RESEARCH ARTICLE   OPEN ACCESS

Fraud Detection using Data Mining Techniques in E-Commerce

Megha SR

(Assistant Professor, St. Joseph's Evening College (Autonomous) , Bangalore)

Abstract:

E- Commerce is a booming industry. And yet it faces the threat from fraudulent attacks. Over few years many researchers have come up in trying to prevent these fraudulent attacks. Data mining techniques have given a new perspective to look into this problem and have also helped the ecommerce industries to reduce the fraud rates using these techniques. In this course of paper we will encounter different fraudulent attacks that will help us in understanding various types of fraud and also we will explore our knowledge in understanding the techniques involved in trying to detect these fraudulent attacks efficiently using data mining techniques.

Keywords — **Fraud detection, Data Mining, Machine learning, Bayesian networks.**

I. INTRODUCTION

In the year 1990s *[1]* due to the drastic development of the web based networks there was an urge of doing business differently. Thus led to what we know today as web commerce or e-commerce. In short e-commerce is nothing but buying and selling products and services online. When we look back we see that there is always essential involvement of technology in how we do business. Over the centuries we can see that we have travelled a milestone from exchanging goods for barter to the very recent exchange of bitcoins and plastic money that has led to the change in how we look at business. Over the few millennia people travelled overseas to trade. But now it happens in finger tips. Internet will change in how we do business and you cannot ignore it! *[1]* *(Minoli and E. Minoli, Web commerce technology handbook. New York: McGraw-Hill, 1998.)*. Access to internet has exponentially increased over a period of time. As we eliminate the limitations of geographical boundaries the transportation and communication have become more efficient. Thus has resulted in the rise of millions of companies online. This revolution has challenged the traditional markets where the capital investments are huge. E commerce revolution has created an equal platform for small establishments

to reach millions of people over the net and also eradicates the misconception of competence of the with the big institutions that has huge financial background. So in today's world e-market is open to all who's creative and innovative ideas can put forth an impact globally with the use of technology. So businesses today see technology, not only an essential supporting tool but also a strategy in order to increase the horizon of business opportunities.

E commerce is a collaboration of various aspects whose origin can be traced from both science and commerce streams. This collaboration involves data management and various security protocols that enable safe and efficient transaction of data related to the sale of goods and services over the internet. This also ensures the authorized and authenticated data transfer in or among various organizations. Various security mechanisms play an important role in order to verify the source and to authorize the recipient who takes part in communication. So organizations have a new challenge how to transmit or communicate without losing data integrity. Due to the increase in the use of web commerce, the data generated is humongous. So the companies should come up with new strategies in order to process all these data. Over the course of this research paper we will be discussing about various security

---

mechanisms that will help preventing the fraudulent attacks.

II. DATA MINING

At companies today, data resides in repositories in different files, formats and in different environment. It is very complex to quickly extract meaningful data. Data is growing exponentially; as a result of this companies cannot depend in the traditional way of extracting data. 80% of data growth is unstructured data. The challenge is all to pull it together to make sense to the companies to find lucrative business opportunities. Data mining tools helps in organizing, storing, processing and mining meaningful data. In the field of e commerce these data mining can be used to predict upcoming trends, purchase patterns and also customer search histories. The companies can strategically plan out marketing and advertising techniques in order to attract customers to buy online.[8]

Data mining is an efficient way of finding the fraudulent patterns that might occur in an online transaction. This gives the e-tailers a new window of opportunities to gear up there business and marketing strategies.

III. TYPES OF FRAUDULENT ATTACKS

A) Identity theft: Fraudulent acquisition of victim's private details for their financial gain. This is one of the most increasing fraud types. The fraudsters might get hold of the private information such as name and address and try to get credit cards, loans etc., under the victim's name.

B) Account takeover: In this type the fraudster will try to access to the entire private details of the customer and will try to impersonate. so the existing accounts of the victim will be taken over by the fraudster in to have financial gain.

C) Phishing: This is one of the most common types of attacks. Its malicious act of acquiring individual's details by sending mass mails or creating websites through which the sensitive information such as credit card number, passwords etc., are being acquired.

D) Re- Route fraud: The fraudsters will give legitimate billing and shipping address at the time of placing an order online. But they change their shipping address when the product is out for delivery. The companies who give services online are prone to this type of an attack.

IV. EXISTING SYSTEM

FRAUD DETECTION

There is an increased focus to develop new tools and software in order to detect fraud in e commerce industries. The most common methodologies used in this field are Bayesian networks and machine learning techniques which enable us to find out possible fraudulent attacks using probability and statistics.[2] [3] [4].

A) Machine learning:

It's practically impossible for humans to process large amount of data. The industries stores all the data related to the customers in huge data repositories. These repositories contain all the information about the customer, there purchase patterns, private information etc. using these interesting patterns the companies try to extract interesting and alarming data that might help them in decision making in the field of fraud detection. They develop model using machine learning techniques and automates the process of trying to find fraudulent patterns and trying to alarm them. Machine learning uses complex algorithms to find patterns.

ORDER   MACHINE LEARNING MODEL   FRAUD RISK ESTIMATE   ACCEPT OR REJECT

[7] Fig 1 Fraud detection process

Methods in machine learning techniques to build models are:

1. Classification: this method is to find the target classes using complex algorithms and then giving the result based on the degree of risk factors.

For example: risk based on the geographical region based on the statistics received. The cities can be categorized into high risk, low risk and med risk. These kinds of relation and dependencies can be obtained from classification model.

2. Regression analysis: This is the method involving in finding out the dependencies between the variables. These uses complex algorithms to process the mathematical data. This is mainly used for low volumes of data which may require high mathematical logic. Based on the result we may encounter the degree of fraudulent behaviour [5].

B) Bayesian networks:

There are no defined or determined rules to declare a customer as a fraudster. We can actually try to find the probability of a customer being a fraudster by trying to evaluate and formulate using different relations and dependencies. Bayesian networks method does the job for us by using all these dependencies and applying standardized formulae to predict based on the obtained patterns. This requires expert system. With the help of user and the prediction analysis we can detect the possible fraudulent orders placed.[6]

PROPOSED SYSTEM

The fraud detection tools developed using the present data mining techniques are not efficient to give an accurate result. It also requires an expert to handle the intermediate process. The tools today that are used by various companies are not fully automated and often end up giving false fraudulent alarms. As the result of this the fraud investigator's job is to go extra mile to do further investigation by trying to contact customer based on the pattern analysis for additional details. Hence the customers find it inconvenient to disclose their private details to prove their identity. The existing system lacks in confirming the fraudulent orders placed but just gives degree of fraudulent behaviour. Hence we require new technologies that are efficient tools not only to give the degree of fraudulent behaviour but also to prove it as confirm fraud orders/customers.
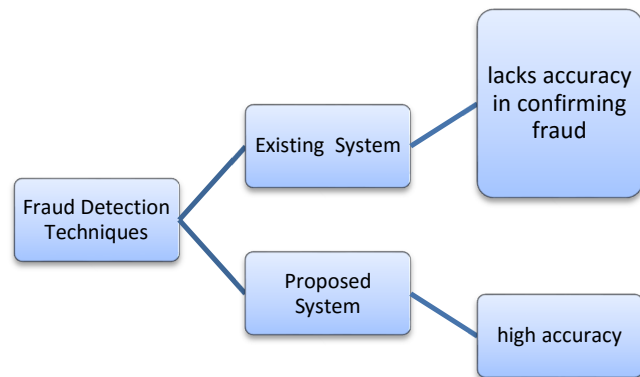
IV. CONCLUSIONS

This paper mainly focuses on the e commerce and the data mining techniques to prevent fraudulent attacks. We have also discussed the possible fraud attacks that are very much common in the field of e-commerce especially for credit card transaction online.

REFERENCES

1. [1]D. Minoli and E. Minoli, *Web commerce technology handbook. New York: McGraw-Hill, 1998.*

2. [2]M. Taniguchi, M. Haft, J. Hollmen and V. Tresp, *"Fraud detection in communication networks using neural and probabilistic methods", IEEE, 2002.*

3. [3]S. Swamy N and S. Lingareddy, *"Fraud Detection using Data Mining Techniques", IJIET.*

4. [4]C. PHUA, V. LEE, K. SMITH and R. GAYLER, *"A Comprehensive Survey of Data Mining-based Fraud Detection Research", 2010.*

5. [5]*"Fraud detection via regression analysis", ELSEVIER, vol. 9, no. 4, 1990.*

6. [6]*"A Survey on Fraud Detection Techniques in Ecommerce", International Journal of Computer Applications (0975 – 8887), vol. 133, no. 14, 2015.*

7. [7]*"How Machine Learning Facilitates Fraud Detection? - Maruti Techlabs", Maruti Techlabs, 2017. [Online]. Available: https://www.marutitech.com/machine-learning-fraud-detection/.*

8. *Data mining in e-commerce: A survey by N R SRINIVASA RAGHAVAN.*