IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A SURVEY OF FRAUD DETECTION TECHNIQUES IN SOCIAL MEDIA

Leena Shibu*, Dr.Ajeet Chikkamannur
* Assistant Professor, Department of CSE , New Horizon College of Engineering, Kadubisnahalli, Bangalore.
Professor and Head of Department, Department of ISE and Engineering, New Horizon College of Engineering, Kadubisnahalli, Bangalore.

ABSTRACT

Due to the increase of fraud which results in loss of billions of dollars worldwide each year, several modern techniques in detecting fraud are continually evolved and applied to many business fields in social media. Fraud detection involves observing the performance of populations of users in order to estimate, detect, or avoid undesirable behavior. Undesirable behavior is a broad term including delinquency fraud, intrusion, and account defaulting. This paper presents a survey of current techniques used in social media fraud detection. The goal of this paper is to provide a comprehensive review of different techniques to detect frauds in social media.

KEYWORDS: Fraud detection, computer intrusion, data mining, knowledge discovery, neural network.

INTRODUCTION

The Association of Certified Fraud Examiners (ACFE) defined fraud as “the use of one’s occupation for personal enrichment through the deliberate misuse or application of the employing organization’s resources or assets .” In the technological systems, fraudulent activities have occurred in many areas of daily life such as telecommunication networks, mobile communications, on-line banking, and Ecommerce. Fraud is increasing intensely with the growth of modern technology and global communication, resulting in considerable losses to the businesses. Consequentially, fraud detection has become an important issue to be learned. Fraud detection involves identifying fraud as quickly as possible once it has been perpebated. Fraud detection methods are continuously developed to protect criminals in adapting to their strategies. The development of new fraud detection methods is made more difficult due to the severe limitation of the exchange of ideas in fraud detection. Datasets are not made available and results are often not disclosed to the public. The fraud cases have to be detected from the available huge data sets such as the logged data and user behavior. At present, fraud detection has been implemented by a number of methods such as data mining, statistics, and artificial intelligence. Fraud is discovered from anomalies in data and patterns.

FRAUD DETECTION TECHNIQUES IN SOCIAL MEDIA

Fraud Detection in Social Media Using Neural Network

A neural network is a set of interconnected nodes designed to replicate the functioning of the human brain [15]. Each node has a weighted connection to several other nodes in adjacent layers. Individual nodes take the input received from associated nodes and use the weights together with function to calculate output values. Neural networks come in many 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. Machine learning, adaptive Pattern Recognition, neural networks, and statistical modeling are employed to develop Falcon predictive models to provide a measure of certainty about whether a particular transaction is fraudulent. ANN (Artificial Neural Networks) provides the ability to generalize from previously observed behavior (normal or malicious) to recognize similar future unseen behavior for both anomaly detection and misuse detection [5]. It is implemented by a back propagation neural network. Problem with neural networks is that a number of parameter have to be set before any training can begin. However, there are no clear rules how to set these parameters. Yet these parameters determine the success of the training. In the most general case, neural networks consist of an (often very high) number of neurons, each of which has a number of inputs which are mapped via a relatively simple function to its output. Networks differ in the way their neurons are interconnected (topology), in the way the output of a neuron determined out of its inputs (propagation function) and in their temporal behavior (synchronous, asynchronous or continuous)[4].

Fraud Detection in Social Media Using Expert Systems

Classification model is built with association rules algorithm This approach can automatically generate concise and accurate detection models from large amount of audit data. However, it requires a large amount of audit data in order to compute the profile rule sets. Moreover, this learning process is an integral and continuous pattern of an intrusion detection system because the rule sets used by the detection module may not static over a long period of time.

Fraud Detection in State Transition Analysis

In State Transition analysis technique the monitored system is represented as a state transition diagram. As data is analyzed, the system makes transitions from one state to another. A transition takes place on some Boolean condition being true (for example, the user opening a file).

However there are also a few problems with state transition systems. First, attack patterns can specify only a sequence of events, rather than more complex forms. Second, there are no general purpose methods to prune the search except through the assertion primitives described above. And finally, they cannot detect denial of service attacks, failed logins, variations from normal usage, and passive listening -- this is because these items are either not recorded by the audit trail mechanism, or they cannot be represented by state transition diagrams[2].

State Transition Analysis is a misuse detection technique, which attacks are represented as a sequence of state transitions of the monitored system. Actions that contribute to intrusion scenarios are defined as transitions between states. Intrusion scenarios are defined in the form of state transition diagrams. Nodes represent system states and arcs represent relevant actions. If a compromised (final) state is ever reached, an intrusion is said to have occurred[3].

Fraud Detection Using Social Network Analysis.

Fraud detection and analysis has traditionally involved a silo approach. Rarely does an investigator look across product lines to identify fraudulent connections. However, with the introduction of social network analysis (SNA), investigators are now able to detect data patterns within and across product lines as a potential crime ring or group is developing, saving companies from losses as the crime ring further develops. At a basic level, a social network consists of nodes (vertices) that are connected to other related nodes by links (relationships). The connection between two nodes is called an edge. If all the nodes in a social network are connected to each other, it is called a fully-connected network. A path refers to a collection of nodes that are connected by a link. The variations can be detected in a financial social network like Denial of Service–Hacker attack in Star network, Networking Fraud Ring (Circle), Money Laundering (Chain). Density measures are extremely useful in determining potential fraud hotspots in retail banking from a maze of account transactions and applied control measures. Credit card transaction monitoring and money laundering are potentially two areas where density metrics could trigger the necessity for deeper investigations.
Applying SNA does come with some limitations regarding data and data processing, proactively fighting fraud and regulatory barriers. Data remodeling is required so that the effectiveness of SNA does not deteriorate as the volume of data observed increases. Database query languages like SQL though quite efficient cause significant overheads due to the joint operation performed on extremely large datasets that could increase investigation time. Transactional systems' data needs to be remodeled to avoid disambiguation ,and improve data consolidation and aggregation for enhanced data availability[4].

Outlier Detection:

An outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism .Unsupervised learning approach is employed to this model.

Usually, the result of unsupervised learning is a new explanation or representation of the observation data, which will then lead to improved future responses or decisions. Unsupervised methods do not need the prior knowledge of fraudulent and non-fraudulent transactions in historical database, but instead detect changes in behavior or unusual transactions.

These methods model a baseline distribution that represents normal behavior and then detect observations that show greatest departure from this norm. Outliers are a basic form of non-standard observation that can be used for fraud detection. In supervised methods, models are trained to discriminate between fraudulent and non-fraudulent behavior so that new observations can be assigned to classes. Supervised methods require accurate identification of fraudulent Outlier Detection[3].

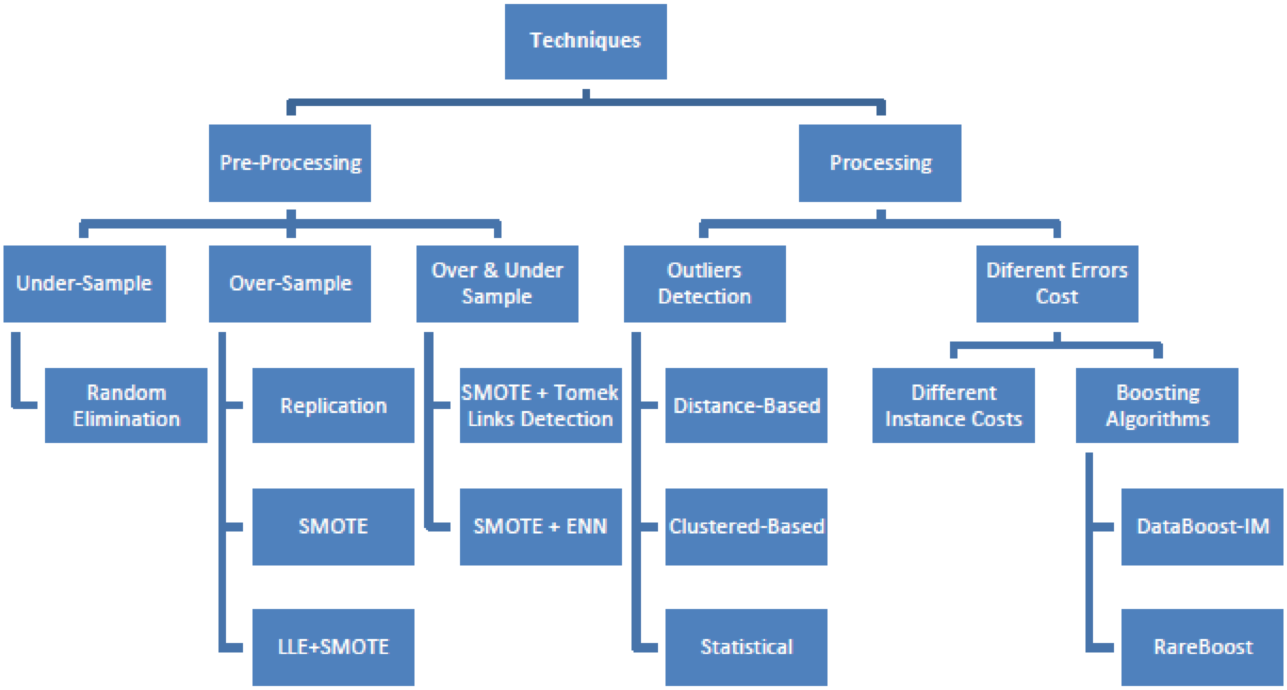

Figure 4 - Data mining techniques to detect fraud

Data mining strategies fall into two broad categories: supervised learning and unsupervised learning. Supervised learning methods are deployed when there exists a target variable with known values and about which predictions will be made by using the values of other variables as input. Unsupervised learning methods tend to be deployed on data for which there does not exist a target variable with known values, but for which input variables do exist. Although unsupervised learning methods are used most often in cases where a target variable does not exist, the existence of a target variable does not preclude the use of unsupervised learning[6].

Other Techniques A genetic algorithm [8] is applied to detect malicious intrusions and separate them from normal use. A genetic algorithm is a method of artificial intelligence problem solving based on the theory of Darwinian evolution applied to mathematical models. This genetic algorithm was designed so that each individual represented a possible behavioral model. This approach provides a high detection rate and a low false alarm rate. Dokas and Ertoz proposed building rare class predictive models for identifying known intrusions [3]. This method can address the inability of standard data mining techniques when dealing with skewed class distribution.

CONCLUSION

Fraud Detection an emerging field of research. However, it is beginning to assume enormous importance in today's computing environment like the combination of facts such as the uninhibited growth of the Internet, the vast financial possibilities opening up in electronic trade, and the lack of truly secure systems make it an important and pertinent field of research. Future research trends seem to be converging towards a model that is a hybrid of the anomaly and misuse detection models; it is slowly acknowledged that neither of the models can detect all fraud attempts on their own.

REFERENCES

[1] 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: DSU Knowledge-Based Svstems, pages 621430, Los Alami, tos, CA, USA, Jan. 1994. IEEE Computer Society.

[2] An Introduction to Intrusion Detection by Aurobindo Sundaram,1996

[3] Survey of Fraud Detection Techniques Yufeng Kou, Chang-Tien Lu, Sirirat Sinvongwattana Yo-Ping Huang Dept. of Computer Science Dept. of Computer Science Virginia Polytechnic Institute and Engineering and State University Tatung University Falls Church, VA 22043, USA Taipei, Taiwan 10451 {ykou, ctlu,ssiriwon}@vt.edu yphuang@cse.ttu.edu.tw,2004.

[4] Implementing social network analysis for fraud 2011 CGI Group Inc.

[5] Credit Card Fraud Detection Using Neural Network Raghavendra Patidar, Lokesh Sharma International Journal of Soft Computing and Engineering (IJSCE)ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011

[6] Using Data Mining Techniques for Fraud Detection Solving Business Problems Using SAS® Enterprise Miner™ Software.

[7] Classification for Fraud Detection with Social Network Analysis, 2009

[8] Detecting credit card fraud by genetic algorithm and scatter search Ekrem Dumana, M. Hamdi Ozcelikb, Elsevier,2011

[9] Syed Ahsan Shabbir, Kannadasan R, “An Effective Fraud Detection System Using Mining Technique”, International Journal of Scientific and Research Publication, Volume 3, Issue 5, May 2013, ISSN 2250-3153.

[10] Anjaneyulu C, Madhusekhar Y, “An Effective Approach towards the Credit Card Fraudulence Disclosure Methods”, International Journal of Computer and Electronics Research, Volume 2, Issue 4, August 2013, page 432-434.

[11] Pratiksha L. Meshram, Tarun Yenganti, “Credit and ATM Card Fraud Prevention Using Multiple Cryptographic Algorithm”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013, Page 1300-1305.