**FraudFind: Financial Fraud Detection by Analyzing Human Behavior**

**Abstract**

Financial fraud is commonly represented by theuse of illegal practices where they can intervene from seniormanagers until payroll employees, becoming a crime punishableby law. There are many techniques developed to analyze,detect and prevent this behavior, being the most importantthe fraud triangle theory associated with the classic financialaudit model. In order to perform this research, a survey of therelated works in the existing literature was carried out, withthe purpose of establishing our own framework. In this context,this paper presents FraudFind, a conceptual framework thatallows to identify and outline a group of people inside anbanking organization who commit fraud, supported by thefraud triangle theory. FraudFind works in the approach ofcontinuous audit that will be in charge of collecting informationof agents installed in user’s equipment. It is based on semantictechniques applied through the collection of phrases typed bythe users under study for later being transferred to a repositoryfor later analysis. This proposal encourages tocontribute withthe field of cybersecurity, in the reduction of cases of financialfraud.

**Architecture**

****

**Existing system**

A key aspect is to classifyindividuals by focusing on reducing the internal risk of fraudthrough a descriptive mining strategy.Besides, the experience of auditors plays an importantrole in the fight against financial fraud. Some work isproposed which points to the creation of new frameworksthat provide systematic processes to help auditors to discoverfinancial fraud within an organization by analyzing existinginformation and data mining techniques using their own

experience and skills. Accordingly, another proposalcreates generic frameworks for the detection of financialfraud FFD, to evaluate the different characteristics of FFDalgorithms according to a variety of evaluation criteria.

**Proposed system**

The proposed framework operates in the continuous auditingapproach to discover financial fraud within an organizationbelonging to the banking sector which will beour main study environment and also focused on the fraudtriangle theory with the human factor considered as anessential element. FraudFind is proposed with the objectiveof analyzing large amounts of data from different sources ofinformation for later processing and registration. The agent is an application installed in the workstationsof the users (endpoints), in order to extract the data thatthey generate from the different sources of information thatreside on their equipment. This application is responsiblefor sending the data entered by the user forordering and classification. Later this organized informationis received by Logstash for its treatment.

**Module Implementation**

1. **Agent**

The agent is an application installed in the workstationsof the users (endpoints), in order to extract the data thatthey generate from the different sources of information thatreside on their equipment. This application is responsiblefor sending the data entered by the user forordering and classification.

1. **Behavior analysis**

If we are given a set of patterns or a set of feature vectors for some set of population then we would like to know if the data set has some relatively distinct subsets or not. In this context we can define cluster analysis as a classification technique for forming homogeneous groups within complex data sets. Typically, we do not know a priori the natural groupings or subtypes, and wewish to identify groups within a data set. We wish to form classifications, taxonomies, or typologies that represent different patterns in the data.

1. **Fraud detection**

Behavioral analytics solutions are designed to understand the normal behavior of each individual account holder, calculate the risk of each new activity and then choose intervention methods commensurate with the risk. The key characteristics that make behavioral analytics effective are automatically monitoring all activity for all account holders, not just devices or transactions; no requirement for prior knowledge of the specific fraud that the perpetrator is attempting; and providing detailed historical context for suspicious activity.

1. **Fraud category**

Periodically, a task that do the alert tracking, checks the information entered and compares it with a fraud triangle library to determine if there is a relation in order to generate an alert that will be stored in the database. The library of the fraud triangle is just a dictionary that contains three definitions: pressure, opportunity and justification. Under these parameters, the sentences and words associated with these behaviors are composed.

**Algorithm**

1. **K means clustering**

K-Means clustering intends to partition *n* objects into *k* clusters in which each object belongs to the cluster with the nearest mean. This method produces exactly *k* different clusters of greatest possible distinction. The best number of clusters *k* leading to the greatest separation (distance) is not known as a priori and must be computed from the data. The objective of K-Means clustering is to minimize total intra-cluster variance, or, the squared error function.

**System Requirements**

# H/W System Configuration:-

#  Processor : Intel (R) Pentium (R)

Speed : 1.1 Ghz

RAM : 2GB

Hard Disk : 57 GB

Key Board : Standard Windows Keyboard

Mouse : Two or Three Button Mouse

Monitor : SVGA

#  S/W System Configuration

* Operating System : Windows 8/7/95/98/2000/XP
* Application Server : Tomcat5.0/6.X/8.X
* Front End : HTML, Java, Jsp
* Scripts : JavaScript.
* Server side Script : Java Server Pages.
* Database Connectivity : Mysql.
* Java Version : jdk 1.8

**Conclusion**

The present work proposes Fraud Find, a conceptual framework to detect financial fraud supported by the fraud triangle factors which, compared to the classic audit analysis, makes a significant contribution to the early detection of fraud within an organization. Taking into account human behavior factors, it is possible to detect unusual transactions that would have not been considered using traditional audit methods. These patterns of behavior can be found in the information that users generate when using the different applications on a workstation. The collected data is examined using data mining techniques to obtain patterns of suspicious behavior evidencing possible fraudulent behavior. Nevertheless, the legal framework and the different regulations that are applied in public and private institutions of a particular region represent a high risk for the non-implementation of this architecture as an alternative solution. Future work will have as its main objective the implementation and evaluation of the framework as a tool for continuous auditing within an organization.