the length of watermark data (L).
5. In the first pixel of the PG, store the ‘L’ value in Blue component.

6. Then store all the ‘W’ bytes in Blue component one by one starting from second pixel to all other successive pixels in PG group until the ‘W’ bytes are completed.

Extract Steps:
Input: Embedded Image
Output: Extracted Image, Extracted Watermarked Data

7. Select the Image (Watermark Embedded) (I).
8. Find the Pixels (PG-Pixel Group) where Red component values falls from 137 to 147, 157 to 167 and 187 to 207.
9. From the first pixel of the PG, get the ‘L’ value from the Blue component.
10. Fetch all the ‘W’ bytes from Blue components one by one starting from second pixel to all other successive pixels in PG group until the ‘W’ bytes are completed.
11. Convert the watermark bytes to data.
12. Check the KNN Property.
13. Output Watermark Data.

ALGORITHM: 2
Fast NN-Preservation Algorithm for Numeric Data Set Embedded Process
Input: Patient Observations, Watermark Data
Output: Modified Patient Observation Data

1) Add the Patient Profiles (P).
2) Add the Patient Observation Data (O).
3) Enter watermark content (W).
4) Convert the watermark data to bytes and find the length of watermark data (L).
5) Sort the Patient Observation Data (O) Patient wise.
6) I=0
7) For Each Patient’s Observation Set in (O)
8) Alter the Observation Data’s third value such that OD(3) = 301 + W(I)
9) Change the OD(1) position = OD(1) position + W(I)
10) I=I+1
11) If I>=L Then
12) Break
13) End If
14) Next
15) Output the New Patient Data Set.

Extract Steps:
Input: Modified Patient Observation Data
Output: Patient Observation Data, Extracted Watermarked Data

16) Select the Patient Data Set (where Watermark Data Embedded) (P).
17) I=0;
18) For Each Patient’s Observation Set in (O)
19) W(I) = Observation Data’s third value - 301
20) Change the OD(3) value= OD(3) value -301
21) If I=0 Then
22) L= W(I)
23) End If
24) I=I+1
25) If I >L Then
26) Break
27) End If

28) Next
29) Convert the watermark bytes to data.
30) Check the KNN Property.
31) Output Watermark Data.

Large color images obtained from two algorithm were used to compare the proposed algorithm and the conventional K-means algorithm. The algorithms are implemented with Visual Studio running on a windows7. The terminating condition İ was set to 0.1 for both algorithms. Due to the limitation of space, here we only give two representative examples. 2000*3000 color image, in which the white, black and other regions represent MRI scan and Brain, respectively.

VI. PERFORMANCE ANALYSIS

6.1. Comparison OF KNN-Algorithm And Fast –Algorithm

The following Graph is explain of comparison of KNN –Algorithm And Fast Algorithm.

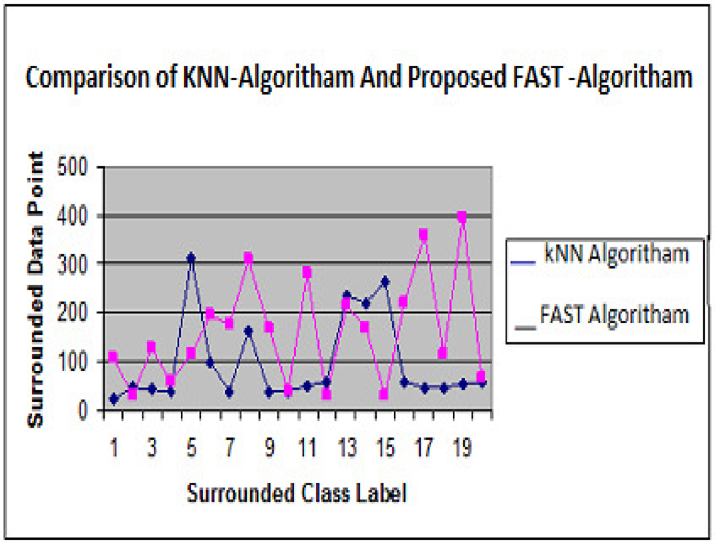

Fig No: 6.1. Comparison of KNN & FAST Algorithms

The following Fig 6.1.describes experimental result for comparison of existing and proposed system analysis in patient data set plotted in graph using KNN and FAST algorithm Class A. The table contains patient id, the data point in surrounded data point and its class label details. In this table, it describes the class label surrounded in patient data set in Enhanced FAST algorithm work with water marking details.

6.2. Image Noise Ratio

Signal to noise ratio is been calculated to find the objective of the non-constant intensity object. The value of the FAST method shows that the noise is removed higher than the KNN and MST model. Since SNR value for Fast method is low the meaningful information is higher in biomedical MRI Image.

METHOD	IMAGE NOISE RATIO
KNN	32.8172
MST	32.0719
FAST	29.6431

Table: 6.1. Image Noise Ratio

SNR refers to the signal strength. Signal refers to the information. The information such as pixel value, additive intensity value, standard deviation value which cannot be measured accurately when there is noise, can be obtained after finding the SNR value. The less, the SNR value the more, the information can be retrieved. The method is considered to be best if its SNR value is less. The experimental result shows that fast method performs better than KNN model and MST. MST model is an extension of KNN model.

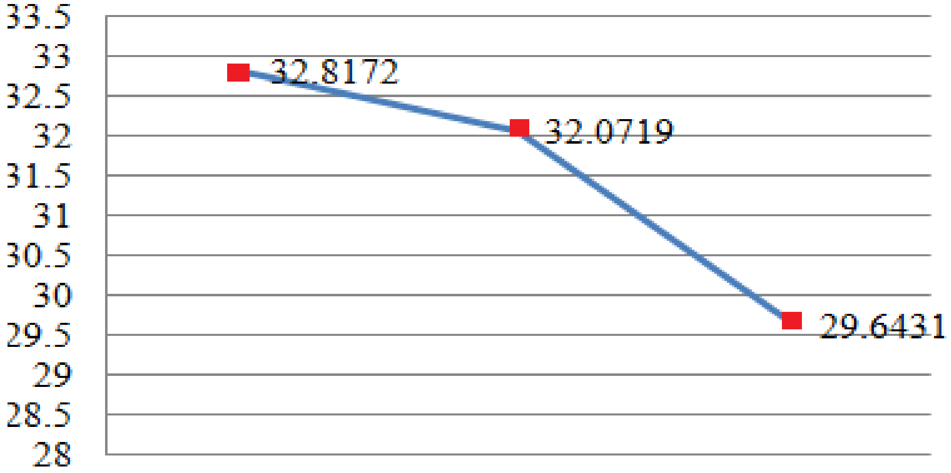

Fig No: 6.2. FAST method performed better then KNN and MST

VII. CONCLUSION AND FUTURE WORK

Watermarking approach is applied over the numerical data sets. The novel K-Nearest Neighbor approach is used in this proposed methodology without modifying the process of existing KNN approach. The patient observation data set is used for the experimental study of the proposed system. Unlike the ancient approach, the watermarking is applied in both image and numeric dataset as well. And also the data about the receiving user also embedded along with the watermarking data. The proposed novel watermarking methodology preserves the nearest neighbor’s property of each object of the original numerical dataset of the patient observation. This leads to protection of any mining related process that depends on the ordering of distances among the objects of the dataset, searching and categorization and visualizing techniques as well. The experimental system takes the patients observation dataset and applied the KNN classification before watermarking. Then the classified data is supplied for watermarking with image and the watermarked data is published. The experimental system shows that the results of the efficiency of right protection of publishing with provable distance based mining.

The watermarking system can be enhanced with the following features. The various applications of this methodology includes data outsourcing like portfolio analysis, government’s sensitive data, defense force data, business promotion applications and banking transactions analytics etc. The Nearest Neighbor approach can be improved with necessary alteration for the better and efficient classification with fast manner. Involvement of the end user needs to be minimized in K-NN with right protection algorithms. Analysis will be used to design AI based algorithms for MST-preserving watermarking that drastically prunes the vast search space.

REFERENCES

[1] D. M. Thodi and J. J. Rodriguez, “Expansion embedding techniques for reversible watermarking,” IEEE Trans. Image Process., vol. 16, no. 3, Feb. 2007.

[2] M. E. Farfoura, S.-J. Horng, J.-L. Lai, R.-S. Run, R.-J. Chen and M. K. Khan, “A blind reversible method for watermarking relational databases based on a time-stamping protocol,” Expert Syst. Appl., vol. 39, no. 3, pp. 3185–3196, 2012.

[3] X. Li, B. Yang and T. Zeng, “Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection,” IEEE Trans. Image Process., vol. 20, no. 12, pp. 3524–3533, Dec. 2011.

[4] E. Sonnleitner, “A robust watermarking approach for large databases,” in Proc. IEEE First AESS Eur. Conf. Satellite Telecommun., 2012, pp. 1–6.

[5] G. Gupta and J. Pieprzyk, “Database relation watermarking resilient against secondary watermarking attacks,” in Information Systems and Security. New York, NY, USA: Springer, 2009, pp. 222–236.

[6] J.-N. Chang and H-C. Wu, “Reversible fragile database watermarking technology using difference expansion based on SVR prediction,” in Proc. IEEE Int. Symp. Comput., Consum. Control, 2012, pp. 690–693.

[7] K. Jawad and A. Khan, “Genetic algorithm and difference expansion based reversible watermarking for relational databases,” J. Syst. Softw., vol. 86, no. 11, pp. 2742–2753, 2013.

[8] M. E. Farfoura and S.-J. Horng, “A novel blind reversible method for watermarking relational databases,” in Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl., 2010, pp. 563–569.

[9] G. Gupta and J. Pieprzyk, “Reversible and blind database watermarking using difference expansion,” in Proc. 1st Int. Conf. Forensic Appl. Tech. Telecommun., Inf., Multimedia Workshop, 2008, p. 24.

[10] Saman Iftikhar, M. Kamran and Zahid Anwar, “RRW—A Robust and Reversible Watermarking Technique for Relational Data”, IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 4, April 2015.