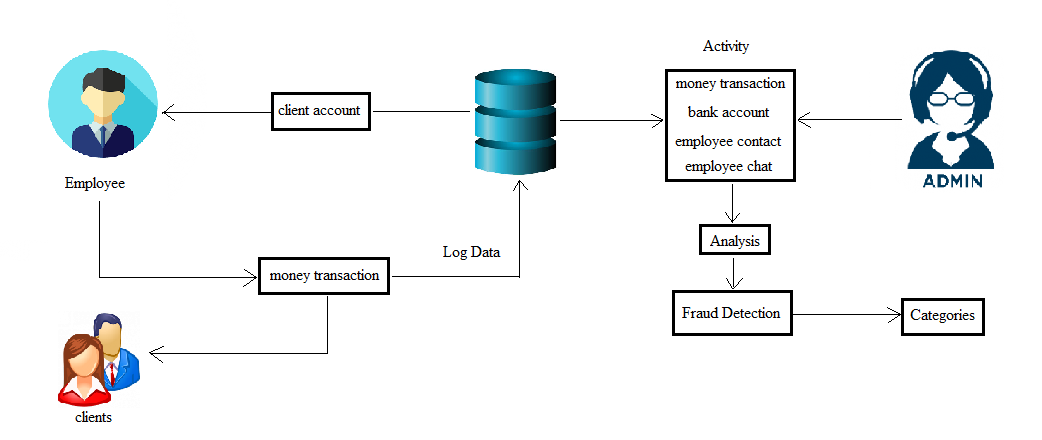
**FraudFind: Financial Fraud Detection by Analyzing Human Behavior**

**Abstract**

Financial fraud is commonly represented by the use of illegal practices where they can intervene from senior managers until payroll employees, becoming a crime punishable by law. There are many techniques developed to analyze, detect and prevent this behavior, being the most important the fraud triangle theory associated with the classic financial audit model. In order to perform this research, a survey of the related works in the existing literature was carried out, with the purpose of establishing our own framework. In this context, this paper presents Fraud Find, a conceptual framework that allows to identify and outline a group of people inside an banking organization who commit fraud, supported by the fraud triangle theory. Fraud Find works in the approach of continuous audit that will be in charge of collecting information of agents installed in user’s equipment. It is based on semantic techniques applied through the collection of phrases typed by the users under study for later being transferred to a repository for later analysis. This proposal encourages to contribute with the field of cyber security, in the reduction of cases of financial fraud.

**Architecture**

****

**Existing system**

A key aspect is to classify individuals by focusing on reducing the internal risk of fraud through a descriptive mining strategy. Besides, the experience of auditors plays an important role in the fight against financial fraud. Some work is proposed which points to the creation of new frameworks that provide systematic processes to help auditors to discover financial fraud within an organization by analyzing existing information and data mining techniques using their own

experience and skills. Accordingly, another proposal creates generic frameworks for the detection of financial fraud FFD, to evaluate the different characteristics of FFD algorithms according to a variety of evaluation criteria.

**Proposed system**

The proposed framework operates in the continuous auditing approach to discover financial fraud within an organization belonging to the banking sector which will be our main study environment and also focused on the fraud triangle theory with the human factor considered as an essential element. Fraud Find is proposed with the objective of analyzing large amounts of data from different sources of information for later processing and registration. The agent is an application installed in the workstations of the users (endpoints), in order to extract the data that they generate from the different sources of information that reside on their equipment. This application is responsible for sending the data entered by the user for ordering and classification. Later this organized information is received by Log stash for its treatment.

**Module Implementation**

1. **Agent**

The agent is an application installed in the workstations of the users (endpoints), in order to extract the data that they generate from the different sources of information that reside on their equipment. This application is responsible for sending the data entered by the user for ordering 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 we wish 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 Specification**

Software Requirements:

* Operating System : Windows/XP/7.
* Application Server : Tomcat 6.0/8.0
* Soft wares : HTML, Java, Jsp, Java scripts
* Database : Mongo DB
* Database Connectivity : Robomongo-0.8.5-i386.

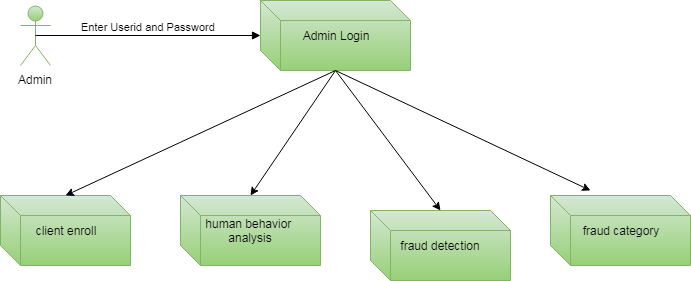
Hardware Requirements:

* Hardware - Pentium
* Speed - 1.1 GHz
* RAM - 2 GB
* Hard Disk - 20 GB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

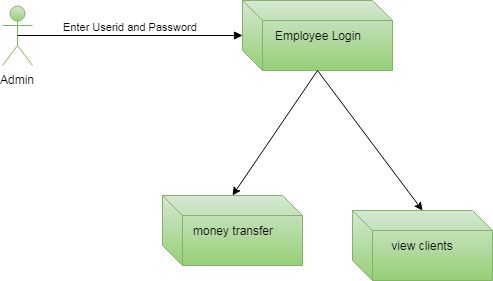
**System Design**

**Component Diagram**

**Admin**

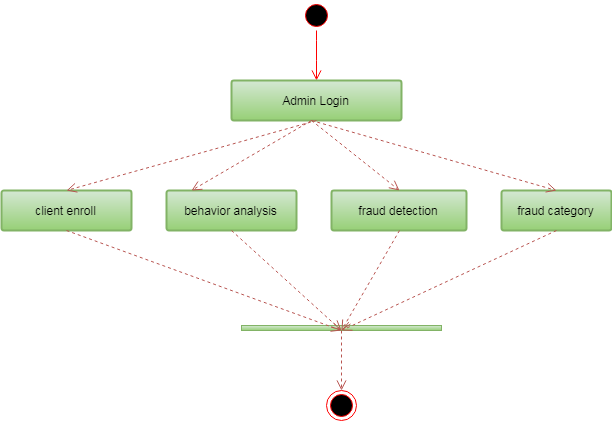
****

**Employee**

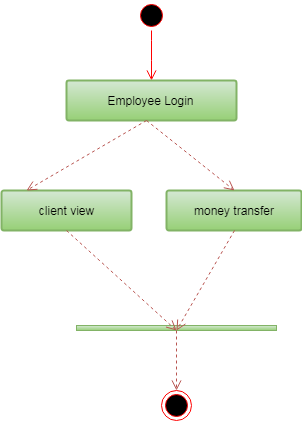
****

**Activity Diagram**

**Admin**

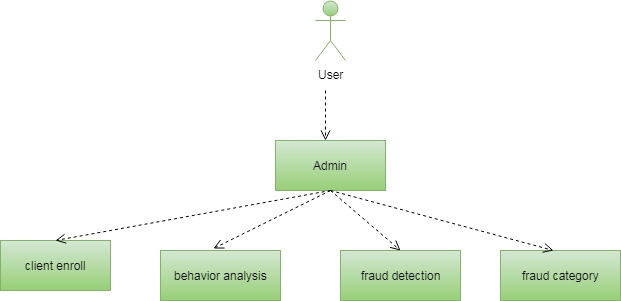
****

**Employee**

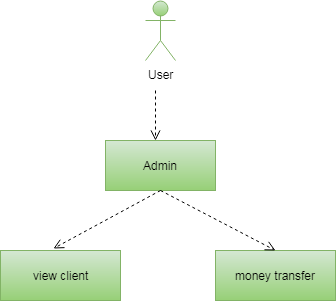
****

**Use case Diagram**

**Admin**

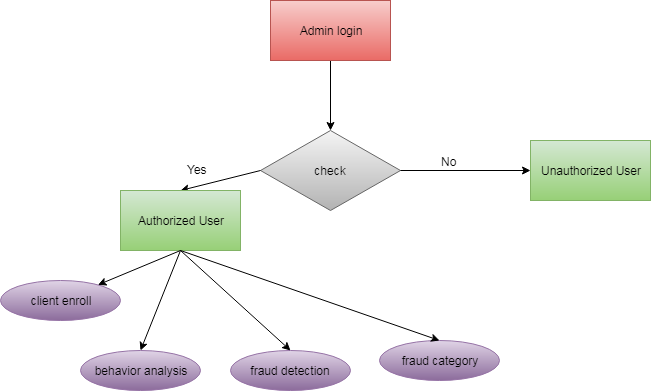
****

**Employee**

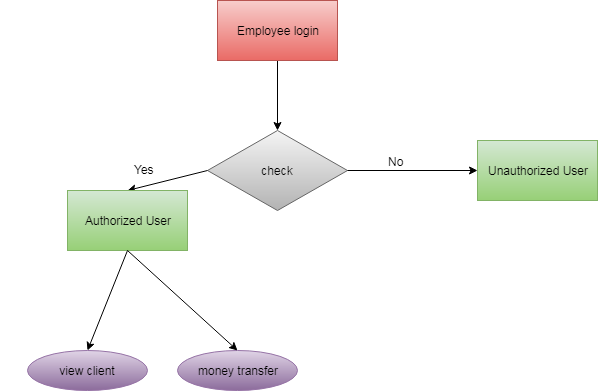
****

**ER Diagram**

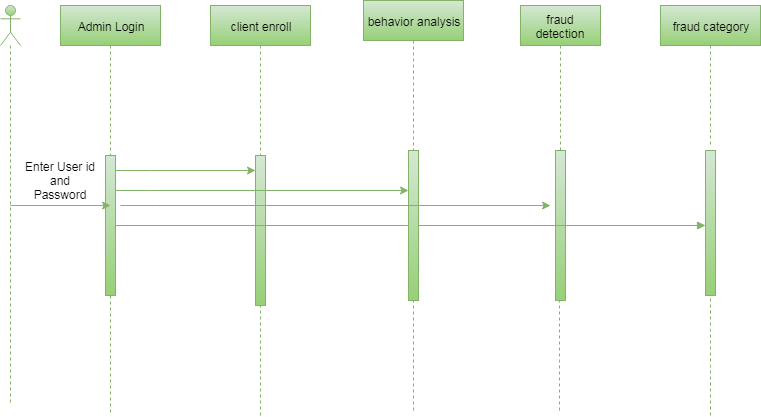
**Admin**

****

**Employee**

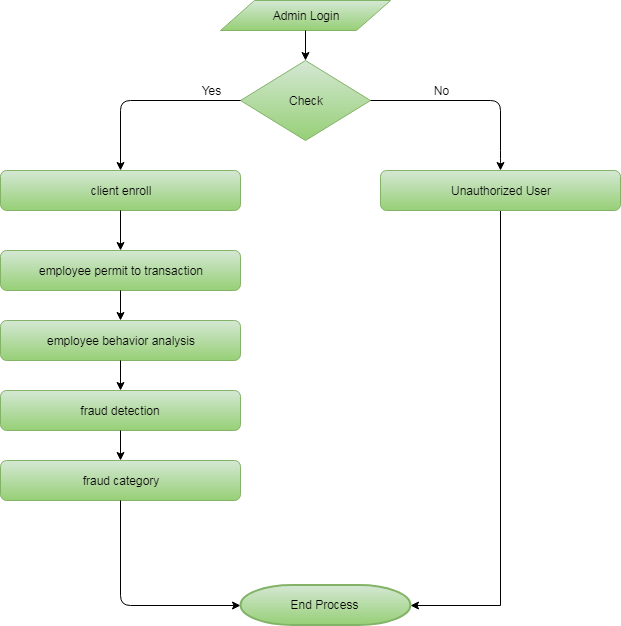
****

**Sequence Diagram**

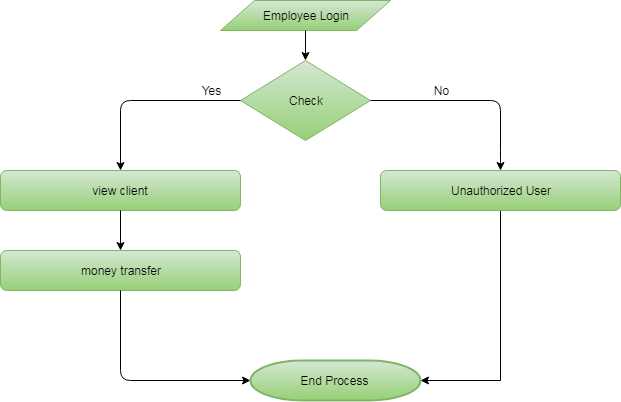
****

**Data Flow Diagram**

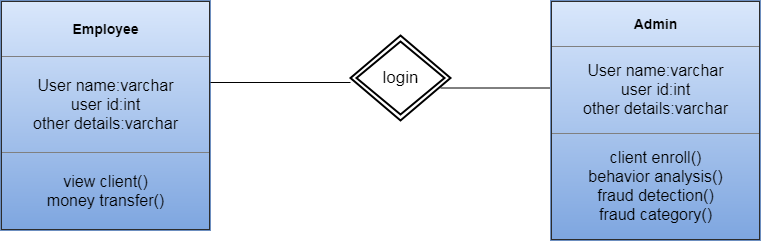
**Admin**

****

**Employee**

****

**Class Diagram**

****

**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.