Academic Science

International Journal of Innovations & Advancement in Computer Science
IJIACS
ISSN 2347 – 8616
Volume 4, Special Issue
September 2015

Identifying the Networks of Criminals Using Management Information System

Vivek Kumar Prasad
Nirma University,Ahmedabad

Matang Ramdevputram A.
Nirma university, Ahmedabad

ABSTRACT

With this proposed solution we can identify the actual criminals and his/her close associates which can be treated as a supported suspected criminals and it can be also used for HR people to select the employees for their company, while checking their criminal background as a result the criminals background candidates will not be selected as NORA used to do. With the help of this MIS (Management Information system) and the proposed techniques and database of criminal's personal information can be used to identify the associated criminals and not the innocent one. Here we are adding an additional layer that can filter the associated criminals and also imposed security onto the database.

Keywords: Management Information system, Nonobvious relationship awareness.

Introduction

Our proposed system will match the person's data or information from the database maintained for the criminals, like if any person or candidates who is eligible for the jobs or the passenger that will be traveling in plane or trains or etc, their data will be crossed checked with the criminal records .So that the criminals can be identified in the very first instant and issue an alert. Though the data should be used for public awareness and security in data has to be imposed and database should be kept encrypted and here we are using the RSA[1][2] algorithms and also the concept of the PPDM [3](privacy preserving in data mining techniques) can be used, so that the data can be fetch by the authenticated person only. Database should contain the information like fingerprint, face outlook, signature, Eye scanning, Phone numbers, list of number of time that person contacted particular person or how closely associates with the criminal back ground person. Our main intension is to catch the main culprit and innocent person cannot be arrested.
Research also shows that Systems Research & Development (SRD) developed its Non-Obvious Relationship Awareness (NORA)[4] technology to help casinos identify cheaters by correlating information from multiple sources about relationships and earlier transactions.

Problem Statement

Their might be several disadvantages using these kind of the techniques like if the database of the criminals went to the wrong hands then the misuse of the data can be done and the data cannot be leaked to the unauthorized person.
Concepts like , KDDM[5] poses a threat to privacy, in the sense that discovered patterns classify individuals into categories, revealing in that way confidential personal information with certain probability. Moreover, such patterns may lead to generation of stereotypes, raising very sensitive and controversial issues, especially if they involve attributes such as race, gender or religion. An example is the debate about studies of intelligence across different races.
The exploratory KDDM tools may correlate and disclose confidential, sensitive facts about individuals. For instance, a central task in KDDM is inductive learning; this takes as input a training data set and produces as output a model (called as a classifier) which is then applied to new, unseen cases to predict some important and perhaps confidential attribute (for example, customer buying power or medical

diagnosis). The classifiers are typically very accurate when applied to cases from the training set, and can potentially be used to compromise the confidential properties of these cases.
Also, knowledge of totals and other similar facts about the training data may be correlated to facilitate compromising individual values, either with certainty or with a high probability.
Another problem is that many a times it will not give the accurate results as it depends of the data resources which results into the arrest / catching of the innocent person and other issue is the scalability to hold large amount of the data.

Proposed Solution

Our main goal here is that innocent people should not be tracked or trapped along with the actual culprit, just because he/she was in the contact list with the actual criminal. Our approach is to make use of HMM (Hidden Markov Model)[6] [7] and classify the type of relation the innocent person is having with the actual criminal.
An HMM is a double embedded stochastic process with the two hierarchy level. It can be used to model much more complicated stochastic processes as compared to a traditional Markov model. An HMM has a finite sets of states governed by a set of transition probabilities. In a particular state, the outcome or an observation can be generated according to an associated probability distribution. It is only the outcome and not the state that is visible to an external observer. HMM-based applications are common in various areas such as speech recognition, bioinformatics, and genomics. Even the HMM can also be used in anomaly detection. Also they can be used to classify TCP network traffic as an attack or normal using HMM

Here we are assuming that if the person A is in close relation with the person B, then in B's call list the Person A's number will be repeated more times and if the relation is not bounded for a long time then average and once in a while will be denoted as least.

The classification can be done based upon the following table 1.0, which shows the call list of the person A:-

Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	Day10
X	X	Y	Y	X	Z	Y	X	X	X

Table 1.0

After counting the data we have come up with the following data

A	X (6 times calls has been reported)
A	Y (3 times calls has been reported)
A	Z (1 times calls has been reported)

And the range can be denoted with respect to some levels like, Close relation if the number of calls in last 10 days are in between 1 to 6 and average relation if the number of calls will be from 1 to 3 and least relation can be 1 to 2. Here we are considering only three person X, Y ,Z associated with A, in actual scenarios it can vary and the data could be changed according to the scenario.

Then the above figure can be again changed to the following figure1.0

International Journal of Innovations & Advancement in Computer Science
IJIACS
ISSN 2347 – 8616
Volume 4, Special Issue
September 2015

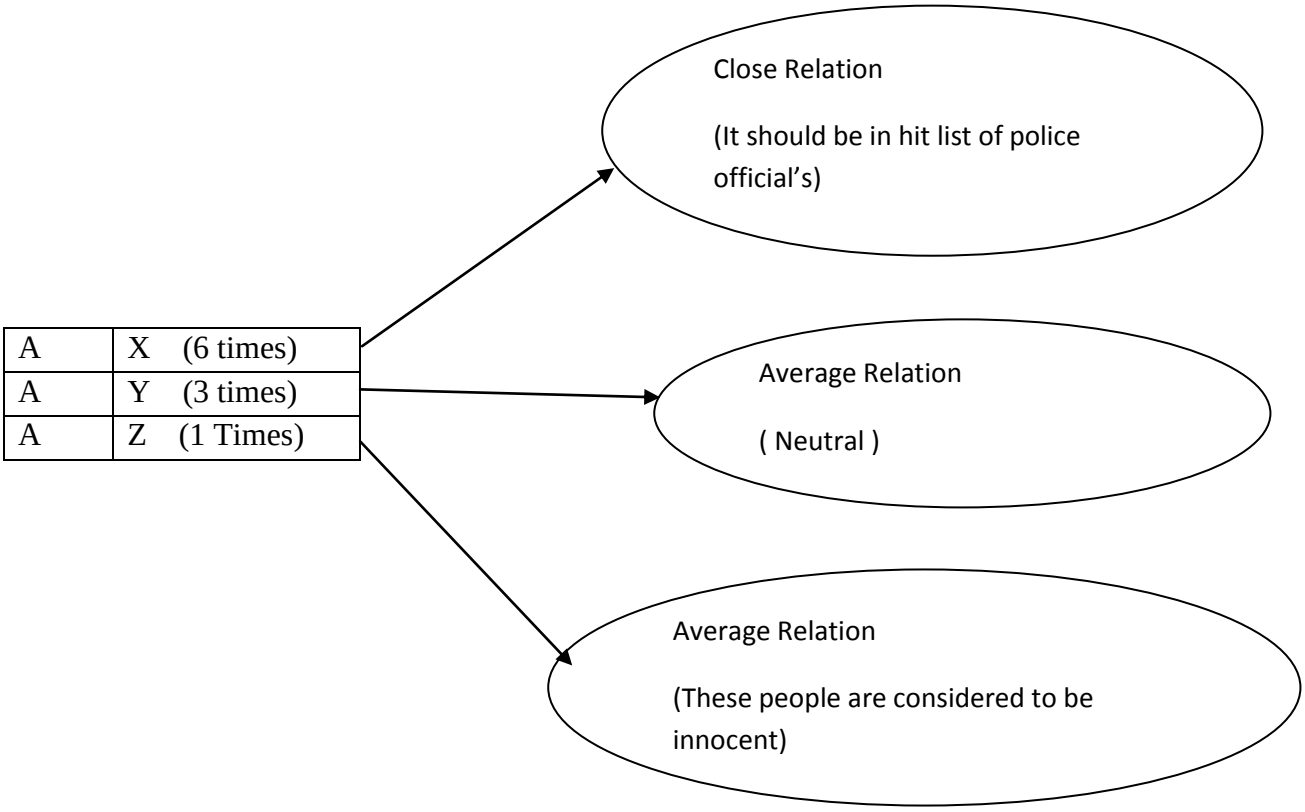

Figure 1.0

Next ,if the labels has been classified ,it can be incorporated into the system and can be added as an additional layer into the existing system as shown below, so that it will give idea to the police officials to suspect upon few people instead of suspecting on more number of people which will make their task tedious and can divert their way of investigation. The following figure 1.1 shows that once the data of the transaction system and the data of the criminals are matched then the criminal associates can be identified with more accuracy.

International Journal of Innovations & Advancement in Computer Science
IJIACS
ISSN 2347 – 8616
Volume 4, Special Issue
September 2015

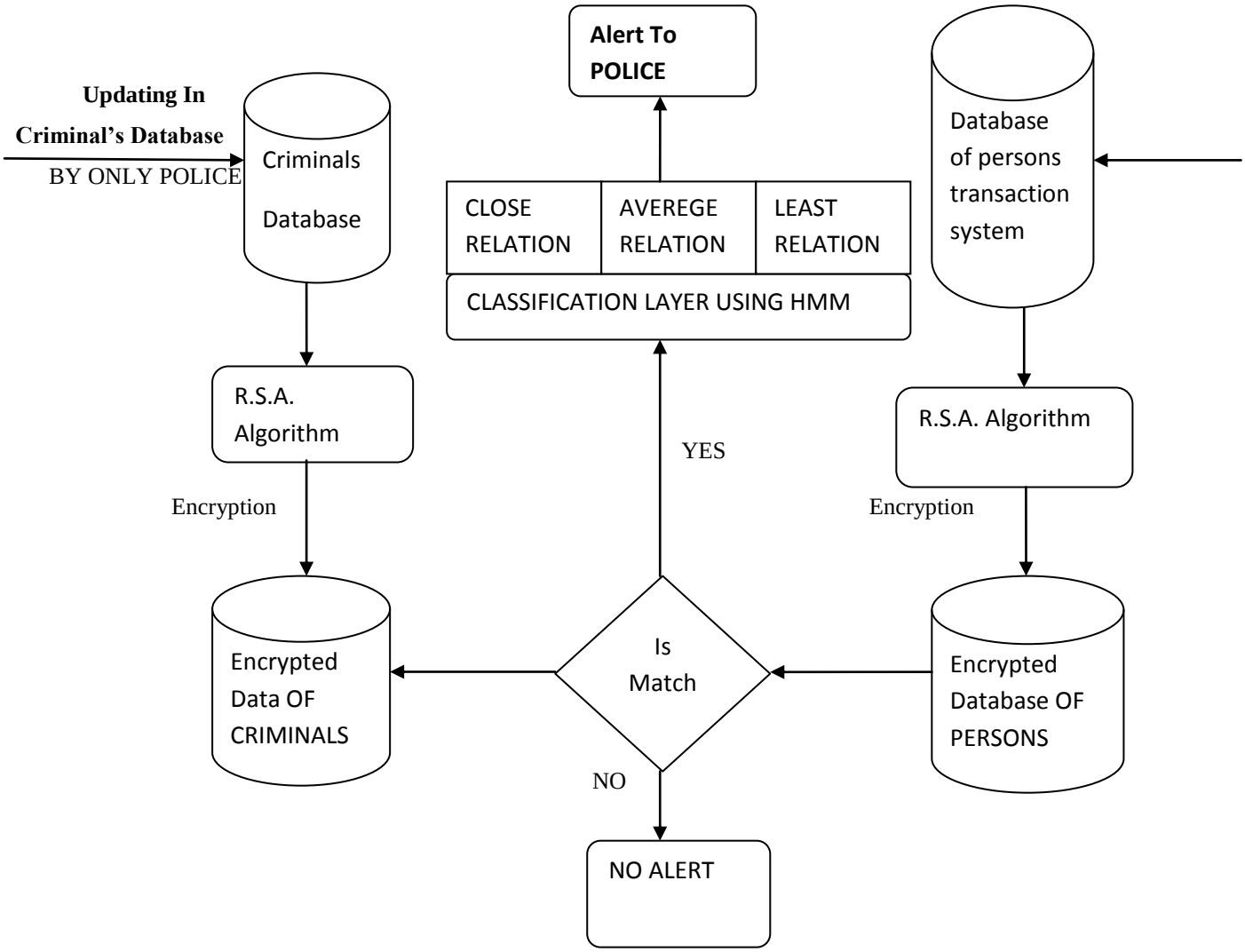

Figure1.1 Data flow diagram of the system

Future work

The Data Mining concepts and machine learning concepts can be implemented and K-means clustering [8] can be implemented to make these relationship more powerful . Cluster analysis is a statistical technique used to generate a category structure which fits a set of observations. The groups which are formed should have a high degree of association between members of the same group and a low degree between members of different groups. While cluster analysis is sometimes referred to as automatic classification, this is not strictly accurate since the classes formed are not known prior to processing, as classification implies, but are defined by the items assigned to them and pattern recognition, Pattern recognition[9] is the science of making inferences from perceptual data, using tools from statistics, probability, computational geometry, machine learning, signal processing, and algorithm design. Thus, it is of central importance to artificial intelligence and computer vision, and has far-reaching applications in engineering, science, medicine, and business.

Conclusion

Our approach is to hide the database from the unauthorized users and to show the strong relationship between the criminals and his close associates, so that the innocent people who so ever has come in contact with the criminal for few times knowingly or unknowingly should not be disclosed as an associate to the actual criminal. The mentioned work can be very useful for the police officials to suspect upon few people instead of suspecting on more number of people which will make their task tedious and can divert their way of investigation.

References

[1] E1Gamai T. A public-key cryptosystem and a signature scheme based on discrete logarithms .IEEE Transactions on Information Theory, 1985, 31 (4) :469-472

[2] Chen Dong, Jiang Zhaogen. "The Application of DSA and RSA cipher System in E-commerce. Computer".Computer Technology and Application, 2002. Computer Technology, 2004.

[3] Jaideep Vaidya, Chris Clifton, "Privacy Preserving Naive Bayes Classifier for Vertically Partitioned Data", SDM2004, SIAM, pp.522- 526.

[4] A.-L. Barabasi, Linked: How Everything is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. Plume, 2003

[5] Julisch, K. & Dacier, M. (2002), Mining Intrusion Detection Alarms for Actionable Knowledge, The 8th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD), pp. 366-375, Edmonton, Alberta, Canada.

[6] S.S. Joshi and V.V. Phoha, "Investigating Hidden Markov Models Capabilities in Anomaly Detection," Proc. 43rd ACM Ann. Southeast Regional Conf., vol. 1, pp. 98-103, 2005

[7] C. Phua, D. Alahakoon, and V. Lee, "Minority Report in Fraud Detection: Classification of Skewed Data," ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 50-59, 2004.

[8] H. Xiong, J. Wu, and J. Chen, "K-means clustering versus validation measures: A data distribution perspective," in Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2006, pp. 779–784

[9] Rueda L .G., Oommen B.J. On optimal pairwise linear classifiers for normal distributions: the two dimensional case [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on , 2002-24(2): 274280.