Spirit of Identity Fraud And Counterfeit Detection

M.Swathi[1], K.Kalpana[2]

[1](II Year M.tech , Department of CSE, G.Pulla Reddy Engineering College, Kurnool, Andhra Pradesh, India)

[2](Associate Professor, Department of CSE, G.Pulla Reddy Engineering College, Kurnool, Andhra Pradesh, India)

Abstract*—Now-a-days the shopping as been evolved as70% online and 30% as offline. In this 30% also 20% have been paid by smart cards. This smart cards usage have found wide spread due to flexible usage. The applications for these credit cards are based on internet or manual applications by the customers who require the smart cards and various loans. The applications in above cases found fraud is a specific case of identity crime. The application fraud pattern is represented by some specific features which may be found duplicates relative to the established base of some criteria. In the existing system we use Communal Detection (CD) i.e. Communal detection Spike Detection (SD) i.e. Spike Detection. Communal Detection (CD) finds real social relationships to reduce the suspicion score and tamper resistant to synthetic social relationships. It is white list-oriented approach on a fixed set of attributes. Spike Detection (SD) finds spikes in duplicates to increase the suspicion score, and is probe-resistant for attributes. The existing system detects the whether the applicant is fraud. It is the attribute-oriented approach on a variable-size set of attributes. In the existing system the fraudster datum nre stored in the database manually. In this proposed system, Communal Detection (CD) and Spike Detection (SD) can detect more types of attacks, better account for varying legal actions, and remove the redundant attributes and to store the fraudulent datum in blacklist using CBR algorithm. CBR algorithm analysis using retrieval, diagnosis and resolution to make the data more secure and to find the fraudulent data. The data that already present or fraudulent is encountered and thrown into the blacklist. Together Communal Detection (CD), Spike Detection (SD) and CBR ensure the data provided by the customer is original. This proposed system makes the system more efficient and enhance the security.

***Keywords*— Communal Detection, Spike Detection, Case Based Reasoning, fraud detection**

I. INTRODUCTION

The data mining consists of multiple algorithms for detection. Data mining algorithms are used in the online credit card application for counterfeit detection. The algorithms are used in this system is the spike detection, communal detection and CBR algorithms. These algorithms are used to detect the fraud and save the data in the database as original data or blacklist database. This system updates the database manually. This system does not give a chance to defaulters in credit card application. Identity crime is defined as largely as feasible in this method. At one extreme, real identity theft refers to illegal use of innocent people's complete characteristics. These can be harder to obtain (although large volumes of some identity data are widely available) but easier to successfully apply. In actuality, identity crime can be committed with a mix of both synthetic and real identity details. Credit applications are Internet or paper-based forms with written requests by potential customers for smart cards, mortgage loans, and personal loans. Credit application deceit is a specific case of identity crime, involving synthetic identity fraud and real identity theft. As in identity crime, the credit application fraud has reached a critical mass of defaulters who are highly experienced, organized, and sophisticated. There are two types of duplicates: exact (or identical) duplicates have the all same values; near (or approximate) duplicates have some same values (or characters), some similar values with a little changed spelling, or both. In short, the new methods are based on White-listing and Detecting spikes of similar applications. White-listing uses existent common relationships on a fixed set of attributes. This reduces false positives by lowering some suspicion scores. Detecting spikes in duplicate, on a variable set of attributes. This increases true positives by adjusting suspicion scores appropriately.

A. Data Mining

Data mining, the extraction of hidden predictive information from huge number of records, is a powerful innovative technology with great likely to help companies focus on the most important processed data in the data warehouses. Data mining tools forecast future trend and behavior, allowing businesses to make proactive, decision support systems. The administrator verifies the provided data with the existing datum to find whether it is defaulter or original. If the data is original, it will be added to the database otherwise it will be placed into the blacklist.

B. CBR Algorithm

Case-based reasoning (CBR) is now making a significant contribution to the task of crime detection. CBR systems are able to learn from sample patterns of smart card use to categorize new cases, and this method also has the promise of being able to adapt new patterns of crime as they come forward. The CBR system is the application of adaptive and hybrid culture systems. The CBR problems are previously considered too dynamic, chaotic, or complex to precisely model.

C. *Problem Statement*

The online credit card application is internet based forms. This system detects the default applicant using the data mining algorithms. The existing method used two algorithms they are spike detection and communal detection algorithm. These two combine together detects whether the applicant is default or original. The proposed system combing with these algorithms CBR uses to find the deceitful data and puts it into the blacklist. This system uses for the match analysis from the existing blacklist records to make the method efficient and secure

D. Objective of the work

Data mining is concerned with analysis of large volumes of data to discover without human intervention interesting regularities or relationships which in turn leads to better understanding of the underlying processes. The data mining consists of multiple algorithms, some algorithms uses for the detection of crime in smart card. Online credit card application uses these algorithms communal and spike detection uses to detect the multiple applicant and with the artificial intelligent CBR algorithm uses to make the fraudulent data in the black list.

II. EXISTING SYSTEM

The credit card application the system detects whether the applicant is fraud or original. The existing system implements two data mining layers they are communal and spike detection. The Communal Detection is used to find the suspicious data of the fraudulent people. It also used to find the communal relationship that are near to reflect the family bond.(i.e. parent – child). It is white list oriented. The spike detection, it is attribute oriented. This does not detect the fraud but updates the system regularly and attributes regularly other than the communal detection.

A. Demerits of existing system

The system detects the whether the data is fraud or original. If the system is data is fraud the processes do not proceed to the next level. The system is attribute oriented that the data is updated in the communal detection manually**.** The system does not verify from the blacklist database. Through the spike detection the system updates the attributes regularly. The system is not secure and it detects the original data also as fraud. (for eg.- twins applying the card is also detects as the fraudulent data).

III. PROPOSED METHOD

The communal detection focused on attack in the white list by defaulters when they submit applications with synthetic relationship. The volume and ranks of the white list's real communal relationships change over time, to make the white list exercise caution with (more adaptive) changing behavior legally, the white list is continually being reconstructed. The spike detection is attribute oriented. It cannot be detected by fraud attribute will be updated regularly. The attributes used in spike detection will not be in communal detection. By using the spike detection and communal

detection detects the defaulters in smart card application. In addition to communal detection and spike detection we use case based reasoning algorithm to make this method more efficient. CBR implements retrieval, diagnosis and pledge to make the data more secure. The CBR used to analyze and retrieval of data from the existing blacklist. The counterfeit datum is moved to the blacklist and the original datum is saved in the database.

A. Merits of Proposed Method

With the existing algorithms proposed the Case based reasoning algorithm to make the method secure. This method verifies the data with blacklist data. The blacklist data is verified with the CBR algorithm that is used to find the fraudulent data. This method updates the database without human intervention by using the data mining algorithm.

IV. FRAUD DETECTION ALGORITHMS

A. Communal Detection

If there are two credit card applications that provided the same postal address, phone number, cell number and date of birth(DOB), but in the first application the applicant's name to be John Smith, and in the other application the applicant's name to be Joan Smith. Either it is a defaulter attempting to obtain multiple credit cards using near duplicate data. Possibly there are twins living in the same house who both are applying for a credit card. Or it can be the same person applying two times, and there is a typographical error of one character in the first name. It is crucial because it reduces the scores of these legal behaviors and false positives. There are two problems with the white list. Initially there can be targeted attacks on the white list by defaulters when they submit applications with synthetic communal relationships Second, the volume and ranks of the white list's real communal relationships changes from time to time. To make the White list exercise caution with (or more adaptive to) changing legal behavior, the white list is continually being reconstructed.

B. Spike Detection

Spike Detection (SD) finds spikes to increase the mistrust value, and is query resistant for attributes. Probe resistance reduces the chances a defaulter will discover attributes used in the Spike Detection (SD) score calculation [4]. It is the attribute-oriented approach on a variable-size set of attributes. The redundant attributes are continually filtered; only selected attributes in the form of not-too-sparse and not-too-dense attributes are used for the Spike Detection (SD) suspicion score.

C. CBR Algorithm

1) Retrieval

Nearest neighbor matching is common to many CBR systems. Again using the basic exploratory facilities of CBR test bed, a set of cases which were considered to be very similar, i.e[13]., above a certain percentage of similarity, were recovered.

2) Diagnosis

Applying the general principle of threshold retrieval, a multi-algorithmic approach to final match analysis was developed as a result of the design and testing of a variety of single discrimination algorithms.[13]. It has been suggested that no single algorithm may perform equally well on all search and classification tasks, and that an algorithm's improved performance in one learning situation may come at the expense of accuracy in another.

3) Resolution

If a set of algorithms is asked to diagnose the set of cases retrieved for an unknown credit request, it is possible that the algorithms may disagree on the result, and resolution strategies were implemented to resolve the varying diagnoses into a single result.

V. SYSTEM ARCHITECTURE

The architecture diagram (Fig 1) represents the overall structure of the system. The data is detected for the crime detection using the data mining algorithm communal detection and spike

detection algorithm. These two algorithms combine together to remove the negative false and then proceeded to the proposed system algorithm (i.e) CBR algorithm. This algorithm retrieved and diagnosis the datum. If the data is fraud it is thrown into the black list database. If the data is original the data is stored in the database. The communal detection focused on attacks in the white list by fraudsters when they submit applications with synthetic relationship. The volume and ranks of the white list's real communal relationships change over time, to make the white list exercise caution with (more adaptive) changing legal behavior, the white list is continually being reconstructed. The spike detection is attribute oriented.

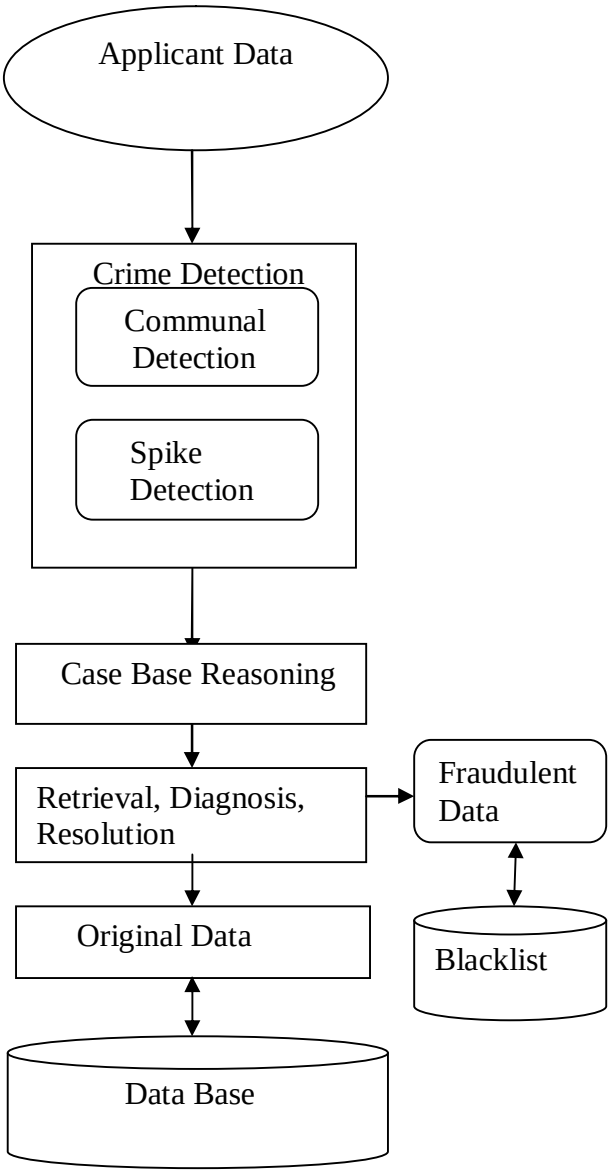

Fig. 1 System Architecture Diagram

It cannot be detected by fraud attribute will be updated regularly. The attributes used in spike detection will not be communal detection. By using the spike detection and communal detection detects the fraudsters in credit card application In addition to communal detection and spike detection we use case based reasoning algorithm to make this approach more efficient. CBR implements retrieval, diagnosis and resolution to make the data more secure. The CBR used to analyze and retrieval of data from the existing blacklist. The fraudulent datum is moved to the blacklist and the original datum is stored in the database.

A. Crime Detection

The crime detection consists of the two algorithms, communal detection and spike detection. The communal detection detects the fraudsters. This detection is the relationship oriented. This detection is attribute oriented. The spike detection detects the system fraudsters by updating the system attributes. These system finds the data whether the data is original or not. These two detections are mainly involved in existing crime detection.

B.Finding Legitimate User

The CBR is used is the fraud detection system that the data is original or not that the data is original or not by retrieving the data from the blacklist verification. This method finds the fraudulent data by the artificial intelligence. The CBR algorithm involves with the data mining concept with match analysis

c. Blacklist Verification

With the provided sets of details are taken into consideration to avoid the identity crime. The data is verified using the above algorithms to make the credit card application enormously efficient. If the data is original further processes will be enforced or otherwise the data will be found as fraud and it will be enrolled in the black list.

VI. CONCLUSION

The system detects fraud detection online credit card application. This system is used avoid the duplicates from the fraudsters while applying the credit card. Data mining algorithms are used this system. The existing algorithm communal detection and spike detection used to detect the multiple applicants. In proposed system combing with the existing algorithm spike detection and communal detection the CBR algorithm is used to make the system more efficient and secure. The CBR algorithm is used to throw the fraudulent data in the blacklist and retrieve the datum from the blacklist database. The identity thief has limited time because innocent people can discover the fraud early and do the necessary tasks, and will rapidly use the same real identities at different places.

VII. FUTURE ENHANCEMENTS

The detection of credit card application is used with the data mining layers. This system is used only on the application, in future the fraud detection in credit process (i.e) the card is used by the unauthorized user. This system can be developed with the data mining system and with the help of the biometric system.

REFERENCES

[1] Resilient Identity Crime Detection Clifton Phua, Member, IEEE, Kate Smith-Miles, Senior Member, IEEE, Vincent Cheng-Siong Lee, and Ross Gayler

[2] A. Bifet and R. Kirkby Massive Online Analysis, Technical Manual, Univ. of Waikato, 2009.

[3] R. Bolton and D. Hand, "Unsupervised Profiling Methods for Fraud Detection," Statistical Science, vol. 17, no. 3, pp. 235-255,2001.

[4] P. Brockett, R. Derrig, L. Golden, A. Levine, and M. Alpert, "Fraud Classification Using Principal Component Analysis of RIDITs,"The J. Risk and Insurance, vol. 69, no. 3, pp. 341-371, 2002, doi: 10.1111/1539-6975.00027.

[5] R. Caruana and A. Niculescu-Mizil, "Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria," Proc. 10th ACM SIGKDD Int'l Conf. Knowledge Discoveryand Data Mining (KDD '04), 2004, doi: 10.1145/1014052.1014063.

[6] P. Christen and K. Goiser, "Quality and Complexity Measures for Data Linkage and Deduplication," Quality Measures in DataMining, F. Guillet and H. Hamilton, eds., vol. 43, Springer, 2007, doi: 10.1007/978-3-540-44918-8.

[7] C. Cortes, D. Pregibon, and. Volinsky, "Computational Methods for Dynamic Graphs," Computational and Graphical Statistics,vol. 12, no. 4, pp. 950-970, 2003, doi: 10.1198/1061860032742.

[8] Experian. Experian Detect: Application Fraud Prevention System,Whitepaper, http://www.experian.com/products/pdf/experian_dete ct.pdf, 2008.

[9] T. Fawcett, "An Introduction to ROC Analysis," Pattern Recognition Letters, vol. 27, pp. 861-874, 2006, doi: 10.1016/j.patrec. 2005.10.010.

[10] A. Goldenberg, G. Shmueli, R. Caruana, and S. Fienberg, "Early Statistical Detection of Anthrax Outbreaks by Tracking Over-the-Counter Medication Sales," Proc. Nat'l Academy of Sciences USA (PNAS '02), vol. 99, no. 8, pp. 5237-5240, 2002.

[11] G. Gordon, D. Rebovich, K. Choo, and J.Gordon, "Identity Fraud Trends and Patterns: Building a Data-Based Foundation forProactive Enforcement," Center for Identity Management and Information Protection, Utica College, 2007.

[12] "Multiple Algorithms for Fraud Detection",Richard Wheeler and Stuart Aitken, Artificial Intelligence Applications Institute,The University of Edinburgh,80 South Bridge,Edinburgh EH1 1HN, Scotland, 2006.

[13] Schaffer. C, A Conservation Law for Generalized Performance,2009.

[14] O. Kursun, A. Koufakou, B. Chen, M. Georgiopoulos, K. Reynolds, and R.Eaglin, "A Dictionary-Based Approach to Fast and Accurate Name Matching in Large Law Enforcement Databases," Proc. IEEE Int'l Conf. Intelligence and Security Informatics (ISI '06), pp. 72-82, 2006, doi: 10.1007/11760146

AUTHORS PROFILE

M.Swathi obtained her B.Tech degree from Jawaharlal Nehru Technological University, Kakinada in the year 2009. She is pursuing her M.Tech in Computer Science and Engineering from Jawaharlal Nehru Technological University, Anantapur, India. during 2011 – 2013.