ISSN 2229-5518

Detecting Telecommunication Fraud using Neural Networks through Data Mining

Mohammad Iquebal Akhter, Dr. Mohammad Gulam Ahamad

Abstract--- Neural computing refers to a pattern recognition methodology for machine learning. The resulting model from neural computing is often called an artificial neural network (ANN) or a neural network. Neural networks have been used in many business applications for pattern recognition, forecasting, prediction and classification. Neural network computing is a key component for any data mining tool kit. Applications of neural network based data mining tools are abound in finance, marketing, manufacturing, information systems and so on. This essay discusses in detail the role of artificial neural networks in prevention of telecommunication fraud.

Keywords--- Artificial iintelligence, Artificial neural networks (ANNs), Coral systems, Data Mining, Expert System, Falcon (Or Credit Card Issuers), Fraud Management Systems, Ghosting, Knowledge Base System, Rule Based System, Telecommunication fraud, Thresholding Systems.

1 An overview of telecommunication fraud

THE telecommunication industry has expanded dramatically in the last few years with the development of affordable mobile phone technology (Pieprzyk J, Ghodosi H and Dawson E, 2007, pp 446-447). With the increasing number of mobile phone subscribers, global mobile phone fraud is also set to rise. It is a worldwide problem with substantial annual revenue losses of many companies. Telecommunication fraud which is the focus is appealing particularly to fraudsters as calling from the mobile terminal is not bound to a physical location and it is easy to get a subscription. This provides a means for illegal high profit business for fraudsters requiring minimal investment and relatively low risk of getting caught. Telecommunication fraud is defined as the unauthorized use, tampering or manipulation of a mobile phone or service.

At the beginning of the twenty first century, the convergence of computing and communication technologies has altered considerably the way in which industrialized communities function. It has created unfold benefits for education, delivery of health services, recreation and commerce and changed considerably the nature of modern workplaces and patterns of employment. Telecommunication fraud can be simply described as any activity by which telecommunications service is obtained without intention of paying. This kind of fraud has certain characteristics that make it particularly attractive to fraudsters. The main one is that the danger of localization is small. This is because all actions are performed from a distance which in conjunction with the mess topology and the size of network makes the process of localization time consuming and expensive. Additionally no particularly sophisticated equipment is needed if one is needed at all. The simple knowledge of an access code, which can be acquired even with methods of social engineering, makes the implementation of fraud feasible. Finally the product of telecommunication fraud, a phone call is directly convertible to money.

- *Mohammad Iquebal Akhter, Lecturer, Department of Computer Science & Engineering, University College, Umm Al-Qura University, Makkah, Saudi Arabia, E-mail: akhter72@gmail.com*
- *Dr. Mohammad Gulam Ahmad Professor Department of Computer Engineering, Salman Bin Abdulaziz University, Alkarj, Saudi Arabia, E-mail: gulamahamad@hotmail.com*

2 Telecommunication Fraud Detection

Fraud is a multi billions problem around the globe. The problem with telecommunication fraud is the huge loss of revenue and it can affect the credibility and performance of telecommunication companies. The most difficult problem that faces the industry is the fact that fraud is dynamic. This means that whenever fraudster's feel that they will be detected they find other ways to circumvent security measures. Telecommunication fraud also involves the theft of services and deliberate abuse of voice and data networks. In such cases the perpetrator's intention is to completely avoid or at least reduce the charges for using the services. Over the years, fraud has increased to the extent that losses to telephone companies are measured in terms of billions of American dollars. Fraud negatively impacts on the telephone company in 4 ways such as financially, marketing, customer relations and shareholder perceptions.

There are various techniques available for managing and detecting telephone fraud these include: 1) Manual review of data, the problem with this technique is the fact that there are too many data records for a team to filter the fraudulent data. Typically a telecom company will have in order of 1 million or more records of telephone calls generated by their customers for a single month within a specific region. As a result this is a time consuming and laborious technique for detecting fraud. 2) Conventional analysis using a fixed rule based expert system together with statistical analysis. A rule based system is a set of rules that take into account the normal calling hours, the called destinations as well as the normal duration of the call and etc. 3) Adaptive flexible techniques using advanced data analysis like artificial neural networks (ANNs). Fed with

raw data, a neural network can quickly learn to pick up patterns of unusual variations that may suggest instances of fraud on a particular account (Liatsis P, 2002, pp 474-475).

3 Types of Telecommunication Fraud

The telecom industry suffers major losses due to fraud (Prasad S K, Routray S and Khurana R, 2009, pp 259-260; Żytkow J M and Rauch J, 1999, pp 251). There are many different types of telecommunications fraud and these can occur at various levels. The two most types of fraud are subscription fraud and superimposed fraud.

3.1 Subscription Fraud

In subscription fraud, fraudsters obtain an account without intention to pay the bill. This is thus at the level of a phone number, all transactions from this number will be fraudulent. In such cases abnormal usage occurs throughout the active period of the account. The account is usually used for call selling or intensive self usage.

3.2 Superimposed Fraud

In Superimposed fraud, fraudsters take over a legitimate account. In such cases the abnormal usage is superimposed upon the normal usage of the legitimate customers. There are several ways to carry out superimposed fraud, including mobile phone cloning and obtaining calling card authorization details. Examples of such cases include cellular cloning, calling card theft and cellular handset theft. Superimposed fraud will generally occur at the level of individual calls; the fraudulent calls will be mixed in with the justified ones. Other types of telecommunications fraud include ghosting (technology that tricks the network in order to obtain free calls) and insider fraud where telecommunication company employees sell information to criminals that can be explained for fraudulent gain.

4 Methods of Telecommunication Fraud

There are two methods by which telecommunication fraud can be implemented. They are discussed below in detail.

4.1 Theft of communication Services:

The phone phreakers of 3 decades ago set a precedent for what is becomes a major criminal industry (Wall D S, 2001, pp 30; Broadhurst R G and Grabosky P N, 2005, pp 32). The market for stolen communications services is now large. There are those who simply seek to avoid or to obtain a discount on the cost of a telephone call whilst there are others such as illegal immigrants who are unable to acquire legitimate information services without disclosing their identity and their status.

4.2 Communications in Furtherance of criminal conspiracies:

Just as legitimate organizations in the private and public sectors rely upon information systems for communications and record keeping so too are the activities of criminal organizations enhanced by technology.

5 Approaches to face fraud

Dealing with the fraud is a very complex task mainly due to its transversal nature to the operator's structure (Samarati P, 2010, pp 201). Traditional fraud techniques are evolving and adapting to the new network infrastructure. The fraud techniques are considered because basic ideas remain despite the underlying technology.

Deceptions in telecommunications include subscription frauds where the cheater accesses the services without being subscribed. User can also suffer line or identity theft being charged for services used by others. Telecommunication operators can oversee users that exceed their download quote and rate performing illegal service redistribution, sometimes for an economic profit. Finally cloning or unauthorized access to services may lead to compromising privacy.

Anyway the most common types of fraud on telecommunications are subscription fraud and identity theft. After that voice mail fraud and calling card fraud prevail. The analysis of different fraud techniques points out that the tendency is a convergence of the fraud which increases the complexity of its detection. Fraud management systems have proved to be a suitable tool to detect fraud in different networks with diverse techniques such as self organizing maps (SOM), general data mining, rule based systems profiling through Artificial intelligence techniques like neural networks or decision trees based on the hierarchical regime switching models, Bayesian networks, fuzzy rules or other data mining techniques. There also exist works on how to discover new rules to detect fraud in telecommunications and on the privacy concerns of applying detection techniques to user's data.

Fraud detection can also be done at 2 levels call or behavior and with two different approaches user profile or signature based (Perner P, 2006, pp 535). Most of the techniques use the CDR data to create a user profile and to detect anomalies based on these profiles. The mined large amounts of CDR have in order to find patterns and scenarios of normal usage and typical fraud situations. These scenarios were then used to configure monitors that observe the user behavior with relation to that type of fraud. These monitors are then combined in a neural network which raises an alarm when sufficient support of fraud exists. This type of system can be classified in a rule based approach since it relies in the triggering of certain rules due to abnormal usage. The rule based system has the drawback of requiring expensive management of rules. Rules need to be precise (avoid false positive alarms) and constantly evolving (detect new scenarios) which result in very time consuming programming.

The most common and best succeeded methods for fraud analysis are signature based. These methods detect the fraud based on deviation detection by comparing the recent activity with the user behavior data which is expressed through the user signature. The work can be adapted and extended by reformulating the notion of signature and by introducing the notion of statistical based distances to detect anomalies. Furthermore the

computational cost can be reduced by using simple statistical functions avoiding processing costly histograms. A clear problem with a histogram approach is that discretization intervals or bucket must be clean and what is right for one customer may be wrong for another.

Other approaches have also been widely applied to fraud analysis like neural networks. Another applied technique is link analysis. Here the client links are updates over time establishing a graph of called communities of interest that can easily reveal networks of fraudster's. These methods are based on the observation that fraudsters seldom change their calling habits but are often linked to other fraudsters.

6 Telecommunication Fraud Detection Using Neural Networks

Artificial intelligence technologies have been a part of fraud detection systems in several industries for half dozen years (Liebowitz J and Prerau D S, 1995, pp 177). Hecht Nielson Neuro-computers of San Diego, California provide a neural network based fraud detection system called Falcon or credit card issuers. In the telecommunication area, a Longmont, Colorado company called Coral Systems offers a knowledge based systems for cellular fraud detection. This system uses expert fraud rules to sport potentially fraudulent calls outside the normal pattern of a caller's activity. A caller who makes calls in a rapid succession is probably the Victim of a clone phone, a common practice among cellular thieves for stealing and using a legitimate cellular amount. If a caller who makes 45 calls a day on average suddenly makes 50, no fraud alert is sounded. But if a different caller who may make only a few calls a week suddenly makes 50 in a day the account is flagged for fraud investigation. The system pulls information from carrier switches and looks for deviations from a caller's normal patterns of phone use. The system can be tuned to produce alerts for any call which looks suspicious or can be tuned to provide as few false positives as possible, which raises the fraud rate among flagged calls to 90-95% but may miss many frauds. The systems have been out since 1992 and according to Coral systems are in use by 5 customers. While this system does not currently use neural networks Coral is investigating them for future use.

7 Fraud Detection System using Artificial Neural Networks

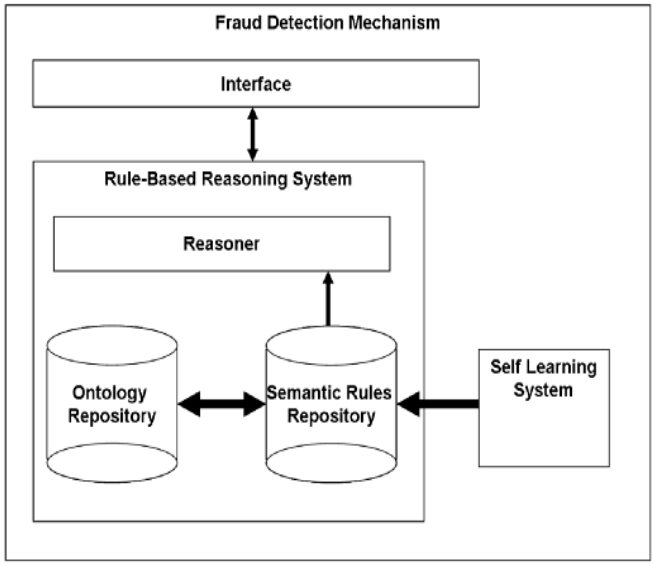

Fig. 1. Fraud Detection System using ANN Technology Source: Oodan A P (2003), Telecommunications quality of service management: from legacy to emerging services, The Institution of Engineering and Technology, London, pp 493.

An overview of the Fraud Detection System developed by using Artificial Neural Network Technology is shown in the below figure:

The process begins with gathering historical data on fraudulent and non fraudulent calls. This data is preprocessed to make it suitable for neural network learning. Next the neural network is trained using the preprocessed data to build a model which incorporates numerous patterns of fraudulent behavior. The model is applied to incoming business where it adaptively learns new patterns of fraud improving its model as the types of fraud evolves.

Fraud detection mechanisms may be home brew, proprietary or a mix of both which is probably the healthiest approach (Oodan A P, 2003, pp 493). A web search for telecom fraud management will result in a bewildering number of hits. Simple thresholding systems work well when integrated into business as usual processes and could well account for 80% or so of the fraud management load. For more complex situations where product interaction is taking place, some form of correlation is required using neural or artificial intelligence techniques. One should also consider the near real time nature of emerging frauds and the need to close down potential high cost fraud quickly. Globally the usual experience is that fraud management system pays for their investment in very short times provided they are directed at the principal sources of fraud.

Data sources for fraud detection systems are principally Call Detail Record (CDR) based, deriving usually from the host billing system. For a more real time data source the

PSTN itself is the usual source. In the narrowband world this will mean monitoring Number 7 signaling messages at strategic points in the network. Network topology then becomes an issue. In a network with widespread use of Signal Transfer Point (STP) working, signaling is concentrated in a relatively few points and monitoring may be economically infeasible so a more focused approach is called for. The usual axiom is to follow the money. Therefore monitoring international routes would be a priority. One could also consider monitoring the Intelligent Network platform as this usually provides a signaling concentration point and may also deal with high value number translation services such as Premium Rate. As voice over IP services roll out, fraud detection system manufacturers are turning their attention to the specific needs of extracting or creating CDR's from telephony services or other network components. The fraud detection mechanism consists of a rule based system for detecting fraud and a self learning system that makes this system adaptive. A rule based tool was a white box approach that allowed detecting the frauds with low rate or false alarms (Prasad R, 2001, pp 265).

More specifically the rule based system contains an ontology repository which is practically the knowledge base of the system and which stores all the domain and fraud specific knowledge that is required by the system (Filipe J and Obaidat M S, 2008, pp 112-113). The actual fraud detection methods and techniques are stored in the Semantic Rules Repository in the form of if-then-else rules that reason over the knowledge contained in the ontology repository. The two repositories are interrelated as the representation of rules depends on the contents of the ontology.

Finally the self learning system provides the overall system with the capability to learn new rules about fraud from the submission of fraudulent and non fraudulent domain specific data and to automatically detect irregular observations in the data thus providing a feedback mechanism for enriching and updating the repository of rules. The self learning system integrates suitable algorithms for statistical data analysis and data mining tasks that enable it to update, optimize and extend existing fraud detection rules by analyzing submitted data.

The fraud detection area is an active area of development for neural networks in telecommunications. Many of the most successful systems are hybrid systems which takes advantage of the relative strengths of several AI technologies. Given the payoffs involved it is an application which should come into routine use in the years ahead.

8 How does Neural Network help in Reducing Telecommunications Fraud?

The Forum of International Irregular Network Access (FIINA) estimates that telecommunication fraud results in a loss of United States $55 billion per year worldwide (Turban, 2008, pp 349). South Africa's largest telecom operator was losing over United States $37 million per year to fraud. Subscription fraud in which a customer either provides fraudulent details or gives valid details and then disappears was the company's biggest cause of revenue leakage. By the time the telecom provider is alerted about the fraud, the fraudster has already moved to other target victims. Other types of fraud include phone card manipulation which involves tampering and cloning of phone cards. In a clip on fraud a fraudster clips on to customers telephone lines and then sells calls to overseas destination for a fraction of normal rates.

Minotaur, developed by Neural Technologies was implemented to prevent fraud. minotaur uses a hybrid mixture of intelligent systems and traditional computing techniques to provide customer subscription and real time call monitoring fraud detection. It processes data from numerous fields such as event data records (switch/CDR, SS#7, IPDRs, PIN/authentication) and customer data (billing and payment, point of sale, provisioning) using a multi-stream analysis capacity. Frauds are detected on several levels such as on an individual basis using specific knowledge about the subscriber's usage and on a global basis using generic knowledge about subscriber usage and known fraud patterns. The neural capability of Minotaur means it learns from experience making use of adaptive feedback to keep up to date with changing fraud patterns. In the first three months of installation of this neural network based software: The average fraud loss per case was reduced by 40%. and The detection time was reduced by 83%.. The combination of neural, rule based and case based technologies provide a fraud detection rate superior to that of conventional systems. Furthermore the multi stream analysis capability makes it extremely accurate.

9 Conclusion

The theft of telecommunication services has been one of the most enduring types of telecommunications crime which has been evident since the beginning of telephone systems (Grabosky P N and Smith R G, 2009, pp 84). Fraud detection for mobile telecommunications is a relatively recent area of research. Other works in the area of fraud detection in mobile telecommunications are based on data mining approaches (Althoff K D, 2008, pp 563).Due to its characteristics this type of fraud requires real time and individualized customer analysis. Each technological development designed to thwart criminal endeavors has been quickly followed by the creation of a new form of crime designed to exploit new security. These few directions of future policy may assist in ensuring that the full potential of global telecommunications developments will be realized while at the same time providing both service providers and users with some expectations that their property rights will be respected.

Fraud detection will continue their accelerated use of neural network based systems. Many of the laboratory techniques for using neural networks in metro burst communications, satellite network management and traffic control will come on line. In short the neural networks will become an increasing presence in major aspects of

telecommunication networks improving efficiency, adapting to changing calling patterns, and providing better information about the use of networks. Neural networks a technology which has been used in telephony since the early 1960s is beginning to make it presence felt in designing in telecommunications network of the next century. As a result Artificial Neural Network is a better method for detecting telephone fraud, due to its inherent ability to adapt as well as its speed and efficiency.

ACKNOWLEDGMENT

I am heartily thankful to my Dean Dr. Faisal Bugdadi and head of the Department, Mr. Abdulwahab Alamri, who encouraged me to prepare this paper. I am also thankful to my colleagues and friends for their support. Any suggestions to further improvement of this topic are most welcome.

REFERENCES

[1] Pieprzyk J, Ghodosi H and Dawson E (2007), Information security and privacy: 12th Australasian conference, ACISP 2007, Townsville, Australia, July 2-4, 2007: proceedings, Springer, Germany, pp 446-447.

[2] Liatsis P (2002), Recent trends in multimedia information processing: proceedings of the 9th International Workshop on Systems, Signals and Image Processing, World Scientific Publishing, London, pp 474-475.

[3] Prasad S K, Routray S and Khurana R (2009), Information Systems, Technology and Management: Third International Conference, ICISTM 2009, Ghaziabad, India, March 12-13, 2009, Proceedings, Springer, Germany, pp 259-260.

[4] Żytkow J M and Rauch J (1999), Principles of data mining and knowledge discovery: Third European Conference, PKDD'99, Prague, Czech Republic, September 15-18, 1999 : proceedings, Springer, USA, pp 251.

[5] Wall D S (2001), Crime and the Internet, Routledge, London, pp 30.

[6] Broadhurst R G and Grabosky P N (2005), Cyber-crime: the challenge in Asia, Hong Kong University Press, Hong Kong, pp 32.

[7] Samarati P (2010), Information Security Theory and Practices: Security and Privacy of Pervasive Systems and Smart Devices: 4th IFIP WG 11.2 International Workshop, WISTP 2010, Passau, Germany, April 12-14, 2010, Proceedings, Springer, USA, pp 201.

[8] Perner P (2006), Advances in data mining: applications in medicine, web mining, marketing, image and signal mining : 6th Industrial Conference on Data Mining, ICDM 2006, Leipzig, Germany, July 14-15, 2006 : proceedings, Springer, Germany, pp 535.

[9] Liebowitz J and Prerau D S (1995), Worldwide intelligent systems: approaches to telecommunications and network management, IOS Press, Netherlands, pp 177.

[10] Oodan A P (2003), Telecommunications quality of service management: from legacy to emerging services, The Institution of Engineering and Technology, London, pp 493.

[11] Filipe J and Obaidat M S (2008), E-business and telecommunications: 4th International Conference, ICETE 2007, Barcelona, Spain, July 28-31, 2007, revised selected papers, Springer Verlag, Germany, pp 112-113.

[12] Prasad R (2001), Towards a global 3G system: advanced mobile communications in Europe, Artech House, USA, pp 265.

[13] Turban (2008), Decision Support And Business Intelligence Systems, 8/E, Dorling Kindersley, New Delhi, pp 349.

[14] Grabosky P N and Smith R G (2009), Crime in the digital age: controlling telecommunications and cyberspace illegalities, Transaction Publishers, USA, pp 84.

[15] Althoff K D (2008), Advances in case-based reasoning: 9th European conference, ECCBR 2008, Trier, Germany, September 1-4, 2008 ; proceedings, Springer, USA, pp 563.