International Journal of Recent Trends in Engineering & Research
ISSN(ONLINE) : 2455-1457
IMPACT FACTOR : 3.344

CLASSIFICATION OF SEMATICALLY SECURED ENCRYPTED RELATIONALDATA USING DATAMINING TECHNIQUES

Dr.P.Banumathi[1], T.Jagadesh kumar[2], M.Kungumaraj[3], G.Preetha[4]
[1,2,3,4] *Information Technology, Kathir College of Engineering*

Abstract-Exchanging and publishing data are becoming an inherent part of business and academic practice, which in many areas that have been obtained after exorbitant and challenging procedures, that can provide detectable evidence for the legal ownership of a shared dataset, without the mutual consessions and its usability under a wide range of machine learning, mining, and search operations. The algorithms also preserve great significant properties of the dataset, which are important for mining operations, and so giving warranty for both right protection and utility preservation. The project considers a definite-protection statergy based on watermarking. Watermarking may control the original distance graph. The proposed watermarking methodology which safeguard the related data that are far apart. This leads to preservation of any mining operation that depends on the ordering of distances between objects. It proves fundamental lower and upper hops on the distance between objects post-watermarking.In particular, it establishes a certain limits for equally dimensed property.This analysis used to plan fast algorithms for NN-preserving watermarking that unpleasant vast search space.

Keywords-Security,Data mining,watermarking classifier.

I. INTRODUCTION

Data mining is the process of finding correlations or patterns among dozens of fields in large relationship data bases with iterative computer-assisted process of analysing enormous sets of data and then extracting the meaning of the data as the fastest growing fields in the computer industry. One of the greatest strengths of data mining is reflected in its wide range of methodologies and techniques that can be applied to a host of problem sets.It follows some techniques and trends. Machine Learnig technique is used to test the database of the system automatically, after the system gets well-trained by the user.

II. MODULE DESCRIPTION

The following modules are present in the project.

1. Patient profile addition
2. Patient observation entry
3. Watermark content addition
4. Embed watermark data in patient observation numeric data set
5. Extract watermark data in patient observation numeric data set after knn check
6. Image addition
7. Embed watermark data in patient observation image
8. Extract watermark data in patient observation image after knn check

1. Patient Profile

In this module, the patient details such as patient id, name, gender, date of birth, entry date, address, mobile, occupation, annual income, father name, guardian name, symptom and previous treatment taken details are keyed in and saved in to ‘Patients’ table.

2. Patient Observation Details

In this module, the patient id is selected, entry date becomes today date, test observation data 1, test observation data 2 and test observation 3 are keyed in and saved in to ‘Observations’ table.

3. Watermark Content Addition

In this module, the watermark content details are added. The details are saved in 'Watermark' Table.

4. Embed Watermark Data In Patient Observation Numeric Data Set

In this module, the watermark content details are converted into bytes and stored in patient observations third column along with a numeric value 301 is added. The first observation values are taken as X position and second observation values are taken as Y position and are pointed initially. Then the first watermark byte value is added with X and then the X Position is modified. This repeats for all watermark bytes (each one watermark byte is stored in one patient's all observation data). The process is listed in embed steps of algorithm 1 of algorithms section.

5. Extract Watermark Data In Patient Observation Numeric Data Set After Knn Check

In this module, the modified patient data set is taken and record values for third observation column values greater than 301 is filtered out. Then for each patient, the third observation column value is subtracted with 301 and the numeric value's character is found out (ascii value). This repeats for all the patients and watermark content is appended. The result watermark content is displayed. The KNN value is found out before and after watermarking and checks for same result display. The process is listed in extract steps of algorithm 1 of algorithms section.

6. Image Addition

In this module, the image is browsed and saved in 'Images' table.

7. Embed Watermark Data In Patient Observation Image

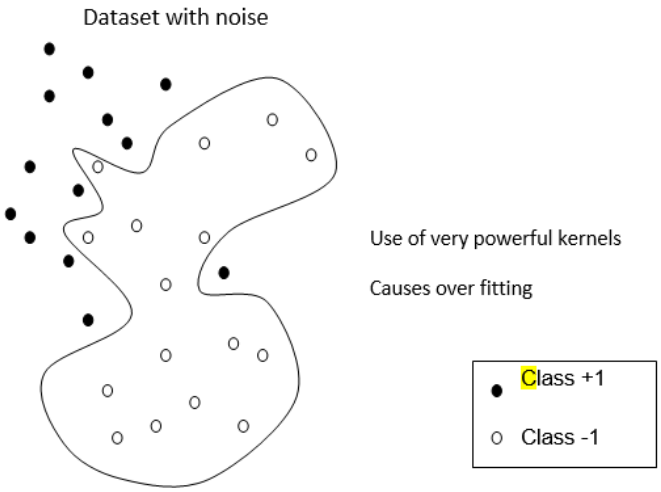
In this module, the watermark content details are converted into bytes and stored in last image browsed and saved in project folder. The process is listed in embed steps of algorithm 2 of algorithms section.

8. Extract Watermark Data In Patient Observation Image After Knn Check

In this module, the watermark content details are found out from the embedded image. The process is listed in extract steps of algorithm 2 of algorithms section.

III. EXISTING SYSTEM

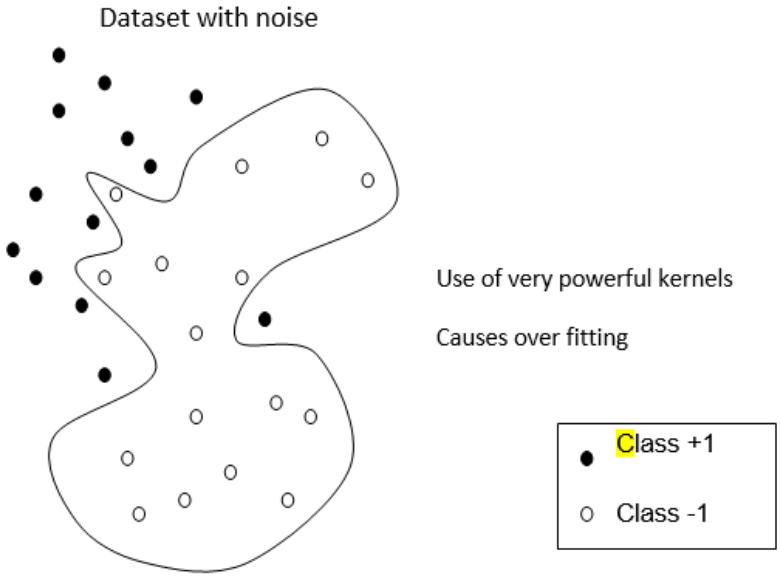
The existing system uses a spread-spectrum approach. This embeds the watermark across multiple frequencies of each object and across multiple objects of the dataset. As such, it renders the removal of the watermark particularly difficult without substantially compromising the data utility. The data locations are altered before applying the watermarking.



The robustness of the watermark embedding depends on the choice of coefficients. The watermark is embedded in the coefficients that exhibit, on average over the dataset, the largest Fourier magnitudes. This makes the removal of the watermark difficult. However, during the data distribution, all kind of receiver users receive the same data.

IV. PROPOSED SYSTEM

Like the existing system, proposed system also uses watermarking without altering the KNN property. In addition, numeric data set is choosen for applying the watermark without altering the KNN property. In addition, if watermark data is corrupted, it can be found out.



V. CONCLUSION

The project uses watermarking without altering the KNN property. Numeric data set is chosen for applying the watermark without altering the KNN property. In addition, if watermark data is corrupted, it can be found out. Watermarking is applied in both image and numeric data set. Data about the receiving user is also embedded in watermarking information. Watermark corrupted information can be found out. The proposed watermarking methodology preserves the Nearest Neighbors (NN) property of each object of the original dataset. This leads to preservation of any mining operation that depends on the ordering of distances between objects, such as NN-search and classification, as well as many visualization techniques. It proves fundamental lower and upper bounds on the distance between objects post-watermarking.

REFERENCES

1. P. Das, N. R. Chakraborti, and P. K. Chaudhuri, “Spherical mini-max location problem ,” Comput. Optim. Appl., vol. 18, no. 3, pp. 311–326, 2001
2. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” IEEE Trans.Image Process., vol. 6, no. 12, pp. 1673–1687, Dec. 2000.

