ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 2, February 2017

A Survey on Fraud Detection Using Graph Database

Arshad Ahmad Malik, Prof. Varshapriya J.N.

M.Tech Student, Department of Computer Engg. (Specialization in Software Engineering), VJTI, Mumbai, India

Assistant Professor, Department of Computer Engg. & I.T., VJTI, Mumbai, India

ABSTRACT: In today's digital world there is dramatic increase in number of fraud resulting in huge amount of financial losses each year worldwide, several modern techniques are continuously developed for detecting fraud. Fraud detection involves monitoring the behavior of users of the system. This paper presents a survey of current techniques used in credit card fraud detection, insurance fraud detection. Sophisticated criminals often escape from by collaborating with each other and find a way to commit fraud, it is very difficult to detect fraud rings in existing techniques. Theproposed system uses graph database which can easily detect fraudrings.

KEYWORDS: Fraud, Sophisticated criminals, Fraud rings, CARDWATCH, Graph Database, eCommerce, mCommerce.

I. INTRODUCTION

The Association of Certified Fraud Examiners (ACFE) defined fraud as "the use of one's occupationforpersonalenrichmentthroughthedeliberatemisuseorapplicationoftheemploying organization's resources or assets [1]. Criminals do fraudulent transactions for financial or personal gain. Fraud occurs in many areas of daily life such as telecommunication networks, mobile communications, on-line banking and e-commerce. Consequentially, fraud detection has become an important issue to be explored. In today's digital world the volume of electronic transactions has raised significantly in last years, mainly due to the popularization of electronic commerce(e-commerce).Thereisalsoasignificantincreaseinthenumberoffraudcases,resulting in huge amount of financial losses each year worldwide. Fraudsters can and will exploit weaknesses wherever they can find them Therefore it is necessary to develop techniques todetect fraud and prevent any kind of damages.

Fraud detection system must identify fraud as quickly as possible once it has been committed. Companies are continuously developing new techniques to detect fraud. The development of new fraud detection methods is made more difficult due to the severe limitation of the exchange of ideas in fraud detection. Data sets are also not available and due to privacy concerns results are often not disclosed to the public. Fraud cases has to be detected from the available data sets which includes the behavior of user, transaction details. At present, fraud detection has been implemented by a number of methods such as data mining, statistics, and artificial intelligence [1]. Fraud is discovered by finding unusual user behavior like buying habit, accessing application from new device or new location etc. These unusual behavior helps to detect fraud and prevent financial losses but sophisticated criminals often find away to commit fraud by finding some loopholes in the existing system.

II. RELATED WORKS

In today's digital world the volume of electronic transactions has raised significantly in last years, mainly due to the popularization of electronic commerce (e-commerce), such as online retailers (e.g. Flip kart, Amazon, eBay).The e-commerce transactions mostly done using credit card, debit card, online banking etc. This increases the chance of fraud. We also observe a significant increase in the number of fraud cases, resulting in huge amount of financial losses each year worldwide. Therefore it is important and necessary to developed and apply techniques that can help in fraud detection and prevention. The design of effective technique to solve this problem is challenging due to several factors such as the heterogeneity and the non-stationary distribution of data. But, several techniques are implemented to solve this problem.

- **Fraud Detection:** As the number of fraud cases increasing day by day fraud prevention techniques must be developed to counter these frauds. Many researchers have developed systems which helps to detect fraud are as follows.

- **Outlier Detection:** An outlier is an observation that deviates so much from other observations which gives suspicion that it was generated by a different or by a fraudulent mechanism [10]. Outliers are a basic form of non-standard observation that can be used for fraud detection. In supervised learning methods, models are first trained using training dataset to find out difference between fraudulent and non-fraudulent transactions. The new observations are analyzed with the help of developed model so that an observations can be assigned to appropriate classes. Supervised methods require correct identification of fraudulent transaction in historical databases and can only be used to detect frauds of a type that has previously occurred. Unsupervised methods do not require the prior knowledge of fraudulent and non-fraudulent transactions, but this method detects the changes in behavior of current transactions. An advantage of using unsupervised methods over supervised methods is that previously undiscovered types of fraud may be detected. Supervised methods are only trained to differentiate between legitimate transactions and fraudulent transactions. Abnormal spending behavior and frequency of transactions will be identified as outliers, which are possible fraud cases [1]. Bolton and Hand proposed unsupervised credit card fraud detection, using behavioral outlier detection techniques [16].

- **Neural Networks:** A neural network is a set of interconnected nodes designed to imitate the functioning of the human brain [11]. Each node has a weighted connection to several other nodes in adjacent layers. Individual nodes take the input received from connected nodes and use the weights together with a simple function to compute output values. Neural networks come in many shapes and forms and can be constructed for supervised or unsupervised learning. The user specifies the number of hidden layers as well as the number of nodes within a specific hidden layer. Depending on the application, the output layer of the neural network may contain one or several nodes. CARDWATCH [2] is a database mining system used for credit card fraud detection. The system is based on a neural network learning module trained with historical data of particular customer, provides an interface to a variety of commercial databases and has a comfortable graphical user interface. It makes the network process the current spending patterns to detect possible anomalies. R. Brause and T. Langsdorf proposed the rule based association system combined with the neuro-adaptive approach [12].

- **Other Techniques:** Expert systems to encode expertise for detecting fraud in the form of rules. Pattern recognition to detect approximate classes, clusters, or patterns of suspicious behavior. Machine learning techniques to automatically identify characteristics of fraud. Hybrid systems and other data mining techniques is used to detect credit card fraud. No system guarantees hundred percent success rate of detecting fraud. CyberSource introduces a hybrid model, combining an expert system with a neural network to increase its statistic modeling and reduce the number of “false” rejections [13].

III. CREDIT CARDFRAUD

Credit card fraud is categorized into two types: offline fraud and online fraud. Offline fraud is committed by using a stolen physical card at storefront or any other place [1]. Online fraud is committed with the help of internet and mobile shopping [1].At the time of purchase only cards details are needed, and a manual signature is not required to complete the transaction. So, fraudsters can do fraudulent transactions easily without any verification. CARDWATCH [2] a database mining system used for credit card fraud detection.

In 2015 a research note from Barclays stated that the U.S. is responsible for around 47 percent of the world’s card fraud despite only accounting for 24 percent of total worldwide card volume [3]. Cross-border fraud occurs when fraudster’s uses a consumer's credit or debit card data in one country to make fraudulent transactions in another country. In 2014, 47 percent of fraudulent cross-border transactions on U.K. credit cards took place in the United

DOI: 10.15680/IJIRCCE.2017. 0502062

States [4]. Statistics shows that there is significant rise of credit card fraud cases in United States too. In 2014 about 31.8 million consumers of United States had their credit cards breached, it's more than three times the number affected in 2013 [5].

Most researchers believe that the reason behind the high number of fraud cases in United States is because it has been slow to adopt EMV. EVM is a global standard in which credit cards have computer chips that reduces counterfeiting by dynamically authenticating card transactions. Countries that have adopted for EMV have encountered a lesser number of fraud as a result [2]. The United States is implementing EMV, and once it is used at all places, card fraud should drop. But in some other countries, other types of fraud especially online fraud will probably grow. In the United States, online fraud is already a big problem.

The figure 1 shows the survey done on U.S. Card Fraud by Type in 2014. The survey shows that online credit card fraud accounted for 45 %, followed by counterfeit card fraud (37 %), fraud occurred due to lost/stolen cards (14 %) and 4% of frauds occurred by other means [6].

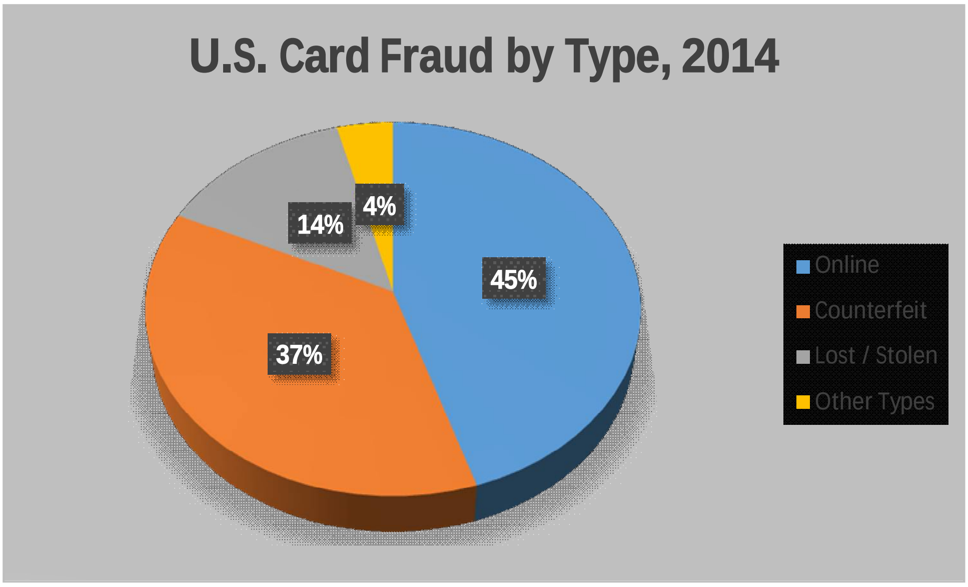

Figure 1. U.S. Card Fraud by Type,2014

In 2014 the consumers who were targeted by fraudsters around 46 % reported that the method by which they were targeted , 54 % said fraud attempts were initiated by phone, 23 % said by email, and 4 % said by mail [7]. Fraudsters most likely target there victims in night when fraud victims are sleeping so that any notification will not see by victim and fraudsters can easily do fraudulent transactions. Phone fraud is the biggest threat for enterprises across industries and borders, with large financial institutions call centers exposed to an average of more than $9 million in potential fraud each year. The fraud attempts that targeted United States call centers for retail and financial institutions increased by 30 % from 2013 to 2014 with 1 fraudulent call in every 2200 calls. For more financial gain the fraudsters mostly targeted the call centers of credit card issuers with 1 in every 900 calls being a fraud attempt [8].

Mobile commerce is now widely adopted as customers need the convenience to buy product with their mobile phones. Due to high adoption of m-commerce the mobile transactions are particularly at risk of fraud. In 2014 mobile transactions only accounted for 14 % of transaction volume but they made up of 21 % of all fraudulent transactions i.e. there is high number of fraudulent transactions as compared to other methods. This is a bad news for merchants who sell their products through mobile channels, as they lost 70 % more revenue due to fraud in 2014 than in 2013 [9].

The figure 2 shows proportion of fraudulent transactions attributed to payment methods in 2014. Most incidents of mobile fraud were related to credit cards. The credit cards were used in 61 % of e-commerce related fraud, 53 % of mobile commerce related fraud [9]. The debit cards were used in 18 % of e-commerce related fraud, 20 % of mobile commerce related fraud [9]. The fraudulent transactions done with the help of debit card, checks and other alternative payment methods are low as compared to credit card payment method. Due to increase in the number of fraud cases, which causes huge amount of financial losses each year worldwide, there is the need of fraud detection system to prevent these losses.

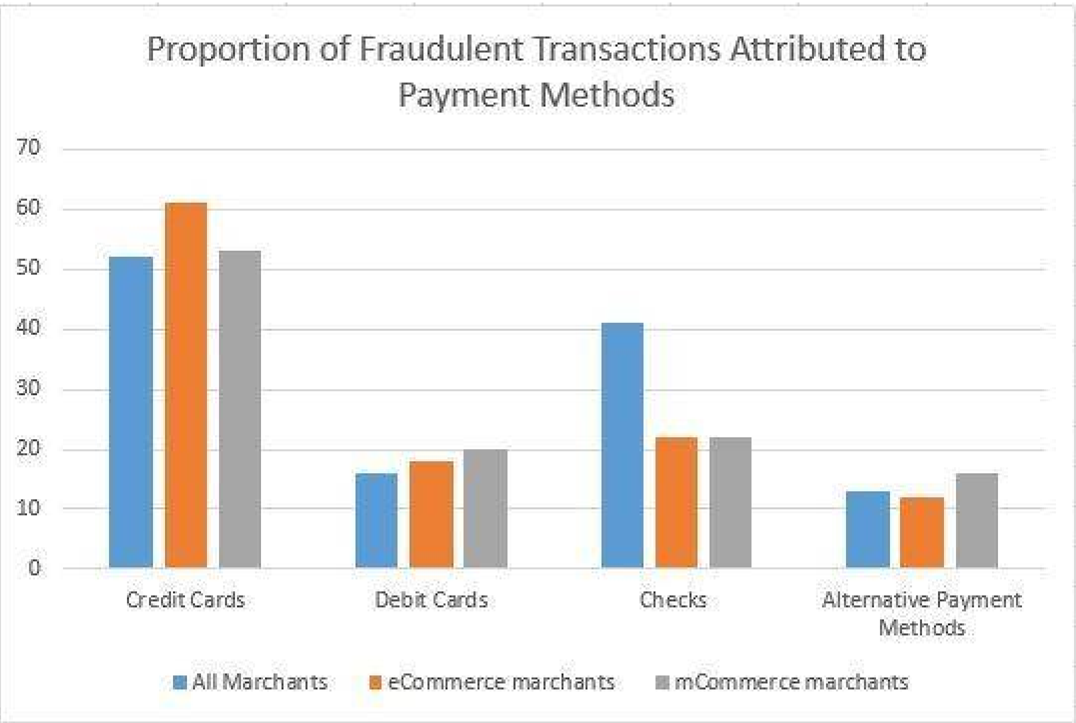

Figure 2. Proportion of Fraudulent Transactions Attributed to Payment Methods

IV. INSURANCE FRAUDDETECTION

Insurance fraud is a major crime that is costing insurance companies billions of dollars every year. According to statistics fraud steals $80 billion a year across all lines of insurance [14]. Existing analysis techniques like data mining techniques are sufficient to detect certain fraud scenarios but sophisticated/educated criminals often escape from these methods by collaborating with each other. Graph databases offer new methods of uncovering fraud rings and other sophisticated scams with a high-level of accuracy, and are capable of stopping fraud scenarios in real-time. Due to confidentiality issue insurance fraud detection techniques is not much disclosed in public. Data sets are also not available for experimental purpose.

While no fraud prevention measures can ever be able to prevent all the frauds, the improvement can be achieved by looking beyond the individual data or transactions i.e. try to find out the relationship among transactions. Most of the times these connections go unnoticed and fraudsters do the fraudulent transactions, as these connections most of the times hold the best clues. Understanding the connections between data and deriving relationship among data doesn't necessarily mean gathering new data. Some meaningful and new conclusions can be drawn from the existing data by simply reframing the problem and looking at it in a new way i.e. as a graph.

Graph database easily expresses the relationship between various data elements. Graph databases can uncover patterns that are difficult to detect using traditional methods which represents data as tables. Nowadays increasing number of companies are using graph databases to solve a variety of connected data problems which requires relationship among data including fraud detection.

Existing techniques like data mining algorithms are sufficient to detect certain fraud scenarios but sophisticated criminals often escape from these methods by collaborating with each other. The rings of sophisticated fraudsters

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 2, February 2017

work together to create fake accidents and raise claim against fake injury [15]. These fake accidents never happened in reality. They consists of fake drivers, fake passengers, fake pedestrians and even fake witnesses. Fraudsters often create and manage rings by recycling participants so as to create many fake accidents [15]. Thus one accident may have a particular person play the role of the driver. In another accident the same person may be a passenger or a pedestrian, and in another a witness.

The figure 3 depicts the same scenario where ring of fraudsters collaborate with each other and play different roles in different claims. The figure 3 shows that Person 1 acts as Driver for Car 1 and as a Witness in Car 3. Person 2 acts as driver for Car 2 and Passenger for Car 3. The sophisticated criminals collude with each other and create many fake accidents which doesn't even happen in real.

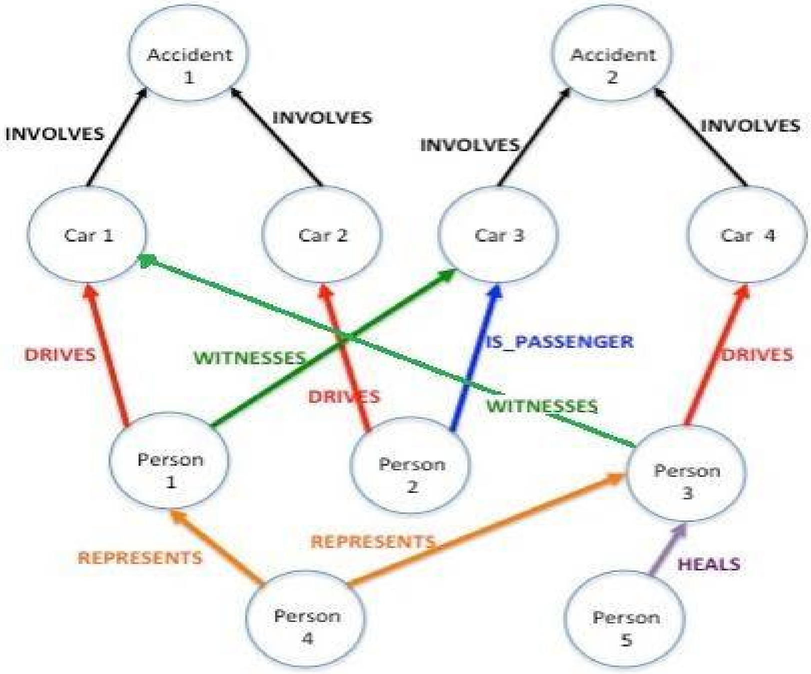

Figure 3. Fraud Ring

Clever usage of roles can generate a large number of costly fake accidents. These fraud rings are difficult to be discovered in traditional fraud detection system. In existing techniques data is stored in table. Uncovering these fraud rings require joining of many tables and for each transaction joining needs to be performed which is quite time consuming process [15].

In figure 4 the scenario depicted is a six-person collusion results in three false accidents. A group of persons collude with each other and create false accidents and each person plays the role of "driver" once and "passenger" twice. Assuming an average claim of $10K per injured person, and $5K per car. If on an average 4 persons are injured in one accident, so the average claim per accident will be approximately $50K. So the ring can claim approximately $150K in total.

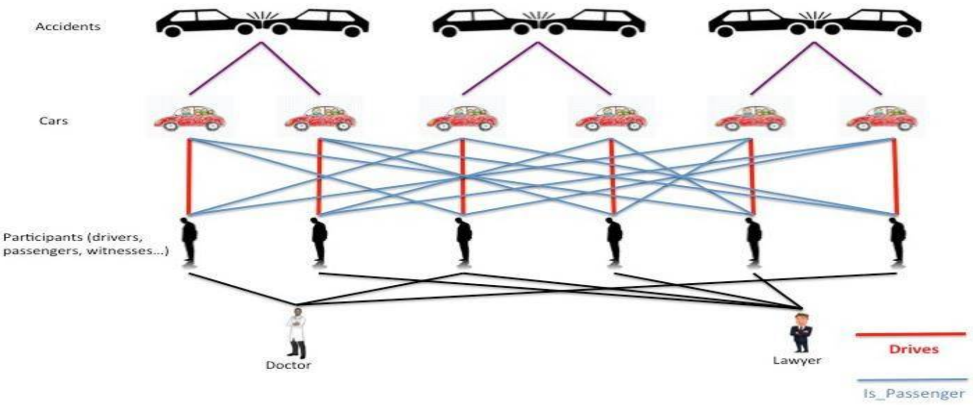

Figure 4. 6-person collusion

In proposed system graph database will be used to detect insurance fraud. Graph database can be used to uncover the fraud ring in real time. This technique not only analyses individual data points but also does the connected analysis to discover the relationship among data. Connected analysis helps to find relationships between people who are otherwise acting like perfect strangers. We will also use social media data to uncover these fraud rings. Graph database provide faster and efficient solution and prevents these frauds in real time. The social media analysis, weather analysis will also help to uncover certain types of frauds. For example, a person submitted an insurance claim to insurance company that his/her car is damaged due to flood. However, the details from the internet source (social media/weather forecast) verify that there was no flood at specified place on that day

V. CONCLUSION

Insurance fraud detection using Graph database helps insurance company to uncover fraud rings in real time which was very difficult and time consuming process using existing techniques like data mining algorithm. This technique helps the insurance company to view the fraudulent transactions in graphical manner which is easy to understand and extract fraud patterns. Traditional techniques, while still suitable for certain types of fraud prevention but it is very difficult to detect fraud rings in real time. As, criminals working in collaboration with each other often escape from these techniques and often go undetected. Graph database provide faster and efficient solution and prevents these frauds in real time.

Social media data like Facebook, Twitter, LinkedIn can be used to find out the relationship among the people involved in accident. Social media also uses graph database to find out the relationship between any two people. The system will use weather API to extract information about the environmental condition of mentioned place to uncover some types of fraud. We would also like to explore the use of machine learning techniques to enable more complex searches, using natural language processing to extract some meaning information, and to allow the system assessing the existence of fraud.

REFERENCES

1. Yufeng Kou, Chang-Tien Lu, Sirirat Sirwongwattana and Yo-Ping Huang. Survey of Fraud Detection Techniques. In Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control Taipei, Taiwan, March 21-23, 2004.
2. E. Aleskerov, B. Freisleben, and B. Rao. Cardwatch: a neural network based database mining system for credit card fraud detection. In *Proceedings of Computational Intelligence for Financial Engineering*, pages 173-200, 1997.
3. Barclays' Security in Payments: A Look into Fraud, Fraud Prevention, & the Future, May 22, 2015.
4. FICO press release, June 25, 2015 at http://www.fico.com/en/newsroom/fico-sees-25-percent-jump-in- cross-border-fraud-on-uk-debit-cards-in-2014-06-25-2015
5. Javelin Strategy & Research 2015 Data Breach Fraud Impact Report at https://www.javelinstrategy.com/coverage-area/2015-data-breach-fraud-impact-report

DOI: 10.15680/IJIRCCE.2017. 0502062

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 2, February 2017

6. Aite Group's EMV: Lessons Learned and the U.S. Outlook, June 10, 2014 at http://aitegroup.com/report/emv-lessons-learned-and-us-outlook.
7. Consumer Sentinel Data Book for January - December 2014 at https://www.ftc.gov/reports/consumer- sentinel-network-data-book-january-december-2014.
8. Pindrop Security Phone Fraud Report at http://www.contactcentrenews.co.uk/2015/06/17/pindrop- security-reveals-financial-and-retail-institution-call-centers-see-30-percent-rise-in-phone-fraud/.
9. LexisNexis True Cost of Fraud mCommerce at http://www.lexisnexis.com/risk/newsevents/press- release.aspx?Id=1422223947870687 on 2014.
10. E. Hung and D.W. Cheung. *Parallel Algorithm for Mining Outliers in Large Database.*http://citeseer.nj.nec.com/hung99parallel.html, 1999.
11. S. Ghosh and D.L. Reilly. Credit card fraud detection with a neural-network. In J. F. Nunamaker and R. H. Sprague, editors, *Proceedings of the 27th Annual Hawaii International Conference on System Science. Volume 3: Information Systems:DSS/Knowledge-Based Svstems*, pages 621-630, Los Alamitos, CA, USA, Jan. 1994. IEEE Computer Society Press.
12. R. Brause, T. Langsdorf, and M. Hepp. Credit card fraud detection by adaptive neural data mining. In *Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence,* pages 103- 106, 1999.
13. CyberSource Company. *Credit card fraud management.* http://www.cybersource.com,1 996
14. Insurance Fraud Organization at http://www.insurancefraud.org/statistics.htm
15. Insurance Fraud Detection at https://neo4j.com/use-cases/fraud-detection/
16. R. J. Bolton and D. J. Hand. Unsupervised profiling methods for fraud detection. In *conference of Credit Scoring and Credit Control VII, Edinburgh. UK, Sept* 5-7, 2001.

DOI: 10.15680/IJIRCCE.2017. 0502062