International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 6 Issue 2, February-2017, Impact Factor: 1.544

Study of Data Mining Techniques for Financial Cyber Crime and Frauds

Neeti Malik

Dept. of Computer Science

ABSTRACT

The Internet in India is growing rapidly. It has given rise to new opportunities in every field we can think of - be it entertainment, business, sports or education. There are different sides to a coin. Web additionally has its own detriments. One of the real impediments is Cyber fraud criminal behavior carried out on the web. The web, alongside its drawbacks, has likewise presented us to security chances that accompany associating with a huge system. Computers today are being abused for criminal operations like email secret activities, Master card fraud, fit, programming theft, etc, which attack our protection and insult our faculties. Crimes in the internet are on the ascent. Building up a budgetary digital fraud recognition framework is a difficult undertaking. At whatever point any online exchange is performed through the charge card, at that point there isn't any framework that without a doubt predicts an exchange as false. It just predicts the probability of the exchange to be a false. The creator read different methodologies for online exchange fraud identification, which joins confirmations from current just as past conduct.

Keywords: Data Mining, Techniques, Financial, Cyber Crime, Frauds.

INTRODUCTION

The Internet in India is growing rapidly. It has given rise to new opportunities in every field we can think of - be it entertainment, business, sports or education. There are two sides to a coin. Internet also has its own disadvantages. One of the significant hindrances is Cyber fraud criminal behavior carried out on the web. The web, alongside its hindrances, has likewise presented us to security hazards that accompany associating with a huge system. PCs today are being abused for criminal operations like email secret activities, charge card misrepresentation, fit, programming robbery, etc, which attack our protection and annoy our faculties. Crimes in the internet are on the ascent. So in the present electronic culture, internet business has turned into a basic deals channel for worldwide business. Because of quick progression of web based business, utilization of charge cards for buys has drastically expanded. Lamentably, fake or illicit utilization of charge card has additionally turned into an appealing wellspring of income for fraudsters. Events of Mastercard misrepresentation are expanding significantly because of introduction of security shortcomings in conventional charge card preparing frameworks bringing about loss of billions of cash each year. Fraudsters presently become exceptionally powerful and utilize advanced procedures to execute Visa fraud. The deceitful exercises overall present special difficulties to banks and other money related foundations who issue Mastercards.

A Gartner review of in excess of 160 organizations uncovers that multiple times more misrepresentation exists on Internet exchanges and those e-posteriors are paying Visa markdown rates that are 66 percent higher than conventional retailer charges. Besides, Web dealers bear the obligation and expenses in instances of fraud, while charge card organizations by and large assimilate the misrepresentation for conventional retailers.

Motivation

The Various digital fraud cases through the charge card comes as often as possible in the everyday news papers and expansive inclusion in the TV media motivated me to work here.

The motivation behind research is first to talk about the diverse monetary digital frauds and fakes which are seen today in the types of Credit card fraud, Phishing and so forth. Besides study the various Data mining procedures like Neural

Network, Clustering strategies, Decision trees and so forth and in the end how these systems can be utilized and applied to identify the budgetary digital fraud and cheats. Fraud Prevention portrays measures to stop misrepresentation happening in any case. Interestingly, fraud identification includes recognizing misrepresentation as fast as conceivable once it has been executed. Misrepresentation location becomes possibly the most important factor once fraud counteractive action has fizzled. By and by, fraud recognition must be utilized constantly, as one will ordinarily be unconscious that misrepresentation anticipation has fizzled. We can attempt to counteract charge card misrepresentation by guarding our cards perseveringly, yet on the off chance that by and by the card's subtleties are taken, at that point we should have the option to distinguish, at the earliest opportunity, that fraud is being executed.

As of now, Data mining is a well known approach to battle fakes as a result of its viability. The errand of Data mining is to examine a monstrous measure of Data and to remove some usable data that we can decipher for future employments. In doing as such, we need to characterize the reasonable objective of Data mining, and discover the correct structure of conceivable model or examples that fit to the given informational index. When we have the correct model for the Data, we can utilize the model for anticipating future occasions by ordering the Data. As far as Data mining, misrepresentation recognition can be comprehended as the order of the Data. Data is broke down with the proper model and decided if it suggests any false exercises or not. A well-characterized grouping model is created by perceiving the examples of previous deceitful practices. At that point the model can be utilized to anticipate any suspicious exercises inferred by new informational index.

A key issue of the proposed work is the manner by which successful the instruments are in recognizing fraud and a deceitful issue is that one commonly portion not know what number of fake cases sneak past the net. In applications, for example, normal time to identification after fraud begins (in minutes, number of exchanges, and so on.) ought to likewise be accounted for. Proportions of this perspective cooperate with proportions of conclusive recognition rate: as a rule a record, phone, and so forth should utilized for a few deceitful exchanges before it is recognized as fake, so a few false negative groupings will fundamentally be made.

LITERATURE SURVEY

Mastercard fraud discovery has drawn a ton of research intrigue and various methods, with exceptional accentuation on Data mining, have been proposed. Gosh and Reilly [1] have created fraud discovery framework with neural system. Their framework is prepared on enormous example of named Visa account exchanges. These exchanges contain model fraud cases because of lost cards, taken cards, application misrepresentation, fake misrepresentation, mail-request misrepresentation and non get issue(NRI) misrepresentation.

E. Aleskerov et al. [2] present CARDWATCH, a database digging framework utilized for Mastercard misrepresentation identification. The framework depends on a neural learning module and gives an interface to assortment of business databases.

Dorronsoro et al. [3] have proposed two specific attributes with respect to fraud discovery an exceptionally restricted time range for choices and countless Visa tasks to be prepared. They have isolated deceitful tasks from the typical ones by utilizing Fisher's discriminant investigation.

Syeda et al. [4] have utilized parallel granular neural system for improving the speed of Data mining and learning disclosure in Visa misrepresentation recognition. A total framework has been actualized for this reason.

Chan et al. [5] have partitioned an enormous arrangement of exchanges into littler subsets and after that apply conveyed Data digging for structure models of client conduct. The resultant base models are then joined to create a meta-classifier for improving discovery precision.

Chiu and Tsai [6] consider web administrations for Data trade among banks. An fraud example mining (FPM) calculation has been produced for mining misrepresentation affiliation rules which give data with respect to the new misrepresentation examples to avert assaults.

Some review papers have been distributed which arrange, look at and abridge articles in the territory of fraud recognition. Phua et al. [7] did a broad overview of Data mining based Fraud Detection Systems and exhibited an exhaustive report.

Kou et al. [8] have evaluated the different fraud location methods for charge card misrepresentation, media transmission fraud and PC interruption discovery. Bolton and Hand [9] depict the instruments accessible for measurable fraud discovery and regions in which misrepresentation recognition advances are most ordinarily utilized. D. W. Abbott et al. [10] think about five of the most exceptionally acclaimed business Data mining apparatuses on an fraud identification application, with portrayals of their unmistakable qualities and shortcomings, in view of the exercises learned by the creators during the way toward assessing the items.

D. Yue [11] lead a broad audit on literary works to find the solutions of the inquiries like (1) Can FSF be recognized? How likely and how to do it? (2) What Data highlights can be utilized to foresee FSF? (3) What sorts of calculation can be utilized to recognize FSF? (4) How to gauge the presentation of the identification? What's more, (5) How successful of these calculations as far as fraud location?

V. Hanagandi et al. [12] create a misrepresentation a score utilizing the recorded data on Visa account exchanges. They portray a cheat non fraud arrangement philosophy utilizing spiral premise capacity organize (RBFN) with a thickness based bunching approach. The info Data is changed into cardinal segment space and bunching just as RBFN demonstrating is finished utilizing a couple of cardinal segments.

FINANCIAL CYBER CRIME AND FRAUDS

Cyber crime encompasses any criminal act dealing with computers and networks (called hacking). Additionally, cyber crime also includes traditional crimes conducted through the Internet. For example, hate crimes, telemarketing and Internet fraud, identity theft, and credit card account thefts are considered to be cyber crimes when the illegal activities are committed through the use of a computer and the Internet. Information Systems Security Association (ISSA), Ireland conduct IRIS cyber crime survey every year. They developed a questionnaire in which respondents indicated the types of cyber crime incident which had affected their organization.

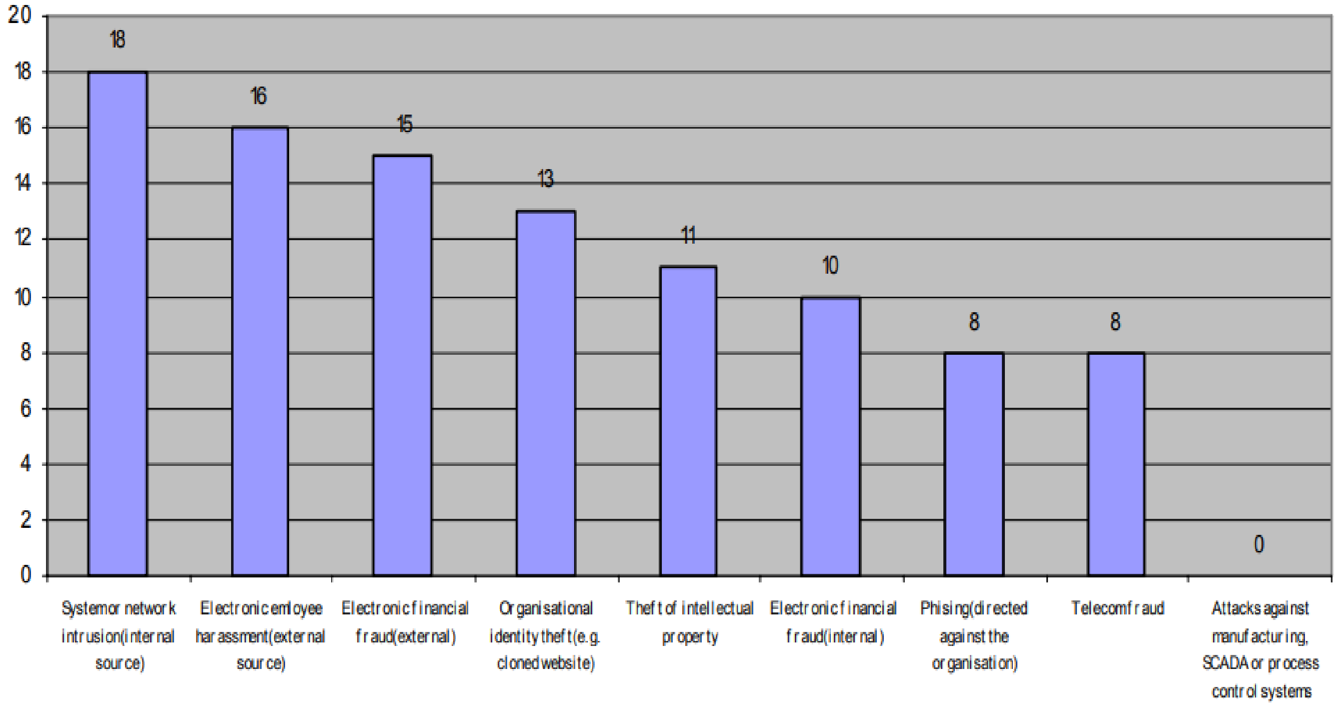

Figure 1: Affecting the Person by Cyber Crime (in %)

An Example of Financial Cyber Crime

One example of monetary fraud is, a site offered to sell Alphonso mangoes at a disposable cost. At first not many individuals reacted to or provided the site with their charge card numbers. These individuals were really sent the Alphonso mangoes. The word about this site presently spread quickly. A great many individuals from everywhere throughout the nation reacted to this site and requested mangoes by giving their Mastercard numbers. The proprietors of what was later demonstrated to be a counterfeit site at that point fled assuming the various praise card numbers and continued to spend tremendous measures of cash a lot to the embarrassment of the card proprietors.

Mastercard Fraud

We just need to type charge card no, expiry date, CVV no into www page of the seller for online exchange. On the off chance that electronic exchanges are not verified the charge card numbers can be taken by the programmers who can abuse this card by mimicking the Visa proprietor.

Fake Applications

This includes the unlawful procurement and utilization of someone else's recognizing data to acquire credit, or the utilization of that data to make an imaginary personality to set up a record. So as to submit wholesale fraud by methods for fake application, the culprit needs to gain not only a name, address or Mastercard number yet one of a kind identifiers, for example, mother's last name by birth, standardized savings number and definite data about an individual's record as a consumer, for example, the measure of their most home loan installment. This is the reason in excess of 40 percent of the data fraud cases that we see are submitted by somebody recognizable to the person in question, as often as possible a relative or somebody in a place of closeness or trust. This assortment of wholesale fraud speaks to three percent of our complete fraud cases.

Types of Telecommunications Fraud

There are a wide range of sorts of telecoms misrepresentation, and these can happen at different levels. The two most common sorts are membership misrepresentation and superimposed or 'surfing' fraud.

Subscription fraud: This happens when fraudster gets a membership to an administration, regularly with false personality subtleties, with no intension of paying. This is in this way at the degree of a telephone number – all exchanges from this number will be false.

Superimposed misrepresentation: This is the utilization of an administration without having the vital position and is typically identified by the presence of 'apparition' approaches a bill. There are a few different ways to complete superimposed fraud, including cell phone cloning and getting calling card approval subtleties. Superimposed fraud will by and large happen at the degree of individual calls – the deceitful calls will be blended in with the real ones. Membership fraud will by and large be identified eventually through the charging procedure – however one would intend to distinguish it a long time before that, since huge expenses can rapidly be kept running up. Superimposed fraud can stay undetected for quite a while.

ROLE OF DATA MINING IN FINANCIAL CRIME DETECTION

Today Industry is confronting gigantic misfortunes because of these kinds of money related violations, so it is ready to discover budgetary fraud through Data mining methods and evacuate it then it very well may be incredible advantage to the business. In this part we have proposed a two-level engineering model for money related fraud discovery. In the principal organize the monetary exchange is confirmed against the standard based framework and is given hazard score by the framework. These standards contain the human understanding and after that this exchange is passed to second phase of Data mining method, which will gain from the past experience of false exchanges and after that choose about the present exchange. So the exactness of forecast expanded as the budgetary exchange needs to go through two phases, one of principle based framework and second of Data mining method based framework.

Two Stage Solutions for Financial Crime Detection

Here we have given a figure 2 of design of 2-arrange answer for money related fraud. In the primary stage, rule based framework contains the static guidelines which is commonly founded on human Data for example human understanding. In the event that the money related exchange goes through this stage, at that point it goes to the subsequent stage.

In the subsequent stage, Data mining strategies create dynamic standards dependent on past false exchanges. Here learning is absolutely powerful so on the off chance that the example of false exchange changed; at that point the model takes in itself from exchanges and creates dynamic standards for forecast of budgetary fraud.

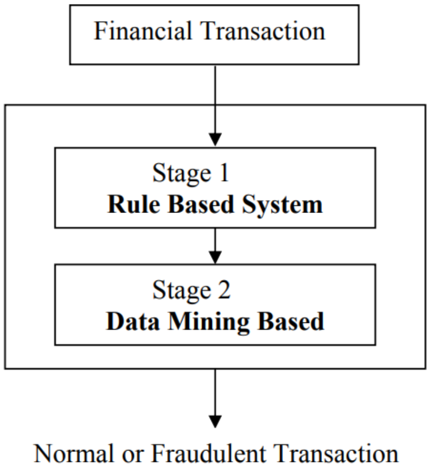

Figure 2: Two Stage Solutions for Financial Crime Detection

Rule Based System: A Rule Based scoring system can be developed for preventing loan default on various parameters like age (i.e. age is less then more points given or more age then less), educational qualification (for higher studies or degrees more points otherwise less), No of Assets owned by borrower at home (for more assets more points otherwise less), borrower's income, margin etc.

Detection Technique: There are many lead indicators available. There is regularly just one "pot" of cash that is burned through the different records - an example of money withdrawals from Mastercards, and after that toward the finish of the credit cycle, a comparable sum reimbursed, typically utilizing a money withdrawal from another charge card. Lead pointers incorporate charge cards that are once in a while used to make genuine shipper buys and have little extraordinary credit adjusts. Another example to search for is an advance record that is left unused. These methods blow up a midway controlled FICO score, giving a bogus impression that the record is considered capable. Discovery needs to happen before the "sting," which is a utilization of the credit and advance records quickly inside a credit cycle. This money related fraud can bring about high misfortunes. Identification must happen before the misfortune, in light of the fact that the sting has a short execution time.

Identification Technique: The strategy for recognition depends on out-of-design exchanges or atypical record use. Similarly as with other budgetary frauds, discovery must happen before any misfortune is supported. There are lead pointers like the "control of credit" portrayed above and in the absence of references, high relationship of coordinating characteristics, and questionable acknowledgment criteria.

The basic components for recognizing these monetary fraud frauds is knowing the conduct of credit, bank, and advance records and building up a comprehension of the classes of clients. Data mining can be utilized to spot anomalies or record utilizations that are ordinary and abnormal. Now and again the record appears "unrealistic," and it regularly is. The nonattendance of phone numbers or other contact data may demonstrate a "ring." These rings empower fake exercises to be separated from their sources and add multifaceted nature to criminal identification. Another sign is the numerous utilization of a similar location or telephone number for various records.

CONCLUSION

Data Mining Techniques like Neural Networks, Decision trees, Link Analysis and so forth can turn out to be useful for money related fraud recognition. These procedures can be utilized with guideline based framework combine so precision of forecast expanded without a doubt. The two tier engineering model is utilized adequately for any money related exchange confirmation. Any money related exchange needs to go through two degree of check, so forecast draws nearer to genuine expectation and furthermore any veritable or typical exchange isn't gotten by the model as fake exchange so ordinary or real client doesn't need to endure. Here, a two-organize answer for budgetary fraud location, which is really half and half approach and contains both human knowledge and machine understanding have been studied. In these kinds of fraud half breed approach demonstrates more dominant than any single stage arrangement and furthermore exactness of forecast is expanded definitely. In this sort of model or framework, we likewise need to deal with that any ordinary or real exchange must not be gotten by as false exchange and make overhead on client. In the event that any client endured, at that point we may lose him.

REFERENCES

[1] S. Haykin: Neural networks- a comprehensive foundation, MacMillan, New York

[2] Fawcet, T. and F. Provost- Adaptive Fraud detection, Data Mining and Knowledge Discovery.

[3] Jesus Mena : Investigative Data Mining for Security and Criminal Detection

[4] Dharwa J. N., Parikh S. M., “Data Mining in Financial Crime Detection” Proceedings of the National Conference on ECTKM-08, 23rd November 2008 at AITS, Rajkot.

[5] M.Minsky and S.Papert – Perceptrons-Expanded Edition: An Introduction to Computational Geometry. MIT Press 1987

[6] Dharwa Jyotindra N., Parikh S. M., Patel A. R. (Dr.) “A Comparative Study of Data Mining Techniques and its Selection Issue” 61-65 Proceedings of the national conference on IDBIT-2008, 23-24 February 2008 at SRIMCA, ISBN:978-81-906446-0-0.

[7] Ian H. Witten, Eibe Frank – DATA MINING Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers, ISBN: 0-12-088407-0

[8] J. Han, M. Kamber – Data Ming Concepts and Techniques, Morgan Kaufmann Publishers, ISBN: 81-8147-049-4.

[9] Margaret H. Dunham, S. Sridhar- DATA MINING Introductory and Advanced Topics, Pearson Education, ISBN 81-7758-785-4

[10] Richard J. Roiger, Michael W. Geatz – Data Mining A Tutorial-based Primer, Pearson Education, ISBN:81-297-1089-7.

[11] F.Yu, Z.Qin, X.Jia, “Data Mining Issues in Fraudulent Tax Declaration Detection”, in: Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xian, November 2003, pp.2202-2206

[12] A. Leung, Z.Yan, S.Fong, “On Designing a Flexible E-Payment System with Fraud Detection Capability”, in: Proceedings of the IEEE International Conference on ECommerce Technology, 2004

[13] W.Chai, B.K. Hoogs, B.T. Verschueren, “Fuzzy Ranking of Financial Statements for Fraud Detection”, in: Proceedings of the IEEE International Conference on Fuzzy Systems, Canada, 2006, pp.152-158.

[14] A. Shen, R.Tong, Y. Deng, “Application of classification models on credit card fraud detection”, in: Proceedings of the IEEE Service Systems and Service Management, International Conference, 9-11 June 2007, pp:1-4.

[15] Y.Kou, C.T.Lu, S. Sirwongwattana, Y. Huang, “Survey of fraud detection techniques”, in: Proceedings of the IEEE International Conference on Networking, Sensing and Control, vol. 1, 2004, pp.749-754.