International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

FRAUD IDENTIFICATION: FRAUD MONETARY DETECTION WITH AID OF HUMAN BEHAVIOR APPRAISAL EXAMINATION

P V GEETHA SRAVANI[1], SHAIK SIKNDAR[2]
[1]PG Student, Dept. of CSE, Eswar College of Engineering, Narasaraopet, AP, India
[2]Asst. Professor, Dept. of CSE, Eswar College of Engineering, Narasaraopet, AP, India
Corresponding authors
[1]geetha.sravani74@gmail.com [2] shaik5651@gmail.com

Abstract— Financial fraud is typically a felony punished by statute by employing unethical practices in which senior management may interfere and staff compensate. Many tools have been established for the study, identification and avoidance of this activity, the fraud triangle hypothesis identified with the classic model of financial audit being the most significant. A review has been carried out in the current literature with the intention of developing our own structure, in order to carry out this analysis. The paper introduces Fraud Locate in this sense, a philosophical construct that helps a community of those inside a banking corporation who conduct fraud with the help of the fraud triangle principle to be defined and illustrated. Fraud Find operates in the process of on-going auditing and is liable for the gathering of details from users' agents. It is based on semantical strategies used in the selection of sentences typed by the study users for further review in a repository. This initiative supports the prevention of financial crime in the area of data protection.

Keywords: fraud find, fraud triangle, human factor, human behavior

I. INTRODUCTION

Fraud is a global epidemic impacting public and private organizations, encompassing a broad spectrum of fraudulent actions and behaviors including deliberate misrepresentation or deceit. Fraud is a matter of willful or malicious deprivation of some other property or money through cunning, manipulation, or other unjust actions, according to the Organization of Accredited Fraud Investigators [1]. In 2016 the PwC Global Economy Crime Study states, about a third of global companies, including wealth misappropriation, extortion, cybercrime, theft and money-washing are victims of some form of economic crime. Around 22 percent of respondents lost about 100 000 to 1 million, 14 percent experienced losses of over 1 million and 1 percent of respondents suffered losses of 100 million dollars. The high rates of damages are a growing trend in fraud costs. 56% of cases in companies refer to internal fraud and 40% to external fraud, when any individual engaged in financial and accounting operations is deemed a possible risk factor for fraud. [2]. [2]. When looking at people's actions in the world of industry, one may infer that the human element is strongly related and aligned with the Donald R. Cressey's hypothesis regarding fraud, where three essential principles are required: pressures, opportunities and streamlining. In the trade[4],[5], as well as in the academic sector, there are today various technologies that have established some work in progress[6] and[7] that have been designed to identify financial fraud. This solutions, all focused on datamining techniques and large data, rely on the usage of various instruments to undertake computational, parametric research and conduct analyses; none of them address the question of financial fraud identification in

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

real time. Fraud Locate illegal behavior in real time, unlike every other proposal, by routine monitoring of users' data for further analysis and care identification, documenting and storage. This paper introduces Fraud Discover, a computational method for the detection and discovery of suspected offenders operating in the financial industry on the basis of a triangle of fraud theory in real time. Any of the applications components for knowledge retrieval, including Rabbit MQ, Logstash and Elastic Search have been evaluated for the designing of the Fraud Locate system. In addition, it would be necessary to find possible bank robbers at lower false positives by the computerization of the fraud triangle and the usage of semithetic techniques. The remainder of the text is organized accordingly. The analytical structure for the description of fraud and the fraud triangle principle is discussed and the relevant works contained in literature.

II. RELATED WORK

The goal of this analysis is to build an architectural model that represents fraud triangles, complements the human aspect and analyses suspect activity to distinguish potential fraud cases. For a potential task to be carried out. Several research that relate to this subject were contained in the literature.

Many records discuss the subject of financial crime and the varying situations. However, it is a deciding factor to recognize individuals that may be engaging in illegal practices. The interference into behavioral research is cited at [6] whose writers instantly email the discovery of various forms of trends in messages using an automated text mining method.

The [7] a generic architecture proposes that the factors of the fraud triangle are assisted. Furthermore, it carries out the conventional detailed study of business transactions previously utilized during the fraud identification audit. In the Internal Threat Prediction Model [8], detecting and classifying suspected theft by suspicious persons is key. One major concern is the classification of citizens by concentrating on a descriptive mining strategy [9] to reduce the internal possibility of fraud. In the battle against financial crime, the expertise of auditors still plays an important part. Some study is suggested that new system could be established to include auditors for discovery of financial misconduct in an organization, utilizing their own experiences and abilities to evaluate current knowledge and data processing techniques [10]. Another plan then creates generic mechanisms for financial fraud identification FFD,

comparing the various features of FFD algorithms based on different classification criteria [11].

New methods recognize typical values in order to enhance efficiency and precision in the identification of the unusual value of a data set by learning and adjusting clustering algorithms such as the K means [12]. Capturing irregular trends related to fraudulent behavior involves an examination of the amount of variations that can be investigated concurrently, since technical advancement has improved dramatically and can be resolved by raising the number of neurons and/or layers to the expense of wider computer usage of increasingly sophisticated neural networks. A significant consideration is the expense of manually identifying suspected fraudulent transactions. This is why, by contrasting data processing technology fully to the strongest in the best practicable way, FFD is critical for the avoidance of the destructive effects of financial fraud [14].

When analyzing the literature, it may be inferred from the expected identification of fraud that similar study would not protect because it is reviewed after the event. This document attempts to minimize this void by way of an online fraud audit by improving the approach that allows irregular activity trends in a timely manner to be detected in light of the human aspect provided by the principle of the fraud triangle. This prototype is an instrument that can be used to diagnose potential instances of financial misconduct in a company.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

III. FRAUD AND THE FRAUD TRIANGLE THEORY

Overall, the empirical concept of fraud is not available. It is regarded nonetheless as a sub-set of internal risks, including corruption, wealth misappropriation, false declarations, etc. [15]. According to the ACFE, fraud is described as "the use of one's job for personal enrichment through intentional misuse or misuse of the resources or assets of the employer's company." However, only financial crime can be considered in a banking setting because of the nature of this article. There are two kinds of deception in financial fraud: internal fraud and foreign fraud [16]. Internal theft involves a deliberate abuse by fraudsters, culminating in revenue misappropriation and other main business tools and requires a number of violations and criminal activities. This is typically done in the accounting statements that were wrongly presented in records in the case of external fraud. In these circumstances the fraudsters commit fraud actions which exploit these faults due to the weakness of internal control mechanisms and the majority of the known abnormalities. Fraud incidence is better clarified with a collection of books on criminal prevention published by a leading authority on criminal sociology, Donald R. Cressey, in The Fraud-Triangle Principle, outlined in Figure 1. Cressey explores whether individuals are committing fraud and evaluates the answer in three key areas: perceived pressure, perceived incentive and rationalization. The principle of Cressey includes the three factors needing to be present consecutively to provoke an urge to render fraud. The first required condition in the fraud triangle is that the motive and impulse behind an individual's dishonest acts are interpreted as pressure. This is also triggered by a certain level of financial tension in individuals. [17]. [17]. The second thing is the perceived risk, the activity behind criminality and the scope for fraud. The third aspect is the belief that a person may rationalize his or her unethical behavior and make his or her immoral decisions look legitimate and appropriate. When there is an increasing correlation between pressure, incentive and rationalization, the likelihood of fraud rises exponentially.

IV. FRAUDFIND FRAMEWORK

In order to find financial crime inside a Banking institution that will be our key testing area, the recommended methodology is a continuous audit method based upon the fraud triangle principle which is regarded as an integral aspect by the human component. The aim of the Fraud Find is to examine vast volumes of data from numerous information sources using the ELK stack for subsequent analysis and registration. ELK is an open source distributed framework used by Elastic Search, Logstash and Kibana apps for real time data analysis, [20] which is listed below. This is the most recent example.

1) Elastic Search is an open-source Java-based search engine that is a distributed, scalable store that runs in real-time. Mainly built to arrange data for quick access [21].

2) Logstash, which centralizes and examines a vast range of organized and unstructured data forms, is an open source platform used for event management [22].

3) The web interface of Kibana is an adjustable board to accommodate the context and can be adjusted. In addition to complex representations, it requires tables and diagrams to be created [20]. In Figure 2, the numerous modules that make up the system can be observed: Agent, QoS, Collect & Convert, Scan & Evaluate and View & Management.

A. Agent

The agent is a program built in user workstations (endpunkte) to retrieve the data they obtain from the numerous sources of data on their equipment. This application sends the user data in Rabbit MQ for ordering and classification. This application is liable. Subsequently, Logstash collects this knowledge for its care.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Figure 1. Triangle of Fraud

B. QoS

Integration of many networks or modules means that information must be obtained or transmitted, such that all messages are trustworthy, stable, rapid and above all permanently accessible. Because of the substantial and repeating amount of knowledge produced by these agents, this module ensures that the supply to Logstash is orderly and confident. In order to coordinate and disseminate the data properly, an intermediate variable was added, Rabbit MQ. Rabbit MQ is a software open source that acts as message broker to submit and receive messages from third parties providing perseverance, assurance of receiving messages and high availability. A conceptual broker for enforcing features such as a load balance and fault tolling can be built from the cluster of Rabbit MQ servers. Rabbit MQ uses the Round - Robin algorithm as default. It is withdrawn from the queue after delivery [23].

C. Collect and rework

The data submitted by the agents is stored in this node. Figure 2 indicates that after the agent details in the QoS module have been ordered, they are registered in a temporary file with raw details that Logstash doesn't understand and doesn't know how to manage. Logstash has resources called codecs and filters to view this content, which conduct operations and transformations on data gathered to turn the information into a compressible format. The data is then submitted to Elastic Search for storage when processed.

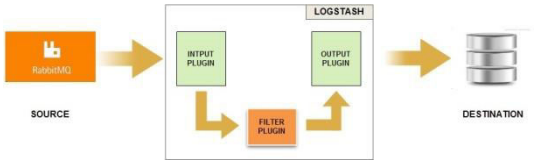

Figure 2. Logstash

D. Search and Analyze

This module has all details stored by Logstash that can be easily searched instantly upon receipt. Elastic Search is a platform focused on the concept of no defect tolerance hardware built for the clustering methodology. This property preserves and replicates the records in such a manner that the data cannot be damaged if the physical infrastructure falls down.

E. Visualize and Manage

Finally, the data found in the Elastic search are provided in this module, using Kibana for this reason. With Elastically Scan, this method enables the details to be viewed and indexed in a personalized manner utilizing histograms, pie charts, measurements and more. This method helps you to evaluate knowledge in real time.

V. FRAMEWORK IMPLEMENTATION

We define a prototype in this segment for the automated identification of financial fraud, now in its implementation process. The proposed implementation system diagram illustrates the definition of the specific modules utilized by open-source projects for practical use.

The information gathered by the agents is initially distributed by means of data queues that must be processed efficiently, securely and reliably. Rabbit MQ, an open source message broker incorporating the Advanced Message Queuing Protocol (AMQP) specification, was used to accomplish that goal. First and foremost, a logical (distributed) broker will group many Rabbit MQ servers on a local network. This helps

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

functionality such as load balance and defect tolerance to be introduced. The AMQP protocol used by Rabbit MQ makes connections between various platforms is another significant aspect.

Logstash receives details submitted by Rabbit MQ for processing (organize and categorize). Logstash is an information collection, encoding and filtering application. It consists of 3 key plugins, input, filter and exit. Next, we have a plugin for input that enables record selection in various formats, such as: archives, TCP / UDP, etc. Second, we have filter plugins that allow Logstash to perform input data transformation. Finally, the output plugin may be used to write processed or transformed data to Elastic Search in a number of formats [24].

Elastic Search receives details from Logstash that indexes this knowledge and analyses it. Elastic Such is a search and storage engine that, along with Kibana, can manage loads of data in real time and provide pace and confidence.

A job doing warning monitoring regularly verifies the entered details and compares it to a fraud triangle library to figure out whether there is a connection in the databases to produce an alert. The fraud triangle library is only a dictionary comprising three definitions: strain, chance and reasoning. The vocabulary and phrases connected with these conducts are composed under these criteria.

VI. ANALYSIS AND DISCUSSION

Performance analysis

Fraud Find includes data extracted from numerous information sources through agents installed in workstations that collect and transmit behavioral data organized to the central server. The typewritten words will then be forwarded to Rabbit MQ, an application that manages message queues and provides fast, stable and accurate data for logstash, a tool for the compilation, review, and final indexing of data from heterogeneous sources.

Much of this is intended to maintain the protection of the users' transactions who attempt by evaluating individual activity and treating the findings to detect suspected acts of fraud. Unusual behavior is not meant to ensure no theft takes place so instead an overview of the risk factors linked to this activity can be regarded, which should be measurable and weighted according to an organization's protection policies. If there is a range of record channels, the logs are unreliable since the formats are incompatible. This is a concern, since managers need access to this data for review and searching in multiple formats is challenging. The Logs are decentralized and each has a particular format and a different means of identifying them, complicating their monitoring and interpretation as spread between the different analytical teams. It gathers all these details in order to analyses them, store it distributed, and utilize care methods like Big Data to achieve reliable outcomes. ELK addresses these issues. In addition, the examination of human behavior, for this research allows the detection of transactions that are part of a trend not detected in data traffic, and are not observed through conventional methods.

Technical analysis

When finding and reviewing details from a source, ELK (Elastic Quest, Logstash and Kibana) offers simplicity and flexibility for records management. Centralized data tracking may be helpful in detecting odd trends in traffic, and helps you to easily scan all saved documents for the appropriate connection to the incident.

Security analysis

The potential loss of privacy is a concern that should be taken into consideration when introducing this approach in a business. In a specific area, laws on data security should be considered. The potential privacy breaches are a concern to be taken into consideration when this approach is developed in an organization. There should be awareness of the legislative legislation for data security in a specific area. The degree of supervision would rely on the corporate internal policy

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

and the regulations regulating each nation and can be decided on the basis of the legal part of an agency or company's advice.

VII. CONCLUSION

The current research presents Fraud Identify, a computational paradigm for financial crime focused on triangle factors of fraud that render a valuable contribution towards the early identification of fraud in an enterprise as opposed to the conventional audit research. Taking human behavior considerations into consideration, irregular transactions cannot be identified that have not been taken into account using standard techniques of auditing. The knowledge created by users utilizing the numerous applications on a workstation will find these behavior trends. The data obtained were analyzed by way of data mining methods in order to acquire trends of suspect behavior. However, a high probability of non-implementation of the architecture as an alternative is the legal system and the numerous regulations implemented in the public and private institutes of a given area. The key focus of future work would be to incorporate and review the system as a method for ongoing audit within an organization.

REFERENCES

[1] "ACFE Asociación de Examinadores de Fraudes Certificados," (Date last accessed 15-July-2014). [Online]. Available: http://www.acfe.com/uploadedfiles/acfewebsite/ content/documents/rttn-2010.pdf

[2] "PwC," (Date last accessed 15-July-2014). [Online]. Available: https://www.pwc.com/gx/en/economic-crime-survey/ pdf/ Global Economic Crime Survey 2016.pdf

[3] N. B. Omar and H. F. M. Din, "Fraud diamond risk indicator: An assessment of its importance and usage," in *2010 Inter- national Conference on Science and Social Research (CSSR 2010)*. IEEE, dec 2010.

[4] "Lynx," (Date last accessed 15-July-2014). [Online]. Available: http://www.iic.uam.es/soluciones/banca/lynx/

[5] "Ibm," (Date last accessed 15-July-2014). [Online]. Available: https://www.ibm.com/developerworks/ssa/local/ analytics/prevencionde fraude/index.html

[6] C. Holton, "Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multibillion dollar problem," *Decision Support Systems*, vol. 46, no. 4, pp. 853–864, mar 2009.

[7] S. Hoyer, H. Zakhariya, T. Sandner, and M. H. Breitner, "Fraud prediction and the human factor: An approach to include human behavior in an automated fraud audit," in *2012 45th Hawaii International Conference on System Sciences*. IEEE, jan 2012.

[8] M. Kandias, A. Mylonas, N. Virvilis, M. Theoharidou, and D. Gritzalis, "An insider threat prediction model," in *Trust, Privacy and Security in Digital Business*. Springer Berlin Heidelberg, 2010, pp. 26–37.

[9] M. Jans, N. Lybaert, and K. Vanhoof, "Internal fraud risk reduction: Results of a data mining case study," *International Journal of Accounting Information Systems*, vol. 11, no. 1, pp. 17–41, mar 2010.

[10] P. K. Panigrahi, "A framework for discovering internal financial fraud using analytics," in *2011 International Conference on Communication Systems and Network Technologies*, June 2011, pp. 323–327.

[11] D. Yue, X. Wu, Y. Wang, Y. Li, and C. H. Chu, "A review of data mining-based financial fraud detection research," in *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, Sept 2007, pp. 5519– 5522.

[12] M. Ahmed and A. N. Mahmood, "A novel approach for outlier detection and clustering improvement," in *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, June 2013, pp. 577–582.

A. Vikram, S. Chennuru, H. R. Rao, and S. Upadhyaya, "A solution architecture for financial

institutions to handle illegal activities: a neural networks approach," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, Jan 2004, pp. 181–190. H. Li and M. L. Wong, "Financial fraud detection by using grammar-based multi-objective genetic programming with ensemble learning," in *2015 IEEE Congress on Evolutionary Computation (CEC)*, May 2015, pp. 1113–1120

[13] A. Vikram, S. Chennuru, H. R. Rao, and S. Upadhyaya, "A solution architecture for financial institutions to handle illegal activities: a neural networks approach," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, Jan 2004, pp. 181–190. H. Li and M. L. Wong, "Financial fraud detection by using grammar-based multi-objective genetic programming with ensemble learning," in *2015 IEEE Congress on Evolutionary Computation (CEC)*, May 2015, pp. 1113–1120.

[14] D. Cappelli, A. Moore, R. Trzeciak, and T. J. Shimeall, "Common sense guide to prevention and detection of in- sider threats 3rd edition–version 3.1," *Published by CERT, Software Engineering Institute, Carnegie Mellon University, http://www. cert. org*, 2009.

[15] P. K. Panigrahi, "A framework for discovering internal financial fraud using analytics," in *2011 International Confer- ence on Communication Systems and Network Technologies*. IEEE, jun 2011.

[16] G. Mui and J. Mailley, "A tale of two triangles: comparing the fraud triangle with criminology's crime triangle," *Accounting Research Journal*, vol. 28, no. 1, pp. 45–58, jul 2015.

[17] D. Al-Jumeily, A. Hussain, MacDermott, H. Tawfik, G. Seeckts, and J. Lunn, "The development of fraud detection systems for detection of potentially fraudulent applications," in *2015 International Conference on Developments of E- Systems Engineering (DeSE)*, Dec 2015, pp. 7–13.

[18] S. GVK and S. R. Dasari, "Big spectrum data analysis in dsa enabled lte-a networks: A system architecture," in *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, Feb 2016, pp. 655–660.

[19] T. Prakash, M. Kakkar, and K. Patel, "Geo-identification of web users through logs using elk stack," in *2016 6th International Conference - Cloud System and Big Data Engineering (Confluence)*, Jan 2016, pp. 606–610.

[20] U. Thacker, M. Pandey, and S. S. Rautaray, "Performance of elastic search in cloud environment with ngram and nonngram indexing," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, March 2016, pp. 3624–3628.

[21] D. N. Doan and G. Iuhasz, "Tuning logstash garbage col- lection for high throughput in a monitoring platform," in *Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2016 18th International Symposium on*. IEEE, 2016, pp. 359–365.

[22] V. M. Ionescu, "The analysis of the performance of rabbitmq and activemq," in *2015 14th RoEduNet International Con ference - Networking in Education and Research (RoEduNet NER)*, Sept 2015, pp. 132–137.

[23] D. N. Doan and G. Iuhasz, "Tuning logstash garbage col- lection for high throughput in a monitoring platform," in *2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Sept 2016, pp. 359–365.

[24] X. M. Li and Y. Y. Wang, "Design and implementation of an indexing method based on fields for elastic search," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, Sept 201