KCA UNIVERSITY

FACULTY OF COMPUTING AND INFORMATION MANAGEMENT

# A MODEL TO DETECT AND PROTECT TOLL FRAUD IN VoIP PBX INFRASTRUCTURE

**BY**

**ALOISE KYENGO MULI**

**RegNo: KCA 15/02903**

## MASTER OF SCIENCE IN DATA COMMUNICATIONS AND NETWORKING

**A research project submitted in partial fulfillment of the requirement for the award of the degree of Masters of Science in Data Communication and Networks**

**2017**

**DECLARATION**

I declare that this research project is my original work and has not been published previously or submitted elsewhere for award of a degree. I also declare that this research project contains no material written or published by any other person except where due citation is made and author duly acknowledged.

Student's Name: Aloise Kyengo Muli Reg. No. 15/02903

Sign:_____ Date: _____

I do hereby confirm that I have examined this master's research project of

**Aloise K. Muli.**

And have certified that all revisions that the research panel and examiners recommended have been adequately addressed.

Sign: _____ Date: _____

Supervisor's Name:

# A MODEL TO DETECT AND PREVENT TOLL FRAUD IN VoIP INFRASTRUCTURE

## CHAPTER ONE: INTRODUCTION

### 1.1 Background

Voice over Internet Protocol (VoIP) is one of the fastest growing Internet applications. It is a more flexible technology compared to the traditional telephony systems due to its high resource utilization and cost efficiency. Also its management as well as administration is relatively easier.

Convergence in communication technologies has resulted to multimedia systems that offer these services: audio, video, and data both on wired and wireless networks. The multimedia infrastructure form the next generation telecomunication networks, whereby all should be IP based.

In large and medium organizations, normal practice is an infrastructure of two communication networks: voice network, and data network. The voice network is for internal communications (INTERCOM), offered by a Private Branch eXchange (PBX). The data network is an organization's LAN (local area network) for the computing services.

APBX in most cases is a small scaletelephone exchange within acorporate business thatinterfaces PSTN (Public Switched Telephone Network) with a company private voice network. Most companies need fewer lines than the number of employees, hence fewer outgoing lines, but many extensions so that employees can converse internally. The PBX calls can be categorized as:

- Intra-office: extension to extension connection
- Local area calling (between the local telephone exchange and the PBX)
- Long distance (toll) calling (national/and international)

The PBX is voice switching (exchange) system. Switching systems (or simply switches) form communication networks that provided a flexible interconnection between two end users to interchange voice messages. The PBX has history of its evolution from manual analogue, automatic analogue, to automatic digital, and now automatic digital on IP standards, i.e. VoIP technologies.

IP telephony switching and routing elements introduced new set of security attacks different to the PSTN attacks.

This study is on call toll fraud in VoIP PBX. Toll fraud is making calls in a telephone system without paying call charges or expecting paying reduced call bills or coercing another party to pay the bills. Telecommunication fraud is cyber crime and statistics as shown that as the usage of

voice systems increase so is the toll fraud menace. In reference to the Communications Fraud Control Association (CFCA) Telecommunication Fraud Survey, annual global Telecommunication Fraud losses in the year 2015 are estimated $38.1 Billion (USD). The report shows the first top five countries from which fraud originates are US, Pakistan, Spain, Cuba and Italy. The report also give Cuba, Somalia, Bosnia & Herzegovina, Estonia and Guinea as the top five countries were fraud terminate.

Africa cyber Security report 2016 gives estimated cost of cyber crime of $2Billion. A number of selected countries show fraud costs given: Nigeria with $550M, Kenya has $175M, Tanzania with $85M, Ghana with $50M and Uganda has $35M, all in USD.

The Kenya Cyber Security Report 2016 shows an increase of cyber crime cost from $150 USD in the year 2015 to $175 USD in year 2016. The story of call fraud is long starting from the time of manual telephone exchanges to current automatic IP based voice systems.

The digital systems offer wide range of security issues and challenges, but their worst enemies are the service providers (assault from within!). The switches are customized computer systems and the subscribers are limited in knowhow of the systems designs and operations, thus, external customer tempering is low. Also the signal being in digital form needs special tools to convert to voice form. Many frauds such **removing billing counters** (meters), changing user identity (spoofing), diverting traffic (DoS), etc. are configuration that unfaithful system administrators/technicians can invoke and misuse a voice call service system.

## 1.2 Statement of the Problem

Toll fraud is the misuse of a telephone system to make free or reduced charge calls or sale calls to another party (Voxox VoIP Technology, 2015). It is theft of long-distance service by unknown 3rd party. The "thieves" hack VoIP PBX and allow calls to be routed through the system to high rate international or premium rate telephone numbers operated by criminals. If a call has originated from or passed through customer equipment, that customer is responsible for the charges associated with the call, whether the call is authorized or not. This means that if you are the victim of toll fraud, you are liable for the costs (TELUS Communications Company, 2017).
Typically, fraudsters will find an IP phone or IP PBX that is not secure and they harvest the signaling credentials (username, password and proxy address) to make calls on that account (ITSPA, 2016). Availability of applications such as spoofing tools and voice distortion make easy to attack poor protected voice networks.
Toll fraud is one of the biggest risks facing business operating in the voice space or for end user organizations with IP PBXs. Today toll fraud is a top security challenge to corporate VoIP PBX. Hijackers command the IP PBX and make free calls, or route calls to other parties, etc. The business under all circumstances must pay for the calls despite it is not the one that has made the calls. Many organizations don't expose the menace of call fraudulence in fear of negative reputation (Koiser, 2016). CFCA Telecommunication Fraud Survey reported in the year 2015 an

annual global fraud loss estimate of $38.1 Billion USD. Pindrop Call Centre Report of 2017 show the call centre fraud rate increase of 113% from 2015 to 2016. It cites the major cause of the increase as that the financial institutes, insurance companies etc. have tightened their security on the online services (credit cards), and hence fraud business has shifted to weak protected voice services.

Toll fraud can earn malicious hackers billions of dollars per year and causes more damage than credit card fraud. FBI (2009) reported toll fraud ring in Philippine that targeted telephone exchanges in US and stolen 12 million minutes of toll calls. New York Times news paper (2014), reports cases of 2012, where hackers hijacked the phone lines of 26 small businesses around Albany and ran up phone bills as high as $200,000 per business over the course of few days. Avaya Inc(2014) reports a case where a retail shop gets a phone bill of $500,000, an amount 400 times its typical monthly bill. In Kenya, VoIP PBX attacks increased by 73% from 450,000to 700,000 between 2012 and 2013 (Serianu Limited, 2014). The figures show death of, or severe damage to, an otherwise health business can happen in minutes!

Apart from financial negative effects, the fraud other effects may be: loss of service, loss of customer confidence, or cause distress (Koiser, 2016). There is increase in VoIP usage which is also directly proportional to the increase of VoIP fraud, which is a great threat to the emerging VoIP technology.

The migration is there, from traditional telephone systems to packet switching networks of which all services are on IP platform. The IP infrastructure is an open system, hence more prone to fraud. The narrated issues above show that toll fraud is a significant security challenge and a growing problem in telecommunication services. It needs proactive addressing. This growing challenge called for development and implementation of a toll fraud detection and prevention system, stop or limit the losses incurred by businesses.

## 1.3 The Study Aim and Objectives

### a) General Objective

The aim the research was to develop a model to detect and prevent toll fraud in VoIP PBX infrastructure.

### b) Specific objectives

i. To find out the existing techniques for toll fraud detection systems
ii. To design and implement a toll fraud detection and prevention model
iii. To evaluate the toll fraud detection and prevention model.

## 1.4 Research questions

1. What are the existing approaches for fighting toll fraud in VoIP PBX infrastructure?
2. How can toll fraud be detected and eliminated/mitigated in IP PBX infrastructure?
3. How can the degree of detection and mitigation of toll fraud in VoIP PBX be assessed?

## 1.5 Significance of the research

This research ended up in the development of a model technique for detecting and preventing /mitigating call toll fraud. The research has contributed to the industry a technique for detecting and preventing call fraud. This will lead to reduction of global losses of billions of money lost in telephony fraud. The contribution is a motivation to the corporate business to migrate to IP PBX, and then reap VoIP many benefits.

The beneficiaries from the research include the corporate business on reduced call bills, telecommunication service providers to have reduction in interconnection levies, communication authorities have a tool to monitor toll fraud, subscribers no longer have inflated call bills, and last but not least, the academia world have a basis for further research (contribution to the body of knowledge).

### 1.6 The research Justification/Motivation

Toll call fraud is making businesses loss lot cash through paying bills of calls they have not made. Hackers are paralyzing companies and hence the economy, affecting production, employment, peoples well being, etc. Apart from financial negative effects, the fraud other effects may be: loss of service, loss of customer confidence, or cause distress (Koiser, 2016). There is increase in VoIP usage which is also directly proportional to the increase of VoIP fraud, which is a great threat to the emerging technology.

### 1.7 Scope of study

The study focused on studying toll fraud on IP PBX, a family member of telephony systems. The success of the study can be extended to the other major switching systems of the PSTN. The tasks were undertaken on a limited to a test model, minimum hardware and software resources. Live calls we tested on working field VoIP PBX.

# CHATPER TWO: LITERATURE REVIEW

## 2.1 Introduction

This chapter discusses related studies conducted on toll fraud in telecommunication services, devices and systems; and is more biased on VoIP infrastructures. The review focuses on bringing out good insights into the concepts of toll fraud detection and prevention/mitigation.

Voice over Internet Protocol (VoIP) is an emerging technology, and its growth rate is high among Internet applications. The VoIP system is on packet switching technology, and offers high resource utilization and cost efficiency compared to the traditional telephony system which is on circuit switching technology.

Although a new technology, various methods had been used to detect and prevent call fraud on voice services. Some of these strategies are discussed in this section.

## 2.2 Theoretical Review

### 2.2.1 Toll Fraud

Toll calls are those calls that are charged (billed). The most common modes of telephone billing are the prepaid and post paid. Prepaid mode is based on advance paying before the services are offered such as airtime tokens. This mode is more prevalent with individual subscribers. Most of corporate business are on post paid mode were calls are made and billing is at the end of the month. Telephone thieves take advantage of this, they make billed calls after hijacking a telephone system and they don't pay. They resell the calls they make. This crime is the toll fraud, and occurs mainly when the business is closed, i.e. at night, early hours of morning or weekend or holidays. The criminals use auto-dialers to identify systems which are easy to hacker into. It is common after the installation of PBX the system is left on default security settings (administrator user and password) (Yu, 2015).

In VoIP security classification, four aspects are considered. The attributes are:

- Confidentiality: privacy during speech/data transmission
- Integrity: no content change; protect both signaling and payload
- Authentication: both caller and called are whom they claim they are
- Availability: always accessible (24/7) by the right user, no denial of service (DoS).

Toll fraud theft falls under authentication feature. A hacker fakes a user ID and initiates a call from the caller system for financial gains. The criminal's common practice is to scan an IP phone or IP PBX that is not secure and get the username, password, IP address to make toll calls (ITSPA, 2016).

### 2.2.2 VoIP Architecture

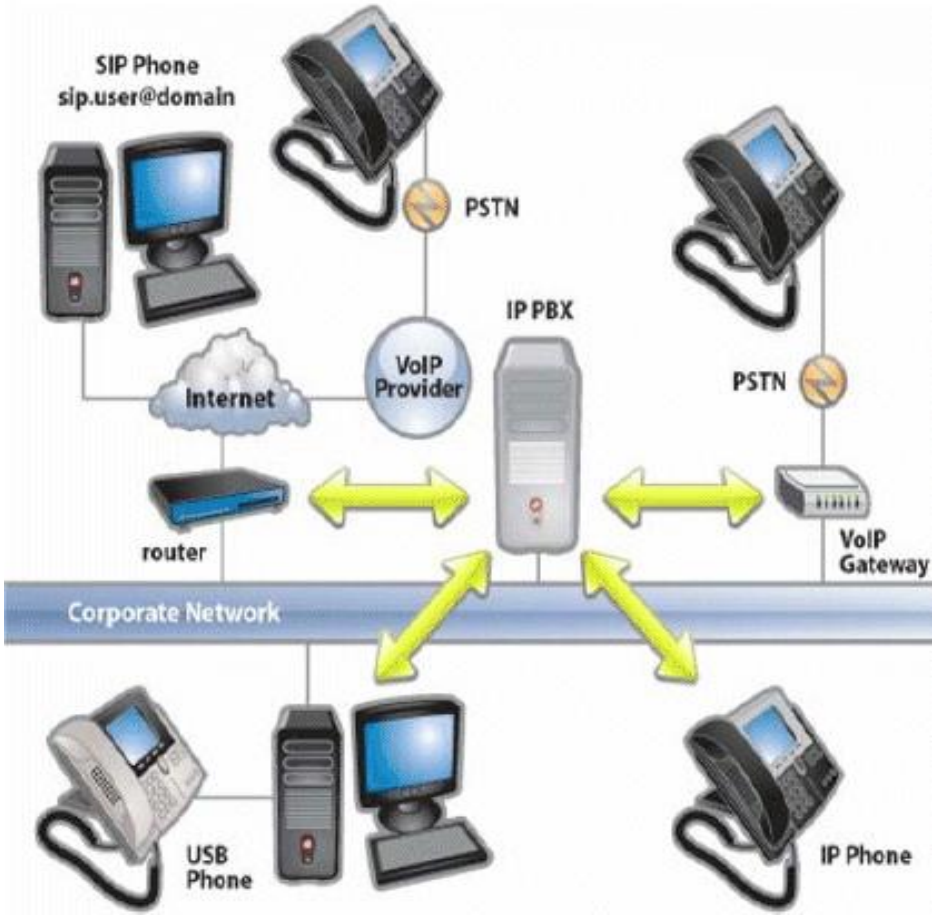



*Figure 2.1 Corporate VoIP Infrastructure (Source: 3COM & Frost & Sullivan)*

The main elements of a VoIP network are the usual packet switched (data) network components; LAN switches, routers, servers (DNS, DHCP, etc), firewalls, modems, etc. The differences are in additional devices for the sake of processing voice over the data network. The additions are VoIP user devices and network elements.

To understand a call session let us get how the process goes: a call connection involves two phases, signaling phase and speech/payload phase. Signaling phase is the establishing a physical/logical link between the caller and the called. It is called a call setup. It takes a defined time depending on the distance between the calling and called parties, link speed and the signaling protocol used.

The infrastructure of VoIP is on various protocols. These protocols are such as SIP (session initiation protocol), H.323 and MGCP) media gateway control protocol.

In speech/payload phase; it is the most important session where user information is passed between the sender and the receiver. Speech is a real-time duplex communication where a sender in caller's device transmits to a receiver in called party device, and a sender in called party device transmits to a receiver in calling party device. The exchange (dialogue) continues until the end of the conversation.

The protocols used in this media exchange phase are; RTP (real-time transmission protocol), SRTP (secure RTP), etc.

There are utility protocols used in services such as name address resolution (DNS), dynamic IP assigning, etc.

### 2.2.3 Session initiation protocol (SIP)

The most common signaling protocol used in VoIP environment. SIP had been developed by IETF (Internet Engineering Task Force) and it is a de-facto signaling protocol in VoIP (Rasol et al., 2016). SIP is multi-platform protocol, an open standard and applications can be written to customize SIP uses (Cisco, 2006). The SIP system involves clients known as user agents (UAs) and one or more SIP servers. The figure below shows an example of a SIP call set up.

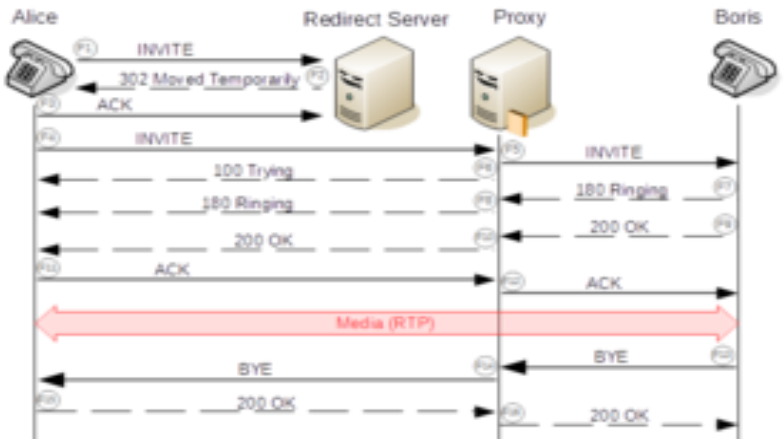



Figure 2.2 SIP call connection

Depending on site installation, a single hardware server may handle user registration, proxy, location, redirection, and/or presence services. The UAs IP phones or PC loaded with soft phones. SIP server can offer proxy services to external SIP servers, such as those in carrier networks that enable hosted services or VoIP trunk interfacing.

## 2.3 Empirical Review

The empirical review is directed on the previous findings on toll call fraud detection and prevention/mitigation. The review is also extending on telecommunication frauds because VoIP is a subset of telecommunication services.

A number of methods had been applied to fight fraud in telecommunication infrastructure. These approaches determine the variables (independent) that contribute to telecommunication frauds (dependent variables). Most of the methods depend on CDR (call detail record) data to create user profiles which are used to detect any variations, deemed as fraud. The following are the past frequently used techniques used for toll fraud:

1. Rule-based approach: compare current and past user profiles
2. Neural networks: more adaptive to user behavior, supervised or unsupervised techniques
3. Visualization methods: human pattern recognition, real-time network element monitoring

Following is explanation of some selected toll fraud detection methods used by various researchers.

### 2.3.1 Machine Learning

The method gives computers the ability to learn without being explicitly programmed(Samuel A, 1959). It explores the study and construction of algorithms that can learn from and make prediction on data. The technique is employed in range of computing tasks where designing and programming explicit algorithms with good performance is difficulty or unfeasible. Examples of applications are such as e-mail filtering, detection of network intruders, etc.

Machine learning depends on the applied input training signal or feedback signal. The machine jobs are normally grouped into three wide types:

i. Supervised training

ii. Unsupervised training and

iii. Reinforced training.

In machine learning a target is called a label, and in statistics it is called a dependent variable.

### 2.3.2 Supervised learning algorithms

This is trained using labeled samples (compare with a lookup table). The algorithm learns by comparing its actual output (o/p) with correct o/p to find errors.

### 2.3.3 Unsupervised learning algorithms

They are used against data that has no historical labels. In this approach, the system is not told the “right answer”, but it is the algorithm to figure out what is being shown. The goal is for the system to explore the data and find some structure (patterns).

### 2.3.4 Bayesian network

This is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed cyclic graph. Example model is a relationship between diseases and symptoms.

### 2.3.5 Neural Networks (NN)

This is an artificial intelligence (AI) technique that mimics the operation of the human brain (nerves and neurons). It comprises of densely interconnected computer processors working simultaneously (in parallel). Their key feature is; they are programmed to "learn" by sifting data repeatedly, looking for relationships to build mathematical models, automatically correcting these models to refine them continuously, Koiser (2016).

NNs can overcome the limitations of the rule-based method. It is more effective against unknown types of toll fraud. Their disadvantage is, have difficulty presenting the interaction of cause and effect of detection.

### 2.3.6 Honeypot/honeynet

#### 2.3.6.1 *Honeypot*

Honey pot is an enticing source of pleasure or reward such as honey in a pot.In computing, it is a computer system that is set up to act as a decoy to lure cyberattackers, and to detect, deflect or study attempts to gain unauthorized access to information systems. Generally it consists of a computer, applications, and data to simulate behavior of a real system that appears to be part of a network but is actually isolated and closely monitored. It is set up to attract and "trap" people who attempt to penetrate other people's computers, (Hoffstadt et al, 2012).

#### *2.3.6.2 Honeynet*

This is a network that contains one or more honeypots. The network diverts attackers from real network and its resources. Applications in honeypots are given appealing (attracting) names such as "Finances", "Customer Accounts", Human Resources", etc. (Gruber et al, 2011).

### 2.3.7 User profiling

A process that refers to construction of a user profile via the extraction from a set of data (Wiens et al., 2015). A user profile stores the description of the attributes of a person. Call detailed record (CDR) is a good example to give. The user proofing and behavior pattern recognition a good statistical online methods that may be chosen for collecting call data.

Two techniques common: differential and absolute analysis. User profile (characteristics) such as: user ID, call source/destination (local, national, international, and special call), call duration, cost per call, number of calls per subscriber, call delays, etc.

### 2.3.8 Outlier

In statistics, outlier is an observation point that is a distant from other observations (Kim et al., 2015). A mixture of two distributions, where some data points are further away from the sample mean. That means extreme measures: maximum or minimum samples; called outliers.

Local outlier factor (LOF) algorithm is used to analyze the extreme points, e.g. fraud detection TPRs and FPRs.

### 2.3.9 Clustering

It is a process of organizing objects into groups whose members are similar in a way. It is a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other groups.Data can be divided using similarity criterion in distance; two or more objects belong to a cluster if they are close according to a given distance.

The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. Clustering algorithms are categorized: exclusive, overlapping, hierarchical, and probabilistic clustering.

### 2.3.10 Rule-based/Rule-engine

It uses predefined rules developed by experts. An alert is raised when a rule is fulfilled. If the rules are well defined the method is effective against toll fraud. The disadvantage of rule-based technique is that it is ineffective against unknown types of fraud.

## 2.4 Previous Approaches to Toll Fraud as before year 2013

Previous researchers have employed different approaches to fight toll fraud in telephony. Rosset et al. (1999) proposed a rule-discovery framework for call criminal deception detection, in which candidate rules are identified first and the most relevant rules are selected on the basis of a suggested algorithm. A rule-based fraud detection framework for VoIP services was proposed by Ruiz et al. (2010). In the structure, a rule engine is generated using a knowledge base.

Artificial intelligence (AI) methods in form of neural networks (NNs) were applied to care of the weaknesses of rule-based solutions. Despite the NNs being more effective against unknown types of toll fraud, they also have limitations. NNs have difficult presenting the interaction of cause and effect of detection. A recurrent neural networks method based on unsupervised

learning was employed by Burge et al. (1997). Tanigushi et al., 1998, suggested a feed forward NNs approach, a Gaussian mixture model, and Bayesian net.

These proposals, (Tanigushi et al., 1998), (Burge et al., 1997), (Grosser et al., 2005),Moreau et al, 1997), (Kapourniotis et al., 2011), and (Fawcett et al., 1997) have provided different techniques of customer data for the detection and prevention of call theft in VoIP networks.

Hollmen et al. (1999) advanced a framework for SOMs to group predictive patterns. Subscriber data records from mobiles communication nets were applied to validate the mode. Graphically the maps are displayed to categorize malicious and normal calls.

The proposal of Alves et al. (2006)dealt with call fraud using two signature approaches. One technique is profiling user present behavior against his/her signature.The other is analysis of subscriber dynamic groups. The variations in the two methods provide a detection of call fraud. The findings yielded success rates of 75% for the first method and, 91% for the second method.

Quite a number of previous work discuss techniques that create profiles from labeled data, train machine learning algorithms and use the result for evaluation of the data in telephony fraud detection. The methods demands specialized knowledge and are also involving as well as time consuming.

Chandola et al.(2009)proposesgeneral view information on anomaly detection as well as fraud detection in for telecommunication networks.

The methods: statistical approaches, NNs and rule-based strategies are common with most of previous works.

Burge et al. (1997) proposed a combination of two methods: NNs and statistical, both depending on user profiling. The system compares the two outputs on past and present behaviours. Hellinger distance is applied in the probability distribution in call clusters.

A mix of NNs and call type grouping to detect fraudulent calls was proposed by Hilas et al. (2008). The technique used supervised and unsupervised training approaches. The technique attributes are: standard deviation, maximum and mean values of calls, calls duration and maximum cost per call.

Gao et al.(2011) proposed applying three approaches: identity authentication, key process monitoring and abnormal call trapping to fight toll fraud.

Fawcet et al.(1997) proposed a technique that works on unsorted data. It scans unusual user behavior, analyze the inputs, and send an alert in case of fraud. The approach needs expert knowledge as well as lot resources, hence disadvantage in fraud detection.

Burge et al.(1997) and Grosser et al.(2005) used SOMs instead of adaptive prototypes. Still the new system lacked the ability to be autonomous.

A Bayesian network and user profiling approach to detect fraud was proposed by Kapourniotis et al. (2011). Their data analysis applied these features of a CDR: Destination Country, Duration, Call Day and call type.

## 2.5 State of the Art in VoIP Toll Fraud

We would like to discuss call fraudulence works done between year 2013 and 2017. The studies are mainly referenced by the methods used by the researchers.The following works had been reviewed:

- Honeynet (Gruber et al, 2013)
- User CDR Profiling ( Wiens et al, 2014)
- User CDR Profiling (Wiens et al, 2015)
- Artificial neural networks (Koiser, 2016).

### 2.5.1 Honeynet method

The team of Gruber, Schanes, Frankauser and Grechenig (2013), analyzed the then status of toll fraud from call detailed records (CDR) collected from an online VoIP honeynet. Honeynet collect information about fraudulent calls, and the data is helpful in analyzing the behavior and security a whole (details on honeynet refer section 2.3.6). A detailed analysis was done on signaling data and media (e.g. speech) data. The researchers used an Asterisk SIP server with four weak passwords to increase the probability (decoy) of toll fraud attacks.

Their tests were based on; user agent (RFC 3261) call initiations (local outgoing calls), calls to PSTN (toll calls), the register on the honeynet, and the attack behavior such as voice pattern (language, e.g. Arabic), calling frequency, direction of calls (originating/terminating), etc. Their data collection period was between August 2011 and 2012. The research findings were that when the PSTN gateway was open (connected) most of the call attempts were international or premium call connections/requests. Otherwise, local calls' attempts were scarce. It was also noted that one weak system account password is enough to route all VoIP traffic via this account (i.e. system fully hijacked).

When the PSTN link is disconnected, they found drastic shrinking of call attempts. After analyzing signaling data to determine countries of call origin, it was found the attempts are international connections. Also it was noted that most of the attackers used VaxSIPUser Agent, a tool for attacking multiple victims.

The researchers instructed on four layer countermeasures on the trapped attacks: harden user credentials, to beat brute force attack, validation of caller data, to be sure no spoofing, validation of signaling data (cross check origin and destination addresses), and context protection (check voice pattern, e.g. uncommon language in your system, means sent alert).

The study recommended a combination of improved password policies, blocking rules, deeper header inspection of the **signaling protocol** and voice pattern detection do prevent toll fraud. They recommended further research on toll fraud attacks, fraud detection and fraud blocking mechanisms is required to ensure secure VoIP systems.

### 2.5.2 User CDR Profiling Method 1

The study by Wiens, Wiens and Massoth (2014), was based on an autonomous unsupervised user profiling approach for fraud detection using CDR (detailed call records) as data for the analysis and extreme random fluctuations in the data were considered as suspect fraud. Two profiles were used for each customer to measure user behavior in different time spans, of which their comparison yields a change, which has threshold to indicate fraudulence.

The profiles were current behavior profile (CBP) and past behavior profile (PBP). They keep on being updated, and with continuous comparison of these features (CBP compared with PBP), estimation (extreme changes) of fraud or not fraud is made.

The attributes used for the two profiles are: number of maximum calls, maximum call duration, maximum call costs, mean calls, mean call duration, and standard deviation of calls. The more the CDR, the statistics have more accuracy and less fluctuation.

The study experiment had a target population of 76,326 calls in a month. The whole data set was used.

Their findings were: empty profiles compare with loaded profiles were giving false positives (this was solved by keeping the last profile), unexpected fluctuations were detected over weekends and holidays (this can be predicted and adjusted for), and poor detection for low usage users.

Their conclusion is that; the technique allows detection of toll fraud using unlabeled data, needs no maintenance, not complex, highly modifiable and low positive rate. Administrator decision is required on detection of toll fraud.

The study recommended further research on an autonomous limit adaption to be developed, and additional attribute, **call destination**, (Wiens et al, 2015) to be incorporated in the approach.

### 2.5.3 User CDR Profiling Method 2

This work, by Wiens, Kubler, Wiens and Massoth (2015), is an extension of Wiens et al (2014). They enhanced the technique by proposing two online user profiling approaches, **call destination profiling (CDP)**, and user behavior pattern recognition. The key point is more on behavior of the called directory number (DN). That is, number of calls to a certain destination (DN), e.g. number of calls that have been conducted to a given destination number destination (premium rate service) in a given time is suspicious.

A local telecommunication supplied CDR data of two weeks. The study experiment set week one data for machine training (no fraud detection) and week two set fraud detection. Two simple behavior patterns were defined, past behavior profile (PBP) and current behavior profile (CBP).Thresholds were estimated by analyzing the resulting values of CBP for fraud and non-fraud cases and for each call type.

The test steps are:

1. Detection deactivated first
2. Profiles are initialized using data set of the first week
3. Thresholds are calculated from CBP values as described before
4. Detection is now activated
5. Data set of the second week is now used as input.
6. Results from detection method are compared to the known cases of fraudulent behavior.

Their detection results: the method was able to identify all attacks, 100%, i.e. the known attacks, otherwise, 95% true positive.

Study findings: 17,110 fraud cases were reported and analyzed. One customer was making frequent international calls, but noted the number was a call centre and excluded (whitelist). A TPR of 98.4% and a FPR of <0.01% was recorded. Not all fraudulent instances were detected, hence study recommended further research.

Also, further research recommended on extensive analysis on CDR profiling, and analysis to be done on a large enough prepared and **labeled data set**.

### 2.5.4 Artificial neural networks method

The study, by Koiser (2016), used a VoIP fraud detection model based on artificial neural networks (NN) to classify VoIP calls as either fraudulent or legitimate based on attributes of the call. A multi-layer feed-forward NN with feedback propagation learning algorithm was used to build the model. CDR, a data set of 15,000 call records was collected from a working VoIP system. call destinations were classified, on-net (internal) and off-net (mobile, national and international).

The implementation of the ANN model was through Matlab NN toolbox. The trained network achieved 100% classification performance on the test data set.

A web application for reading and classifying records in new CDR was developed, which gave visualization (display) of the classified records.

The study established the ANN is a successful technology that can be applied in VoIP fraud detection. This is because it was able to detect abrupt changes in established calling patterns which may be consequence of fraud.

These attributes where used to make the customer profile: average call duration, number of answered calls, calls to/from different area code. Number of weekday calls (Mon-Fri), Number of daytime calls (9 am-5 pm), average number of calls received per day, average number of calls originated per day, number of unique area codes called during period P. the features recoded per VoIP system make huge amount of database. The study used Cross Industry Process Model (CRISP) for Data Mining. Also a web application was developed, using Agile system development methodology, for visualization (mapping) fraud and non-fraud calls of the VoIP system.

The study established that the ANN is a successful technology that can be applied in VoIP fraud, because the model was able to detect abrupt changes in established calling patterns which may be as a consequence of fraud. The developed web application read and classified call records in new CDR files, and used the trained Matlab NN to display results (normal and fraud calls).

The conclusion was that, the implementation of the fraud detection tool based on artificial intelligence will be a big step towards detection and mitigation of VoIP fraud.

The limitations of the technique include: the ANN ability for fraud detection will depend on accurate training which require high computer resources and takes a long time (especially cases of big data), normal calls occur more often than fraudulent calls which is challenge to obtain sufficient data for fraud calls, and involving in human labor.

The study achieved detection as the first step in mitigation of toll fraudit then recommends future work to focus on toll fraud prevention, considering the big money lost by VoIP fraud victims. It also recommends methods (models) combinations of which comparing results from the many models might increase confidence (accuracy).

## 2.6 Literature Review Summary and Research Gap

The toll fraud is a growing illegal business, and many tools for executing the crime are online. The literature done review showed there is need for more research on toll fraud. Toll fraud is a growing lucrative illegal business. Cyber criminals are developing day and night advanced attacking tools. The research on the fraud should be ahead of the criminals, otherwise businesses; especially SMEs will be financially crippled through the VoIP PBX.

Most of the previous work is on machine learning methods. The techniques are trained with labeled data, which require experts to acquire. Few of the techniques proposed are on labeled data. Quite a number of them require some sort of training, making automation hardly possible.

From the literature review, our work is based on the following recommendations on further research as tabulated below:

| Year | Author(s) | Method | Recommendations | Remarks |
|---|---|---|---|---|
| 2013 | Gruber et al | Honeynet | A deeper inspection of the **signaling** protocol | |
| 2014 | Wiens et al | User CDR profiling | Analysis on call **destination** profiling | |
| 2015 | Wiens et al | Destination CDR profiling | Extensive analysis on CDR profiling, and analysis to be done on a large enough prepared and **labeled data set**. | |
| 2016 | Koiser | Artificial neural networks | | Web application for mapping (**visualization**) of normal and fraudulent calls. |

Table 2.1: Further Research Works

The above works has driven us to research on call toll fraud detection and prevention in VoIP networks, mainly the IP PBX. The idea of using call setup delay time (CSDT) as a variable to be measured (timed) and manipulated as a distance vector is derived from the above further research recommendations. The bolded words in Table 2. Signaling, destination, and labeled data set relate to our study.

Signaling in telephony involves setting up a connection (link) between calling party and called party. Both dial tone in fixed line telephone system and network signal in mobile telephone system, are signals that commence telephone call processing. They imply the machine is ready to serve you. Obvious you dial digits as per the directory number (called party) you desire to communicate with. The telephone system once it receives defined digits it determines call destination through digit analysis process. Whether local, national, international or special call the telephone exchange will send an alert signal to the called. A ringing signal is send to the called telephone and a ringback send to the calling phone. When the called answers voice (media) communication starts. The process up to ringing/ringback alerts is the **signaling** phase.

A finite time is taken before the speech commences. That time is the **call setup delay time (CSDT)**. Our study is based on the time lapse between 1$^{st}$ digit dialed and the 1$^{st}$ ringback alert, called post selection delay (PSD) time (ITU, 1999). PSD time for ISDN calls are given in Table 2.1 as per ITU-T Recommendation E.721 (05/99)

| PSD time (s) | Call Type |
|---|---|
| 3 | Local |
| 5 | National |
| 8 | International |

Table 2.1: Call Setup Delay Time

The idea of call delay time approach was inspired by the recommendation on further research on call destination (Wiens et al, 2014). Both call source and call destination are part signaling information exchanged between signaling server(s) and the calling and called call devices. The information is such as source and destination IP addresses, called and calling DNs, etc. The profiling involves both source and destination attributes, and the additional variable, time delay.

The time variable measured, once link speed is given, the geographical location distance between a customer and signaling server can be calculated. Whether the server location is premises or cloud based, the geographical location distance remain fixed. Consider the premises based KCA University VoIP PBX in the Kenya map and World Map attached in Appendix II.

The figures show the distances covered when from a source, e.g. a call from KCA university PBX. Assume a caller using extension number 2111. The table below shows example different destinations. A call PSD time within the institution will take the shortest time. But a call to USA will take appreciable time.

Consider these four destination profiles in Table 2.3

| **CallSource** | **CalledNo.** | **GeoLocation** | **CallType** | **Appriximate PSD time (s)** |
|---|---|---|---|---|
| KCA ext. 2111 | 2000 | KCA | Internal | <3 |
| ,, | 2048 | KCA | Internal | <3 |
| ,, | 22111 | Nairobi | Local | 3 |
| ,, | 045-20111 | Athi-River | Local | 3 |
| ,, | 02-331111 | Nairobi | National | 5 |
| ,, | 044-22111 | Machakos | National | 5 |
| ,, | 041-21111 | Mombasa | National | 5 |
| ,, | 057-20011 | Kisumu | National | 5 |
| ,, | 0722-111111 | Kenya (Safaricom) | National (mobile) | 5 |
| ,, | 0731-111111 | Kenya (Airtel) | National (mobile) | 5 |
| ,, | 0770-211111 | Kenya (Telkom) | National (mobile) | 5 |
| ,, | 001-11-221111 | USA | International | 8 |
| ,, | 00-45-33111 | Denmark | International | 8 |
| ,, | 00-44-2211111 | UK | International | 8 |
| ,, | 010-81-331111 | Japan | International | 8 |
| ,, | 0011-61-21111 | Australia | international | 8 |
| ,, | 00-224-330111 | Guinea | International | 8 |
| ,, | 09-27-2011111 | S.Africa | International | 8 |
| ,, | 00-20-2211111 | Egypt | International | 8 |
| ,, | 00-252-220211 | Somalia | International | 5 |
| ,, | 999 | Kenya | Special | <3 |
| ,, | 100 | Kenya | special | <3 |

Table 2.3: Various call destinations delay times

Media velocity can be determined. The media types are copper wire (twisted pair/coaxial cable), optic fiber or radio frequency. The speed of light (c) in vacuum is 3.86 x $10^8$ m/s. A specific medium propagation speed is less than that of light in vacuum. Assuming a velocity factor, n, the propagation speed of data in a medium is **n** x **c** m/s, equal to $\mathbf{c_m}$. Once n and c are known, the distance (d), given by distance (d) = speed ($c_m$**)** x time ($t_D$), or

$$d = c_m \times t_D \text{ meters}$$

this gives a fixed distance between the two subscribers. Or assuming the central point the PBX, the radial distances of all destinations from the exchange can be measured and a profile of distances created for a particular site. The profile will make a labeled data as recommended by Wiens et al (2015). This data can be used to for comparing with measured distance for discriminating spoofed calls from legal calls.

A logical alert mechanism for visualization of fraud and fraud calls to be implemented. This one is simpler and cheaper compared with a web application based done by Koiser (2016).From the literature read, the works of Gruber et al (2013), Wiens et al (2014), Wiens et al (2015) and Koiser (2016) an idea of solving the problem of toll fraud using call setup time had been created. The time taken to connect calling and called parties by a telephone exchange is finite. That time when computed with the link speed results to the physical distance between the two objects. That distance cannot be mimicked (spoofed). Thus, calling subscriber distance and called subscriber distances can be calculated using this concept. This is a research gap, and adds knowledge in this field. We propose to use the concept and tandem it with CDR to fight call fraud.

Deducing from the theory and literature review the model above, figure 2.3 show our conceptual framework. The conceptual framework presents basically three modules, input module, processing module and output module.

## 2.7 Conceptual Toll Fraud Model

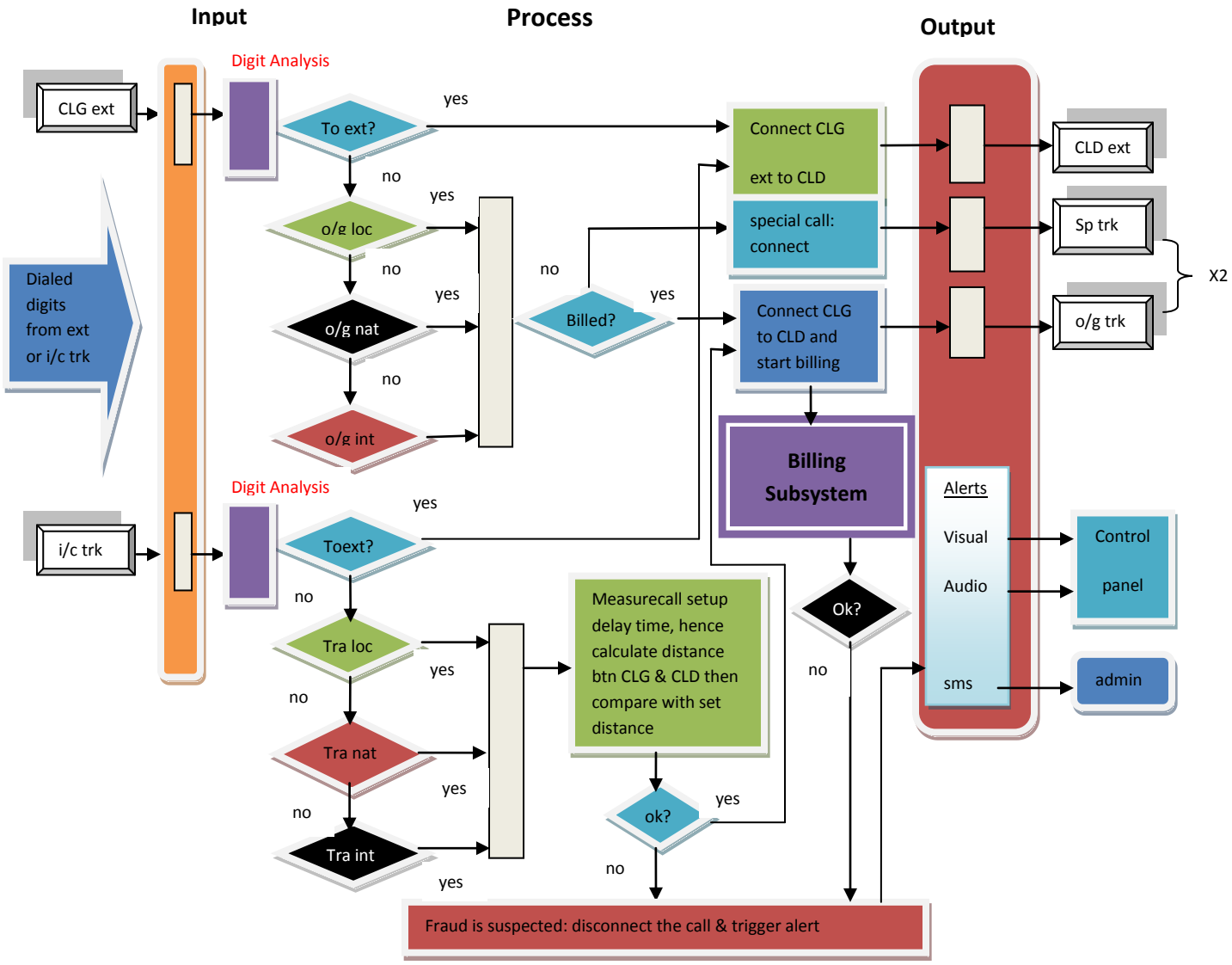



Figure 2.3: Conceptual Framework Model

### 2.7.1 Input module

This module presents the inputs as independent variables. These can be classified as normal calls and malicious (fraudulent) calls. Calls in a telephone exchange, whether legal or illegal, are categorized as local calls, national calls and international calls depending on radial distance from the telephone system. Another classification is in free calls or charged (toll) calls. Various reasons can cause toll fraud calls. The agents of the toll fraud calls are either internal (employees) or external users.

### 2.7.2 Processing module

At this stage the input variables are manipulated by the model. The model analyses various calls with objective of detecting a deviation from normal characteristics. Call origin and destination attributes are used to determine the distances between the calling, called parties and the telephone exchange. If calling/called party destination is spoofed, a rule will be satisfied and the module will trigger an alert.

A billing subsystem accounts for all chargeable calls by calculating call duration and converting the time to charge units. These units determine the individual customer call bill, and customers pay as prepared or postpaid. Our model is to monitor any unusual billing. That is a toll call is made but not billed or unusual bulky billing of a particular number during a weekend/holiday, etc. These behaviors will be monitored and detected. The triggers will be processed and disconnection of the fraudulent call and alerts send to the administration system.

### 2.7.3 Output module

Our output module presents the dependable variables. When a toll fraud call is detected and confirmed, it is disconnected. The module also sends alert to supervision panel and an SMS to the administrator.

## 2.8 The VoIP PBX Model Operation

The model implements two methods for detecting and preventing toll fraud. The first technique is to evaluate the distance between a connected subscriber and the exchange, and the second technique is monitoring billing parameters in call accounting subsystem.

In case one, both distances between calling (CLG) party and the PBX, and between called (CLD) party and PBX are calculated. First the originating (calling party) distance is determined. The distance is an addition to the usual calling customer identification attributes such directory number, IP address, username, password, etc. The advantage of this attribute is that it is fixed and cannot be spoofed. The called party distance is also evaluated through use of dialed digits. The digits will determine the call type, hence destination of the call. If an internal (ext-ext) the local signaling server will determine the called extension location (DN and IP address). The distance between the extension and the PBX will be evaluated. For local, national, international or special calls the same procedure is used. That is, the first few dialed digits will determine call type, and the PBX will select an outgoing trunk to the specific destination. In this case the called party is outside the enterprise. In all the cases the distance is calculated using call setup delay time. Note that signaling server (SVR) can be premise based or cloud hosted (cloud computing), but still the call setup delay time will still remain fixed.

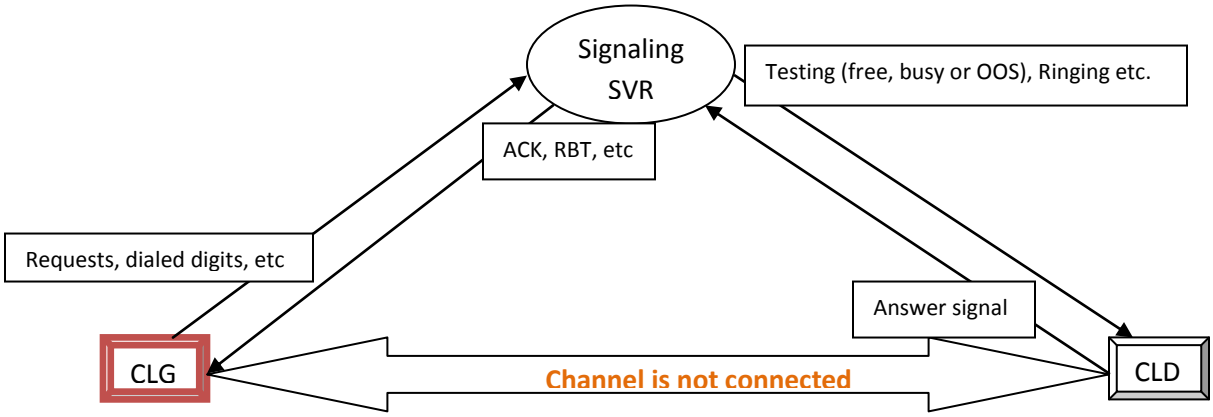



Figure 2.4: Signaling Phase

The second method is based on call charging done by a billing subsystem of the PBX or a main telephone exchange. The signaling server tests status of the called party and if free: sends ring back to the calling subscriber, send ringing to the called party and select a free network speech path to connect the two on called party answer. On hearing ringing called subscriber answer, and the voice device send answer signal to the exchange. Once the server receive the answer signal, it will: disconnect ringing and ring back signals, change-over signaling path to speech path, send initial metering pulse to call accounting system, commence call timing, and free itself to serve others in need of signaling.


Figure 2.5: Speech Phase

The connection time depends on the two parties. Normally the calling party is the one who is charged, so more is the one to control call duration. Subsequent charge pulses depend on call duration and charge rate (band).

### 2.5.1 General Number Dialing Plan

A call destination is determined by the directory number (DN) dialed by the caller. The first few digits are used to route the call the location of the called party. The table below, Table 2.1, show the general structure of the possible digits and the call types they translate to.

Table 2.3: Call Digit Analysis

| s/n | DN Length | Call Type |
|---|---|---|
| 1 | 1xxx….. | Local calls; numbers of a telephone exchange/PBX, e.g. DN 220011 or Extension 2081 |
| 2 | 2xx…… | |
| 3 | 3xx…… | |
| 4 | 4xx…… | |
| 5 | 5xx…… | |
| 6 | 6xx…… | |
| 7 | 7xx…… | |
| 8 | 8xx…… | |
| 9 | 9xx | Special calls, e.g. 999 (police, emergence, etc.) |
| 10 | 0xx…… | National calls: Fixed or Mobile |
| 11 | 00…… | International calls |
| 12 | *xx….. | Supplementary services, e.g. *144# |
| 13 | #xx….. | |

Some details follow on the two toll fraud detection and prevention approaches.

### 2.8.2 Call Setup Delay Time

Time lapse between first digit and last digit dialed and instant when ring-back tone is received (Eyers et al., 2000). To account for this, three aspects are considered; packet transmission time, propagation delay and packet delivery time.

#### 2.8.2.1 Packet transmission time

- Is time taken to sent first bit to last bit of a message by a sender
- i.e. packet transmission time, $t_t$ = packet size (p)/bit rate (speed=$c_{t)}$
- e.g. assuming 100 Mb/s Ethernet, and packet size of 1526 bytes
- Transmission time, $t_t = p/c_t = 1526/100 \times 10^6 = 122.08$ µS.

#### 2.8.2.2 Propagation delay

This is the time taken by a bit to propagate from a sender to the receiver. Its travelling rate depends on the medium used (optic fiber, copper TP, RF, etc.).

The propagation time, $t_p$ = distance (d)/propagation speed ($c_p$)

Example is the Ethernet communication on copper UTP cable 100 meters between a workstation and a LAN switch,

$t_p = d/c_p$

$= 100/200x10^6 = 0.5\ \mu S.$

### 2.8.2.3 Packet delivery time

The latency time starting when the first bit leaves the sender until the receiver records the last bit. When a physical medium link, packet delivery time, $t_p$ = transmit time ($t_t$) + propagation delay ($t_p$).

In case of network connections through different network components/devices, the delivery time depends on sum of each link, queuing time and processing delay. The WANs delivery time is in order of milliseconds (mS).

The ITU Recommendation E.721 (1999) specifies the average call setup delays. These are, 3.0 mS for local calls, 5.0 mS for national calls, and 8.0 mS for international calls. The parameters can be profiled in the CDR and be a factor to detect toll call fraud.

On a specific site where an IP PBX is installed, different tests can be done to create a training profile (lookup table) for the site. The possible call connections are:

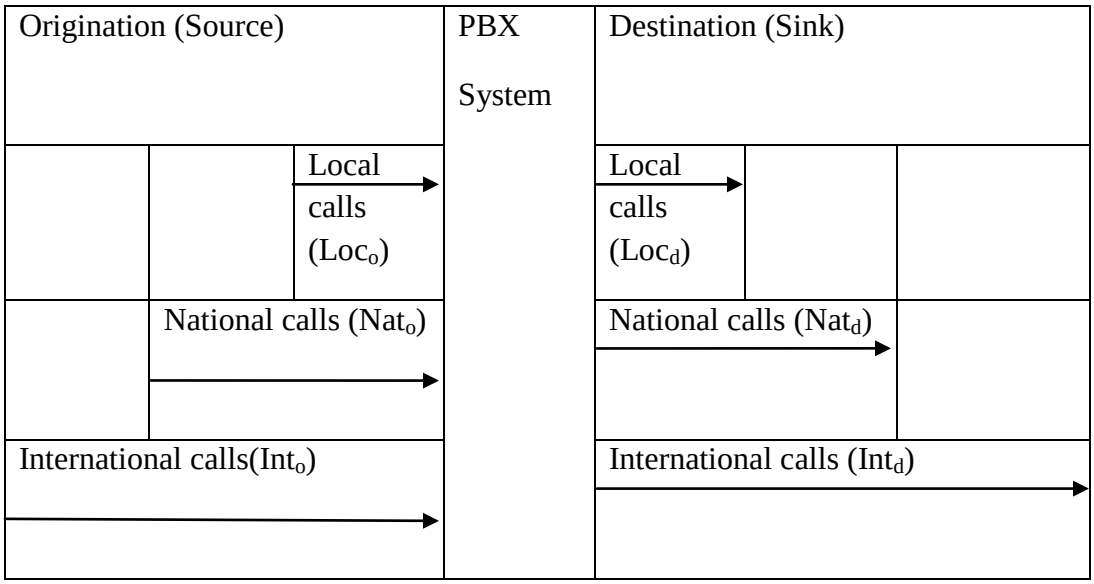



Table of call connection type (Conn_Type) vs average call setup delay (Ave_Delay) can be create such as below:

| S/N | Conn_Type | Ave_Delay (mS) | | Action |
|---|---|---|---|---|
| 1 | $Loc_o$ – $Loc_d$ | | | Normal |
| 2 | $Loc_o$ –$Nat_d$ | | | Normal |

| | | | | |
|---|---|---|---|---|
| 3 | $Loc_o$ – $Int_d$ | | | Normal |
| 4 | $Nat_o$ – $Nat_d$ | | | Abn |
| 5 | $Nat_o$- $Int_d$ | | | Abn |
| 6 | $Int_o$ – $Int_d$ | | | Abn |
| 7 | Special | | | Normal |

Table 2.1 Call type vs average delay

The delay time can be used to calculate average distance between the calling and the exchange as well as the called party distance. The comparison of the distances can validate a fraud or no fraud call connection.

The most common VoIP signaling protocols are SIP and H.323 (Malliah et al., 2011). These determine the IP call setup delay.

## 2.9 Call Billing Fraud Detection

The normal calls are monitored and charged by call accounting subsystem. The charging system may be implemented in the PBX or in the PSTN local exchange. The billing depends on these parameters: call type, call duration, day type, time of day and charge rate (band). For chargeable (toll) calls, the metering starts on calling subscriber answer. The answer increments subscriber's meter with one unit and the subsequent units depends on call duration and charge rate.

Toll fraud is common at night, weekend and holidays when most enterprises close shop. The study is focusing on monitoring calling behaviors of clients over these times/days. Too long call duration means high bill a deviation from normal billing. The call accounting subsystem will detect the anomaly and report (alert) by raising alarm on control panel and sending SMS to administrator.

# CHAPTER THREE: RESEARCH METHODOLOGY AND IMPLEMENTATION

## 3.1 Introduction

This chapter discusses on how we planned to tackle the research problem (toll fraud). It provides the work plan and describes various activities necessary for the completion of the research project. The tasks include research design, target population, sampling strategies, data collection instruments, data collection procedures, data processing and data analysis in relation to the research objectives.

Research methodology is a step-by-step process of solving a problem. There are many types of research and are classified on basis of methodology and the knowledge it creates, application areas, problem to be solved, etc., (Catherine, 2002).

The most common methodologies of research are: basic research, applied research, problem oriented research, problem solving research, quantitative research, qualitative research, simulation research, etc.

## 3.2 Methodology for achieving objective one

### 3.2.1 Research approach:

The goal of this objective was to find out the existing methods of fighting toll fraud in VoIP infrastructure. A literature review was done using ***literature based research methodology.*** A ***systematic literature review method*** was designed where the target population was all sources where call toll fraud literature and related data was found. The source materials are such as latest publications, journals, text books, face to face discussions, etc. A number of data bases were accessed to download most of the data collected. These data bases are such as Google Scholar, Google Search Engine, Springer, IEEE Explorer, etc.

The search criterions were limited to the relevant literature review of before year 2013, while the latest literature review (state of art) of up to year 2017. The two categories form two research groups of the target population. The literature before year 2013 gave most of the research study's background information. The state of art (SOA) review identified four major works (methods), given on Table 2.1 of Chapter Two. This made our sample size to of 4 (four). Guber et al, (2013) used Honeynet method to detect toll fraud and recommended further study on **signaling** protocol. Wiens et al, (2014) approach on toll fraud in VoIP networks was on user profiling and the study recommended further work on call **destination**. The Wiens et al, (2015) studied the problem using technique of destination profiling and recommended use of **labeled data**. Final sample is that of Koiser (2016) who used a method of artificial neural networks to detect toll fraud in VoIP infrastructure. Koiser demonstrated way of classifying calls as genuine or fraudulent.

Detailed analysis and criticism resulted to identification of a research gap in fighting call toll fraud in VoIP environment. That is to say, we synthesized signaling, destination and labeled data information. Our method CSDT (call setup delay time) is a technique we have used to measure **signaling time**. This time is utilized to determine the fixed distance between the called

(**destination**) and the calling parties. The time (distances) we measured for local, national and international calls make the labeled data (thresholds) that is used in the CSDT method to compare with actual time measured. So signaling time, origin-destination distances formed our basis for our method for detecting and preventing toll fraud in VoIP PBX infrastructure.

### 3.2.2 Target population:

An extensive literature survey was done on important articles, books, and other sources with literature related to toll fraud in VoIP networks.

### 3.2.3 Sampling procedure:

Purposive sampling method was used. This is a non- probabilistic sampling method that selects a number of individuals for a study to represent a large group from which they were selected.

The search was grouped into two units: one on the relevant literature review of before year 2013 and two on the latest literature review (state of art) of up to year 2017. The two categories form two research groups of the target population. The literature before year 2013 gave most of the research study's background information. The state of art (SOA) review identified four major works (methods), given on Table 2.1 of Chapter Two. This made our sample size to of 4 (four). Guber et al, (2013) used Honeynet method to detect toll fraud and recommended further study on **signaling** protocol. Wiens et al, (2014) approach on toll fraud in VoIP networks was on user profiling and the study recommended further work on call **destination**.

### 3.2.4 Sample size

Many publications were reviewed but we narrowed to sample size of 4 (four) articles. Guber et al, (2013) used Honeynet method to detect toll fraud and recommended further study on **signaling** protocol. Wiens et al, (2014) approach on toll fraud in VoIP networks was on user profiling and the study recommended further work on call **destination**.

## 3.3 Methodology for achieving objective two

Simulation method was used to design and implement the toll fraud detection and prevention model. A simulation model can manipulate variables and test to see immediately how results change.

### 3.3.1 Toll Fraud Detection and Protection Model Design

The model is based on a typical IP PBX infrastructure. The model outlines the interconnections, interfaces, relationships and characteristics of its various network elements.

### 3.3.2 Installation

This phase involves installation of hardware, loading the PCs with relevant software and initiating the defined call connections. The LAN was installed physically as shown on Figures 4.1a, b and c. The PCs are loaded with software as follows:

- PC1, Asterisk IP PBX software (has telephony applications)
- PC2, Wireshack, packet analyzer
- PC3, X-Lite, soft phone
- PC4, X-Lite, soft phone.

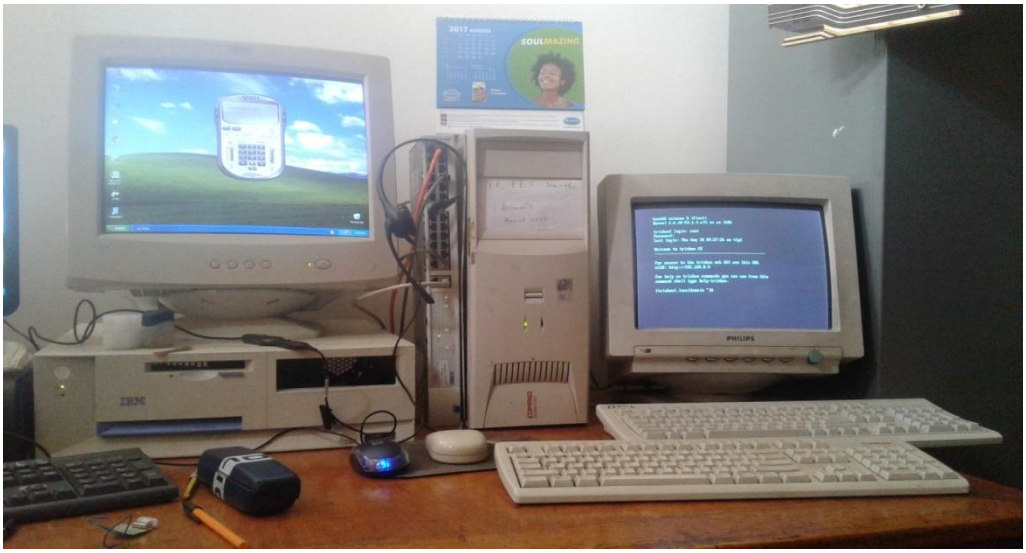



Figure 3.1a: ext. 202 and IP PBX server

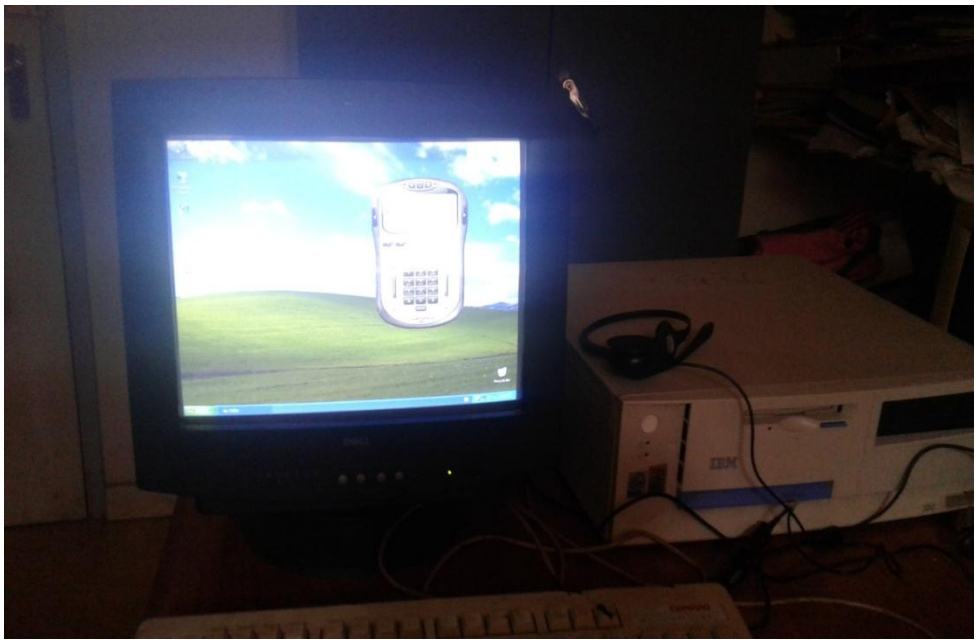


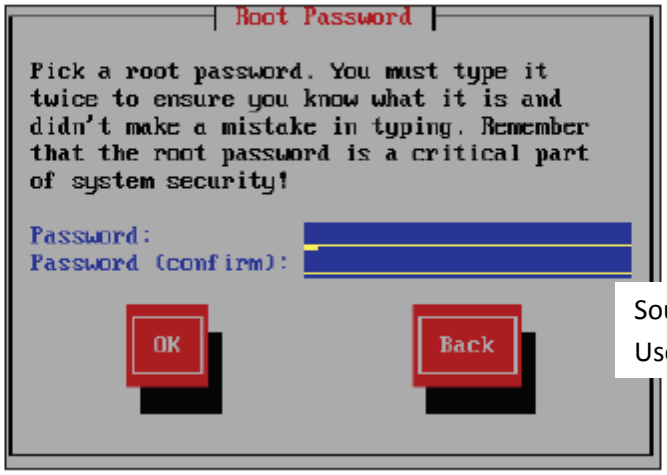

Figure 3.1b: ext. 201

### 3.3.3 Initial Installation/Configuration Procedures

The installation and configuration of the small scale (model VoIP) PBX steps are as follows:

TrixBox CE (VoIP PBX software) is loaded in a desktop computer,PC1, which is the network signaling server. During automatic model installation a very important message given below will appear.

Next, we are prompted for a password for the root user. It is advisable not to forget this password as we will need it to log into our system when it is up and running.




Source: TrixBox Made Easy User Guide (2006)

Figure 3.2: Creating root password

The administrator password was entered, and confirmed.

You will be prompted to configure the network adaptor IP address as given:


Figure 3.3: IP address configuration

The study model had been configured with a static IP address: 192.168.0.5 for its network accessing.

After completion of loading the telephony platform, the system will prompt the installer with command line interface (CLI). This is a user console which gives a full control of the model software (Linux OS and the telephony applications). The CLI screen below prompts you to enter user name, and then your password.

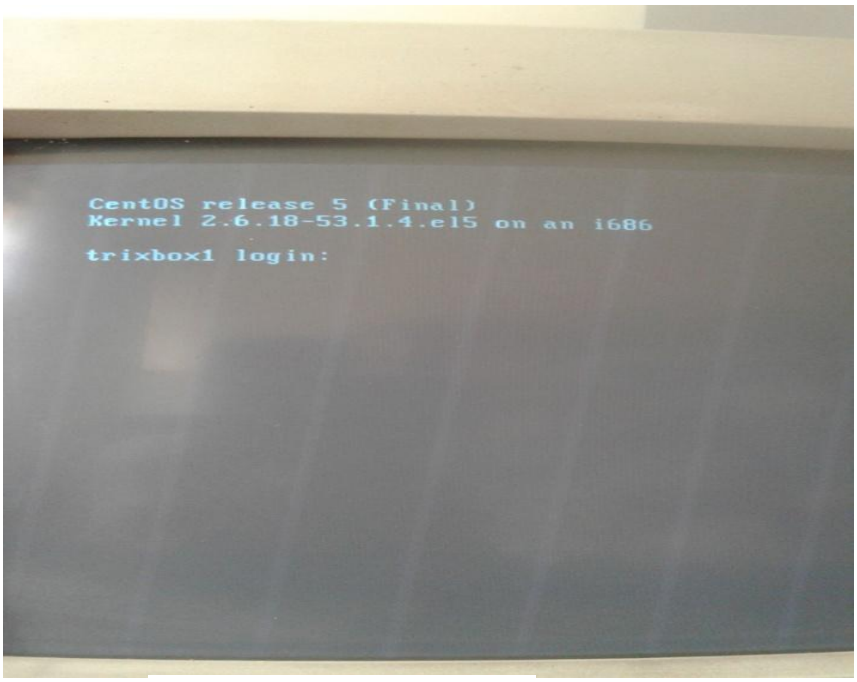



Figure 3.4: Login as root user

Enter "root". The model will prompt you "password". Enter the confidential password you created earlier in the system. Below is the snap shot you will get.

The screen displays,

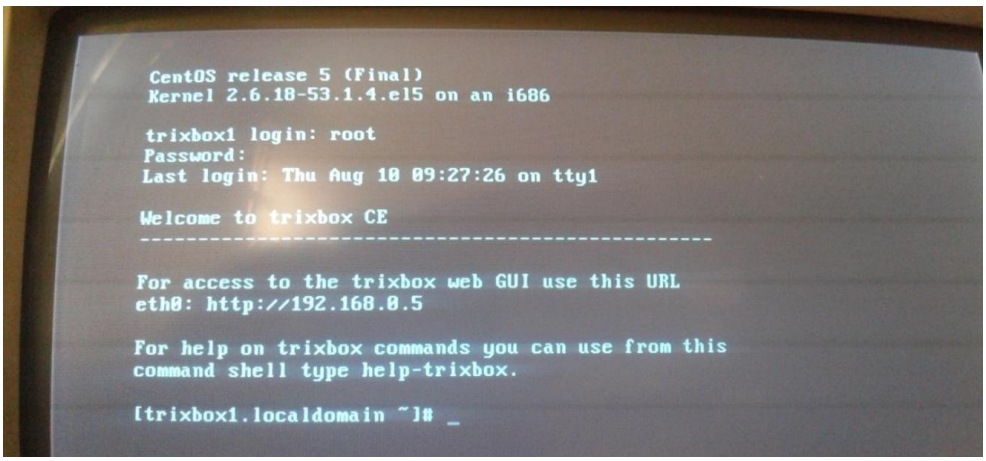



Figure 3.5: Successful login to CLI console

At this stage the IP PBX is running well. Installation and configuration of various user devices, such phones and PSTN trunks, can commence. To do this, CLI will not be the best user model interaction language. The CLI normally is not user friendly, unless you are CLI expert. The graphic user interface (GUI), rescues the user CLI limitation. As shown on the screen, the URL (uniform resource locator), http://192.168.0.5 open web browser of a remote terminal to communicate with the signaling server, TrixBox. The desktop computer, PC1 (or any other) is configured to access the server. Type in the address bar of the PC browser, (Chrome/IE/Mozilla, etc.) the address: 192.168.0.5. The PBX main menu (window) appear as,

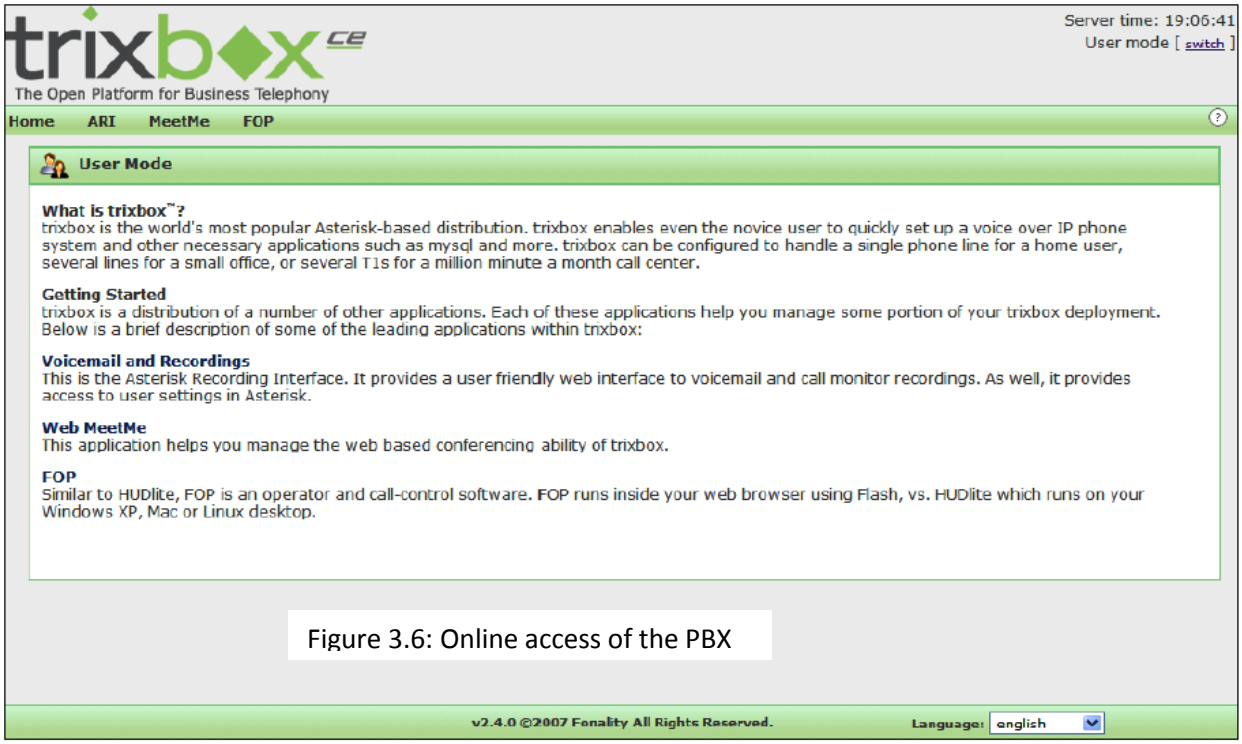



Figure 3.6: Online access of the PBX

Click on the **switch** button, in quotes, on the top on the right hand side, to login in the system. A dialogue box will appear. Figure 4.7 is an example of a login dialogue box.


Figure 3.7: Maintenance login

Enter

Username**: maint**

Password**: password**

Both the username and password are default, and for security to change to power credentials.

### 3.3.4 Creating extensions

After logging in the PBX server, click PBX>>PBX settings>> click on extensions on the left hand side.

"Generic sip device" is showing in front of "Device" in the box.

Click on "submit". After submission, a configuration form appears.

Type User Extension: 201

Type Display Name: Eliza

Type SIP Num Alias: 201

Type secret: 1234 (extension user password)

Click submit>>then click on "Apply red bar configuration changes", shown at the top of the page. This updates the system database on the new changes.

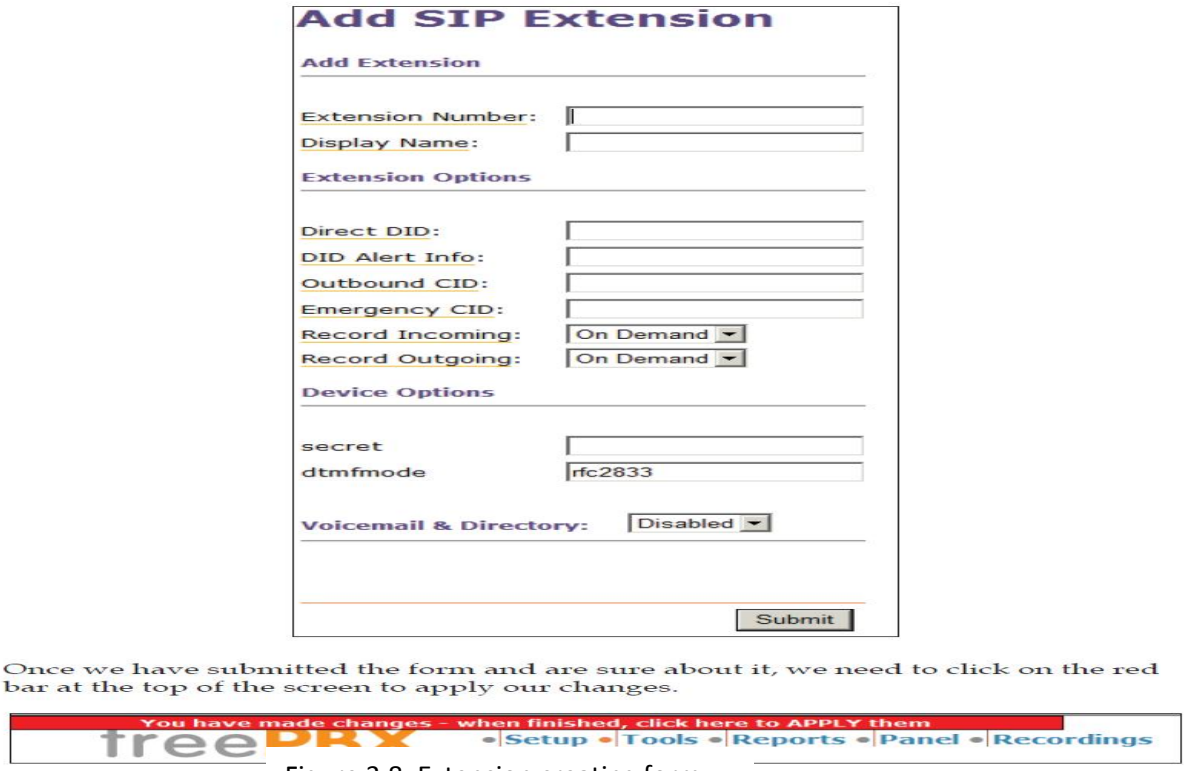



Figure 3.8: Extension creating form

The system will respond by showing the extension configured, “Eliza<201> has been successfully crated.

For our study, at least two extensions are required. The second created extension is “Muli<202> as shown on Figure 4.9

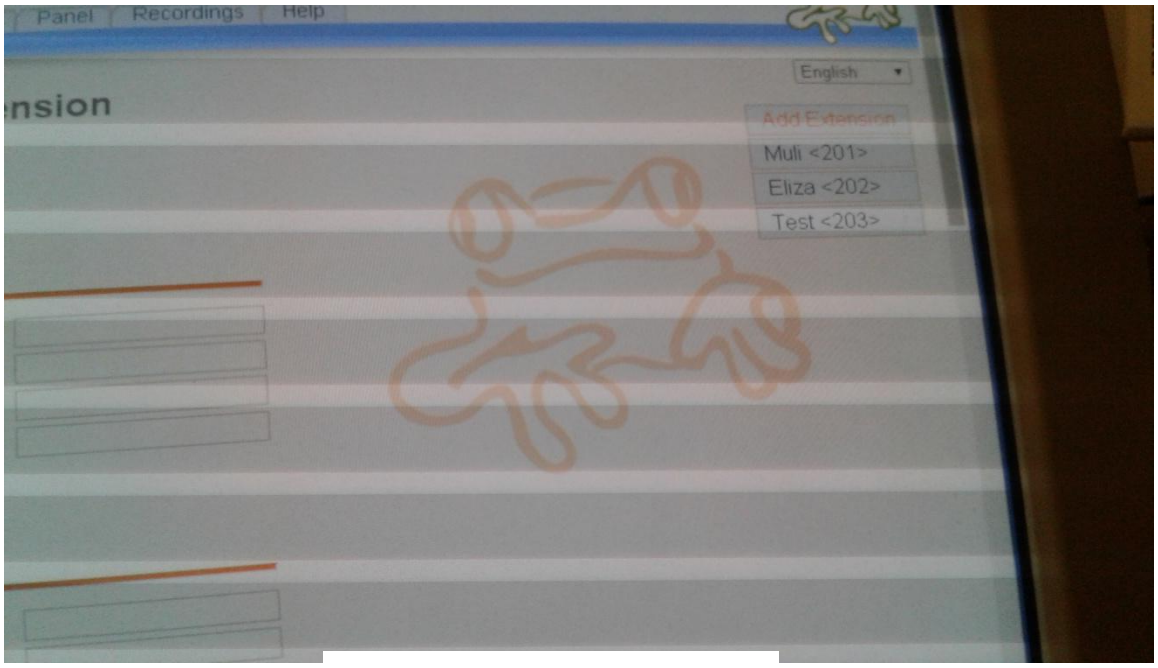



Figure 3.9: Created ext. 201 &202

### 3.3.5 Soft phone installation

Installation and configuration of soft phones is the next step. The examples of soft phones are such as X-lite, 3CX and zoper. X-lite soft phone free application was downloaded and installed.

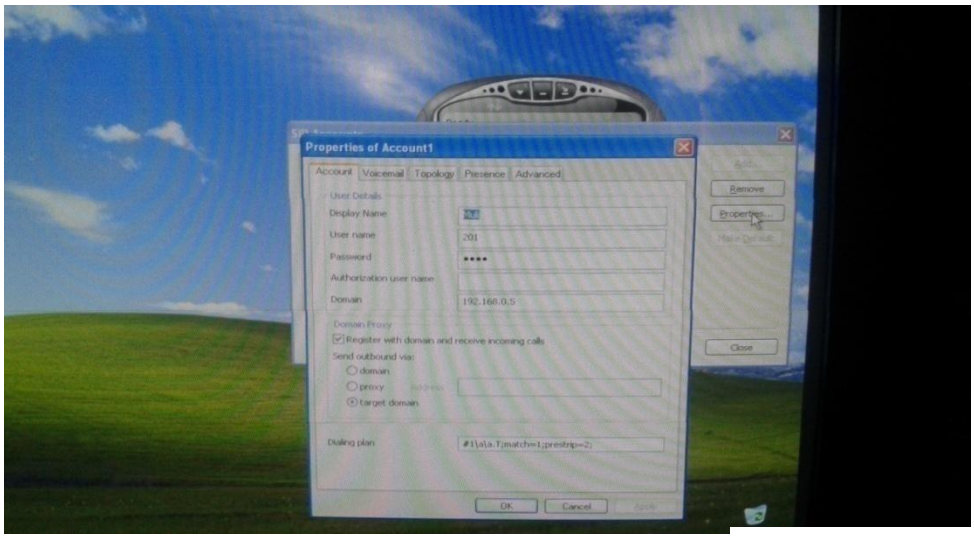



Figure 3.10: X-Lite soft phone

Click on soft phone option>> account settings, then

Account name: Eliza>> User ID: 201: Domain: 192.168.0.5

Password: 1234>>display name: 201

Click OK.

SIP account, extension: 201 had been enabled in X-Lite software.

Configuring the second extension:

Account name: Muli>>User ID: 202>>Domain: 192. 168.0.2

Password:1234>>display name: 202

Click OK.

SIP account, extension: 202 had been enabled in X-Lite software.

### 3.3.6 A test call between the extensions

A USB headset is plugged in to each desktop computer, PC2 and PC4. After initialization a ready display appears, Figure 4.11


Figure 3.11: Ready X-Lite soft phones

A test code *65 is used for phone self-test, line and audio sensitivity/quality tests are done. Extension tests 201 to 202 and vice versa tests are done. Figures 4.12 and 4.13 shows two users communicating through extensions 201 and 202.

Figure 3.12: User ext. 202

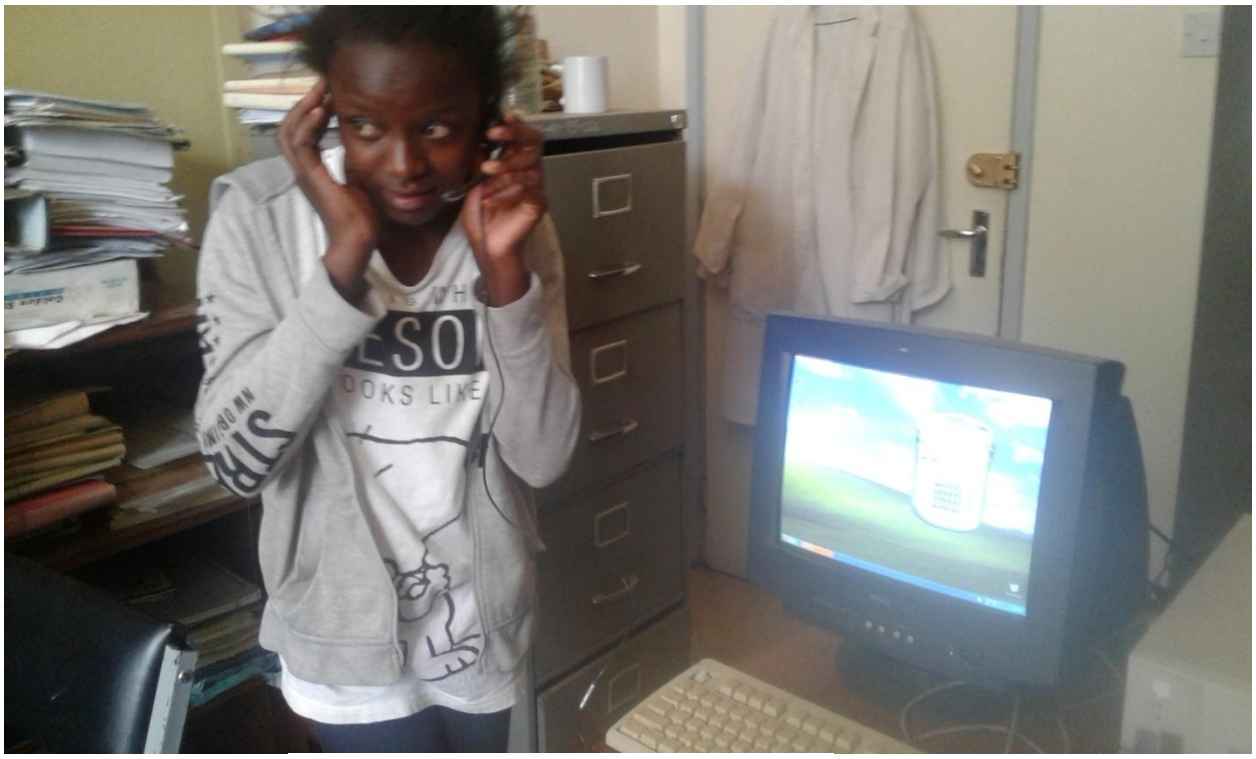


Figure 3.13: User ext. 201 communicating

### 3.3.7 Creating trunks

These are interface circuits linking PSTN or another PBX with the VoIP PBX model. The form to create the trunks is shown below.

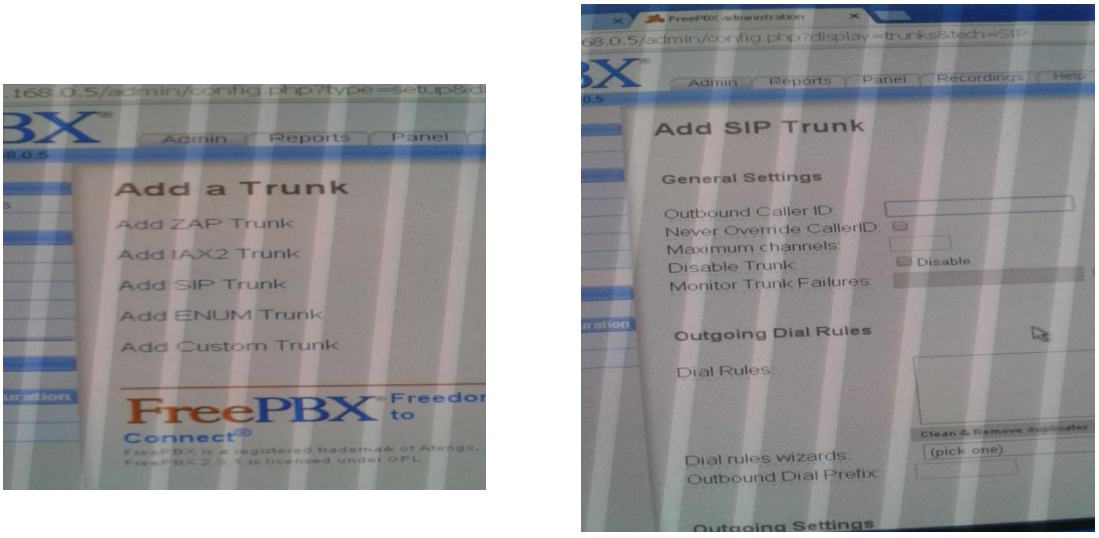



Figure 3.14: Trunk creation

Due to administration and technical issues, linking to the PSTN was not viable. Permission was requested to acquire real CDR data from a particular VoIP PBX of an institution. The institution

is within Nairobi city and it is the center of communication in our study. Call types had been analyzed as how far a subscriber is from the PBX.

## 3.4 Methodology for achieving objective three

### 3.4.1 Research design

A project design ensures the entire planning for the success in achieving the goals of a study. We chose experimental research design to test toll call fraud. Experimental research has two major subclasses, true experimental and quasi-experimental. Quasi-experiments take advantages of natural occurrences. True experiments are popular where quantities (variables) are measurable. We chose true experimental because we had to measure time quantity.

The choice of research design determines the success of a study project, which also affects conclusions one makes from research results (Bordens and Abbott, 2011). A VoIP PBX network model design was used in the experimentation. The experimental method was chosen because of its two major features. The first feature is that it is possible to manipulate one or more variables, and second, it is possible to control over extraneous variables (keep variables constant and measure only one at a time). It made it possible to measure quantities such as telephone call traffic, billing units, call duration, call delay time etc. The values bear relationships that can be expressed mathematically as variables, i.e. dependent variables and independent variables. The variables statistically can be manipulated to project the next behavior/value.

### 3.4.2 The Principles of Experimental Design

A successful research study relies on a proper design and a clear well layout of experimental design. In an experiment a researcher manipulates one or more variables, while holding all other variables constant. The experiment is based on manipulated variables (independent variables), and the response (output) of the manipulation is the depended variable. Our research independent variables were calls to/from the PBX. The manipulation of the variables was the call processing by comparing a call setup delay time with a threshold set time (treatment). The output of the comparison is either a normal call or a fraudulent call. The output is our dependent variable.

The researcher can test whether a causal relationship exists between the manipulated variables and the response. Three principles of experimental design (Kothari, 2004) are followed to achieve the tests. These are:

a. Replication: repetitions to improve precision hence reduce standard error of measured quantity. Example is in our toll fraud detection model, repeated calls (12 samples) had been made to ascertain the distance between a subscriber and the signaling server.
b. Randomization: this is the practice of using chancing methods to assign experimental samples. It groups experimental units into groups such that every treatment has equal

chance of being assigned to any experimental sample. A good example in our research Toll Fraud Detection model were the inbound (terminating) calls are categorized as local calls, national (fixed and mobile) calls and international calls. It draws lots of incoming calls to the model, i.e. the function provides a random sample for toll fraud observation (detection) for any incoming calls to the model.

c. Local control: this is an approach taken to reduce effects of extraneous variables (i.e. variables other than the independent variables and dependent variable). The extraneous variables, called lurking variables can cause deviation of expected experimental results. The control minimizes experimental errors, receives no treatment and is designed to measure research bias. The techniques applied are such as:
    - Use of an appropriate experimental design
    - Compare several treatments
    - Use appropriate experimental sample sizes
    - Proper handling/testing of the selected samples
    - Refinement of sample tests.

    Our study used set thresholds values (treatments) in form of delay time, call duration and call cost to compare with actual value detected to determine a call class (normal/fraud call).

### 3.4.3 VoIP PBX design requirements

The designed and implemented VoIP PBX infrastructure is robust and reliable, hence produce valid data. Thus the hardware and software used is highly reliable.

These are the materials/equipment used for the IP PBX model design:

- Personal computers
- Operating systems (Windows and Linux)
- Trixbox software (IP PBX)
- IP softphones
- LAN switch
- UTP patch cables (RJ45)
- Power supply
- Two headsets for PCs.

Mean distance, from measured call setup delay time, to be calculated. Different call types were initiated to determine destination in relation billing parameters. A site testing was done considering radial mean distance for calls.

### 3.4.4 Location of the study

The diagnostics were done in laboratory and other data collected from live environment (working IP PBX). Experiments are done in laboratories where variables can be controlled to avoid inconsistent results. Some experiments can be done in field equipment, but with care to avoid interrupting services (Bordens and Abbott, 2011).

### 3.4.5 Target Population

Our IP PBX model capacity was composed of a signaling server and two softphones (two extensions). Our target population was all calls to or from the IP PBX model but sampled on 12 calls per call type.

A research population is a well defined collection of individuals/objects taken from a general population who share common characteristics, such call destination/origination, age, sex, etc. In some cases it is usually not possible to study an entire population, hence need to study a sample of the population.

In a telephone exchange, the total call traffic is due to incoming calls to and outgoing calls from the system. In a PBX calls are received or send by extensions and trunks. The telephony system traffic load depends on calling rate. The call type is determined by the customer's request of service. The services are classified as: internal call (extension to extension), which is a free call and charged (toll) calls. The toll calls can be sub-classified as local, national and international calls. There are also special calls such as the customer care, emergence numbers, etc. This category of calls most of them are not charged. All these calls aggregate to the target population.

### 3.4.6 Sampling and Sampling Procedure

A sample is a subset of population, selected by either probability or non-probabilistic methods. Or simply, participants in a research study. Sampling is a statistical inference that permits us to draw conclusions about a population. Sampling reduces research cost; it is cheaper to interview 100 people than 100 million people. Our study used ***simple random sample*** procedure whereby 12 test calls are made per call category. In simple random sample method each member has equal chance of being chosen. The test samples we used were these call connections: extension to extension, extension to local DN, extension to national DN and extension to international DN. For each connection 12 calls were sufficient for our data analysis.

### 3.4.7 Research Instrument

This is a research tool for collecting data. It should be easy to use. A common example is a questionnaire, with simple structured simple questions (Kumar, 2005).The IP PBX our tool that we used to generate and capture data for the study. Also from a working VoIP PBX data was collected for toll fraud analysis. Tables were created as tools for storing and retrieving data for analysis.

### 3.4.8 Validity and reliability of the IP PBX model

Validity is to do with the accuracy of the data collected. The error margin should be very low, hence accuracy of about 100%. Reliability points on data consistency. The series of test calls executed by the instrument must have uniform characteristics. The PBX model is on the standard computer hardware and software platforms; of which the accuracy and reliability is assured. May calls were initiated then compared with standard calls and a good consistency was recorded.

### 3.4.9 Data collection procedure

A series of test calls were initiated to and from the IP PBX model. This involved limited calls (12 calls) per call type. The timing of each call setup delay was recorded. The model is to compare a collected data with pre-loaded fixed data to distinguish fraud calls from legal calls.

In-built CDR is normally in telephony software. One of its important tasks is to collect telephony traffic. Various attributes are to be activated; these are such as call originations, terminations, frequency, duration, etc. The profiling was used in the second method to discriminate fraud calls from legal calls.

### 3.4.10 Data processing and analysis

The IP PBX model generated call traffic which was used for analysis so that it can identify fraud. If a comparison is equal, that call is deemed normal and call progress is allowed. Else a comparison is greater than the expected value that call request is suspicious and a further analysis is called depending on the rule to be applied. The rules are formulated as per call type, origination distance, destination distance, time of the day and day type.

Data collected was analyzed using descriptive statistical methods, especially the measure of central tendency (mean). Data was organized in frequency distribution tables. Graphical representation of data is shown in tables, charts and figures, etc.

### 3.4.11 Ethical considerations

The research was bound by a number of ethical considerations in respect to confidentiality of information such as:

- Privacy of secondary data (CDR) collected from the field. Running IP PBX
- Real user names from the field IP PBX were deleted from the CDRs.
- Unbiased interpretation of the data was made.

### 3.4.12 Data Collection and Experiment Settings

In our study two methods were used for detecting toll fraud in VoIP PBX environment. This applies the hypothesis that as the distance between caller and called users increases, also the call setup delay time (CSDT) increases. That is why we have attached national (Kenya) and world

maps in the appendix so that one can have a realistic visualization. In Method 1, we set three experiments and in method 2, two techniques where employed.

# CHAPTER FOUR: DATA ANALYSIS, FINDINGS AND DISCUSSION

## 4.1 Introduction

This chapter deals is about research findings and discussion of the findings. It is very important sections were goals of the study are reported. Our study specific objectives were to unearth the known techniques of fighting call toll fraud, develop a toll fraud detection and prevention model and evaluate the model in VoIP environment.

## 4.2 Findings for objective one

The study sought to find out the existing methods in detecting and preventing toll fraud in VoIP infrastructure. After literature review at least seven methods recorded as per Table 4.1

| | **Method** | **Weakness** |
|---|---|---|
| 1 | Honeynet | Use many resources, work offline |
| 2 | User CDR Profiling | Too much data to analyze, Work offline |
| 3 | Destination CDR Profiling | Too much data to analyze ,Work offline |
| 4 | Local Outlier Factor | Probabilistic (random) approach |
| 5 | Artificial Neural Network | Use many resources, work offline |
| 6 | Rule-based | |
| 7 | Mining Anomalous Behavior | Work offline |

Table 4.1: Toll fraud detection methods

VoIP fraud commonly is done by a third party making toll calls of which she/he doesn't pay but billed to the corporate business. That is, toll call fraud involves hacking a VoIP PBX and initiating long distance calls of which are billed but the hacker never pays, but the business must pay. The hacker re-sale the calls thus a lucrative business where he/she has not invested but a death to the business, mainly SMEs. We found this is one of the reasons that make corporate business think twice before adopting IP PBX, despite its many benefits.

To compact the toll fraud a number of techniques had been devised. The approaches are such as system authentication, user call profiling, artificial neural networks, honeynets, local outlier factor (LOF), Bayesian network, etc. Most of these methods rely on user call detailed records and it is usually offline data. These techniques have not yet solved the problem of toll fraud.

Our study suggested an alternative technique of solving toll fraud based on real-time (online) data. We have designed and implemented a toll fraud detection and prevention IP PBX model that classifies calls as either normal calls or fraudulent calls.

## 4.3 Findings of objective two

We relied on literature review analysis that provided the model specifications. These specifications enabled us to design the Toll Fraud detection model. The model was installed in a science laboratory. A various tests were done to ascertain its functionalities. Telephone calls (independent variables) were initiated and their timing done. When we compared the measured time with the expected time there was conformity and consistency.

### 4.3.1 Toll Fraud Detection and Protection Model Design

The model is based on a typical IP PBX infrastructure. The model outlines the interconnections, interfaces, relationships and characteristics of its various network elements.

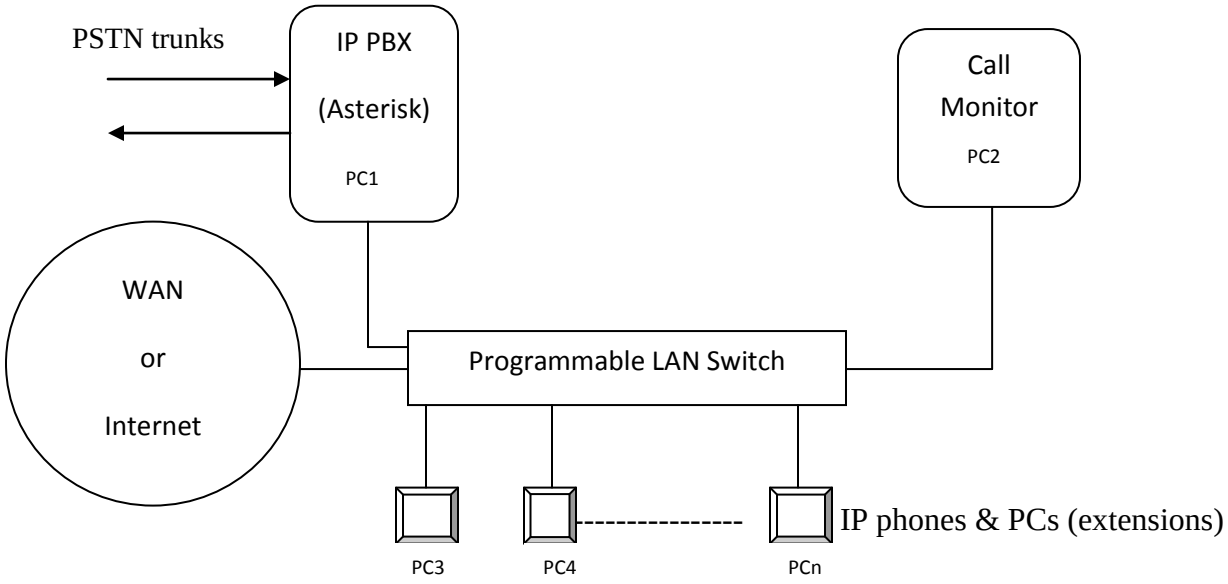



Figure 4.1: Toll Fraud Detection and Prevention Model

### 4.3.2 IP PBX

The IP PBX is open software (Asterisk, also called TrixBox) was loaded in a normal desktop computer. The software has the most essential telephony functions. Among the functions is the CDR, which is required as the primary data for the study.

### 4.3.3 IP Phones and PCs

These are the user end devices (extensions), used to make and receive calls. The IP phones are physical telephone sets with RJ 45 connectors for hooking on the LAN. The desktop computers are loaded with soft phone application such as X- Lite, Zoiper and 3CX phones. This work has used X-Lite softphone. For the audio transducers (microphone and ear piece) headsets had been used and a good audio communication achieved.

### 4.3.4 Programmable LAN Switch

The switch interconnects the network devices by doing the packet switching. It has a remote access port for attaching administration devices, such as a packet analyzer, e.g. Wireshark.

### 4.3.5 PSTN trunks

These are circuits that interface the PSTN (Public Switched Telephone Network) with the private voice network (IP PBX). The number of the trunks depends with how busy the telephone is. That is to say the number of trunks depends on the corporate business demand on external voice communication, and also the number of lines the company can afford.

### 4.3.6 Call Monitor

Offers additional facilities in call recording, which is of primary importance in the toll fraud study. It is connected on the switch so that it can record all interactions between the connected nodes (network elements). The study is using Wireshark, an open source packet analyzer software.

### 4.3.7 WAN/Internet

This is a gateway to the data networks, the Internet or a corporate WAN. Some details such as routers, firewalls, etc. are omitted because the study used a model, rather than full system.

## 4.4 Findings of objective three

### 4.4.1 Method 1: Call Setup Delay Timing

This was the core of our study. Our hypothesis is delay time is directly proportional with distance between any two user agents (subscribers) communicating. As the distance increases so it the delay time taken to setup the call. The measure time when we compared with the given timings by ITU-T, there was conformity. ITU-T (1999) provides timing in seconds 3, 5 and 8 for local, national and international calls respectively. We measured less than 3 seconds for local calls and about 6 seconds for national (mobile) calls.

The expected call setup delay times (ITU-T E.721, 1999) are as per table below:

Table 4.2: Call setup delay time as per call type

| CallType | Loc | Nat | Int |
|---|---|---|---|
| DelayTime (s) | 3 | 5 | 8 |

Graphically as shown below.

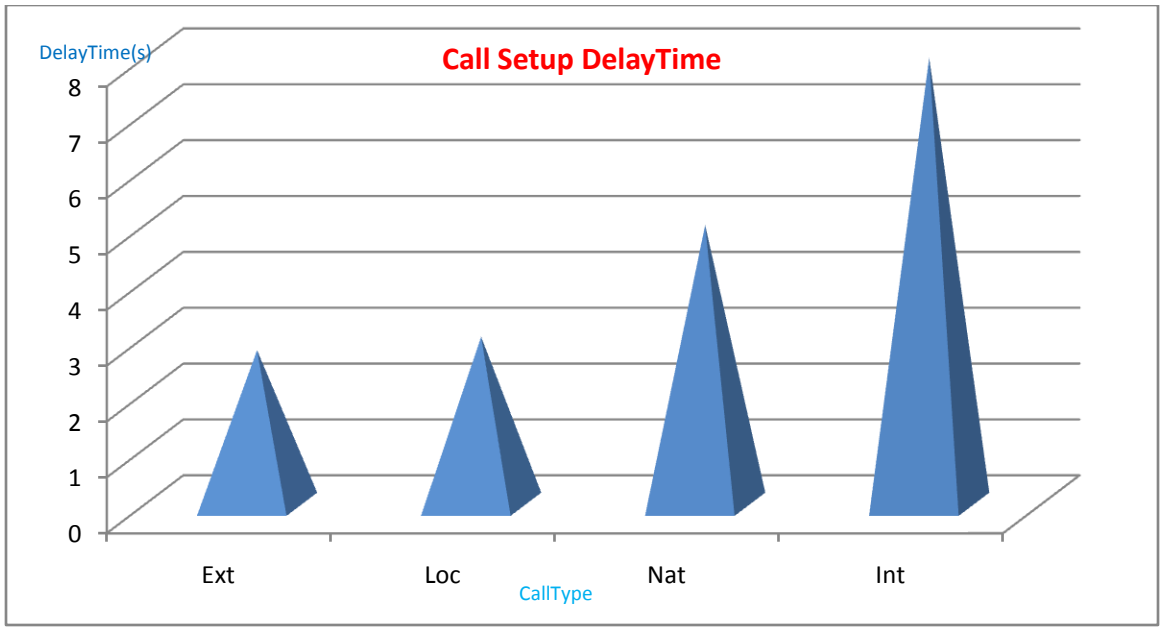



Figure 4.2: Expected call setup delay time

***Experiment 1: CSDT for Local Calls***

The experiment was set for extension (201) to extension (202) calling. A Smartphone stopwatch was used to record the call setup delay time. Ext. 201 dials ext. 202 (or vice versa). Call disconnected, the REDIAL key clicked simultaneously with timing start of stop watch. This is repeated 12 times (12 samples). The results are tabulated below

Table 4.3: Local call setup delay time

| CallNum | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DelayTime(s) | 2.84 | 2.73 | 2.83 | 2.78 | 2.69 | 2.79 | 2.74 | 2.83 | 2.70 | 2.87 | 2.73 | 2.79 |

Graphical representation is shown on Figure 4.3

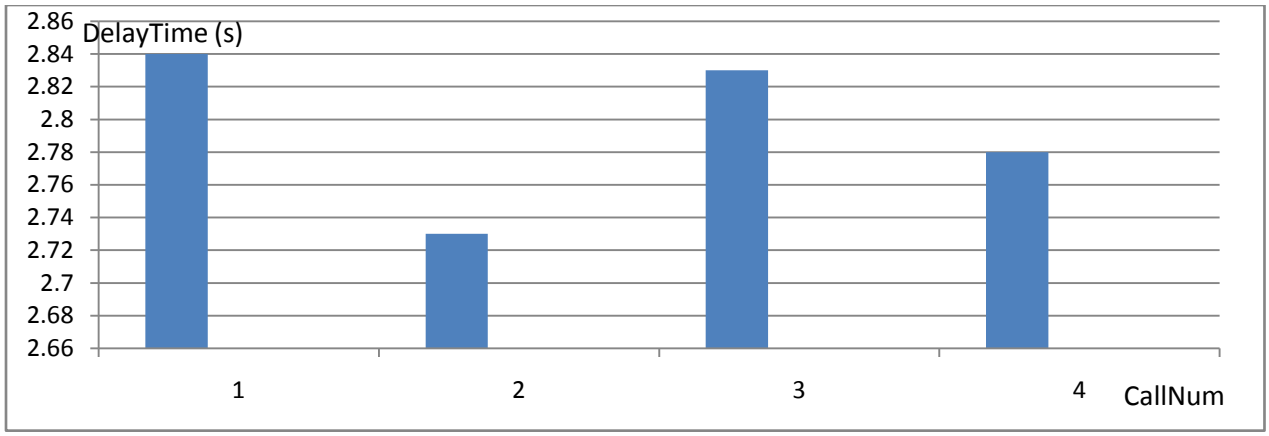



Figure 4.3: Call setup delay time for local calls

***Experiment 2: CSDT for National Calls (*Mobile Toll Calls)**

Ext. 2558 dials mobile numbers 90722659753 (Safaricom) and 90731111111 (Airtel). Call disconnected, the REDIAL key started simultaneously with timing start of stop watch. This is repeated 6 times for each network number (total of 12 samples). The results are tabulated below

Table 4.4: Mobile call setup delay time

| CallNum | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DelayTime(s) | 4.07 | 4.61 | 4.77 | 4.57 | 5.44 | 5.32 | 4.97 | 5.49 | 4.00 | 5.04 | 5.44 | 3.87 |

Graphical representation is shown on Figure 4.4

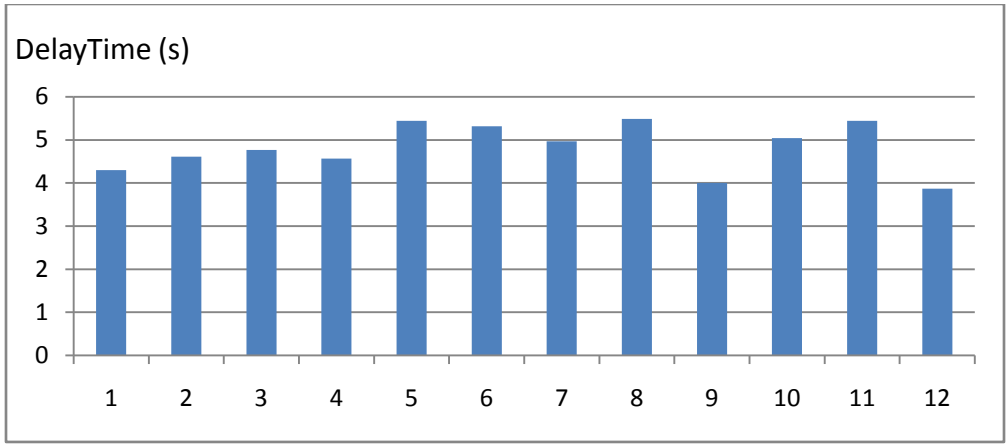



Figure 4.4: Mobile call setup delay time

***Experiment 3: CSDT for International Calls***

We were denied permission to make international calls. The reason was that once you access an international call link, even if the call is not answered, there are interconnect payments between service providers.

### 4.4.2 Method 2: Call Profiling

This method is to be set in tandem with CSDTs to monitor user behavior of calls classified normal and speech phase established. A local call can take long duration and no billing or low charges. While long call duration for national or international calls might be of suspect fraud. Two techniques are applied to classify normal and fraudulent calls.

A call detailed record (CDR) had been acquired from a local university. The relevant user profiles had been extracted (see Appendix I), and a summary data for analysis made.

Call selection criteria:

1. Call costs equal or greater than Ksh. 15.00: Technique 1
2. Call duration equal or greater than 3 seconds: Technique 2

**Technique 1: Call cost equal or greater than Ksh. 15.00**

Eleven samples of the records had been selected with the aim of setting threshold (quota) on allocation of maximum billing as per extension. If the allocated amount is exceeded the system disconnects the call and alerts administrator.

Table 4.5: Call cost data

| CallNum | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CallCost (Ksh) | 20 | 40 | 20 | 17 | 40 | 26 | 48 | 19 | 120 | 32 | 19 |

Cost(Ksh)

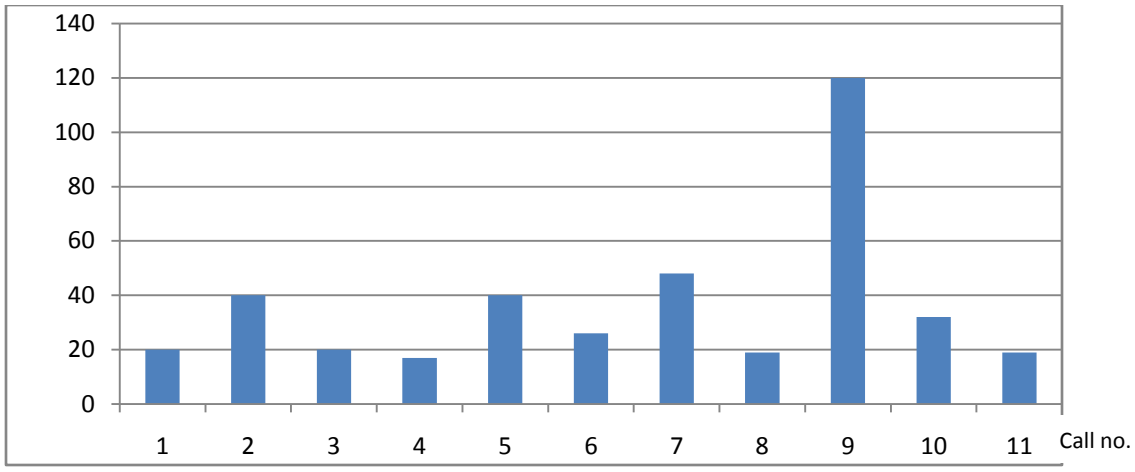



Figure 4.5: Call cost

### Technique 2: Call duration equal or greater than 5 minutes

Callers have tendency of talking long when they are no paying a call. Monitoring long duration calls classify normal and fraud calls depending on a system set thresholds. We selected eleven records for our analysis.

Table 4.6: Call duration

| CallNum | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Callduration(s) | 5 | 11 | 5 | 5 | 11 | 7 | 13 | 5 | 31 | 9 | 5 |

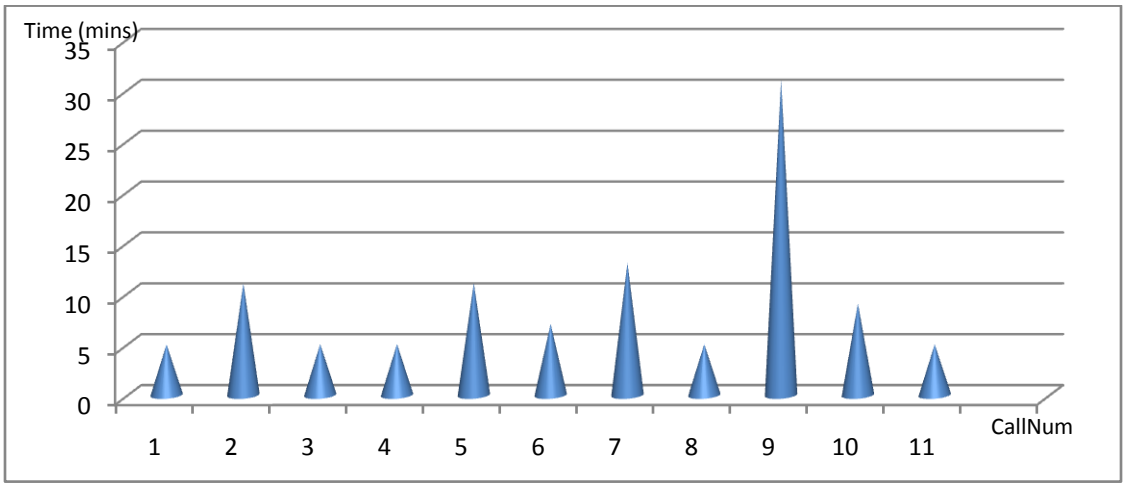



Figure 4.6: Call Duration

### 4.4.3 Analysis of the Data Collected

#### Method 1 Data Analysis

Data collected in Experiment 1 & 2 was used to classify a call (independent variable) as fraudulent or normal call (dependent variable).

The VoIP PBX model subscriber calls are categorized as either normal or fraudulent based on the following attributes of a call:

1. Source number: shows the number that made the call. A registered user caller should have the correct DN within the company calling numbering plan

2. Destination number: the attribute specifies the DN where the call terminates. Through digit translation call type, hence location of the CLD. The call termination is local, national or international destination. The destination should be within the company allowed outgoing calls to be considered a normal call.

3. Call setup delay times (A & B). Elsewhere in this document call setup delay time was defined as the time duration taken as from the first digit dialed up to the first ringing/ring back. The setup delay time A in seconds is measured during initial installation of the VoIP PBX system. The measurements will involve minimum and maximum delay of a particular call type. This forms the lower and upper thresholds for discriminating normal and fraud calls. The following where the average site set thresholds for our tests:

    - Extension to extension (and other local) calls, in range of 1 to 3 seconds
    - National (fixed and mobile) calls, in range of 3 to 6 seconds
    - International calls, in the range of 6 to 9 seconds.

The real-time measurement of a call connection setup time measured form call setup delay time B in seconds. When delay time B is taken it is compared with set time A, if within the threshold there is no fraud, thus genuine call. Otherwise B outside the threshold limits fraud is flagged and the system disconnects the call and at the same time alert signals are send to control panel and the administration.

**VoIP Calls Classification: No fraud status**

Table 4.7: No fraud only normal calls

| Origination number | Destination number | Call type | Measureds etup delay time (A) | Site setdelay time (B) | A>B | Call class |
|---|---|---|---|---|---|---|
| 2558 | 2081 | local | 2 | 3 | no | normal |
| 2067 | 90722659753 | national | 6 | 6 | no | normal |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2081 | 9011120211111 | international | 7 | 8 | no | normal |
| 2048 | 2558 | local | 3 | 3 | no | normal |
| 2222 | 999 | Special/local | 2 | 3 | no | normal |
| 2034 | 90202222222 | national | 5 | 6 | no | normal |
| 2340 | 90731111111 | national | 6 | 6 | no | normal |

Table 4.7 show a sample classification report for a selected CDR data. The call records from the selected CDR file have all been classified as normal calls, thus **no fraud**.

### VoIP Calls Classification: Detected fraud cases

Table 4.8: Fraud calls detected

| **Origination number** | **Destination number** | **Call type** | **Measured setup delay time (A)** | **Site set delay time (B)** | **A within B** | **Call class** |
|---|---|---|---|---|---|---|
| 2558 | 2081 | local | 2 | 1 to 3 | yes | normal |
| 2067 | 90722659753 | national | 2 | 3 to 6 | no | **fraud** |
| 2081 | 9011120211111 | national | 8 | 3 to 6 | no | **fraud** |
| 2048 | 90086790808283 | international | 7 | 6 to 9 | yes | normal |
| 2222 | 999 | Local | 2 | 1 to 3 | yes | normal |
| 2034 | 90202336711 | national | 4 | 3 to 6 | yes | normal |
| 2340 | 900358722659753 | international | 2 | 6 to 9 | no | **fraud** |
| 2081 | 9073905723 | national | 5 | 3 to 6 | yes | normal |

Table 4.8 shows a sample classification report where a number of callers had been detected as fraudulent callers. The 2nd call is marked by the model as fraud for the reason that the comparison of measured is not within the threshold set time range, 3 to 6 seconds for national calls. For this reason the call is suspicious and the system disconnects. Call number three the delay time measured is 8 seconds which is long enough for international but in the call analysis the system decode the destination as national. The 7th call the delay measure 2 seconds though call type decoded as international. This is a suspicious call and the system disconnects it. Figure 4.x show the call classes.

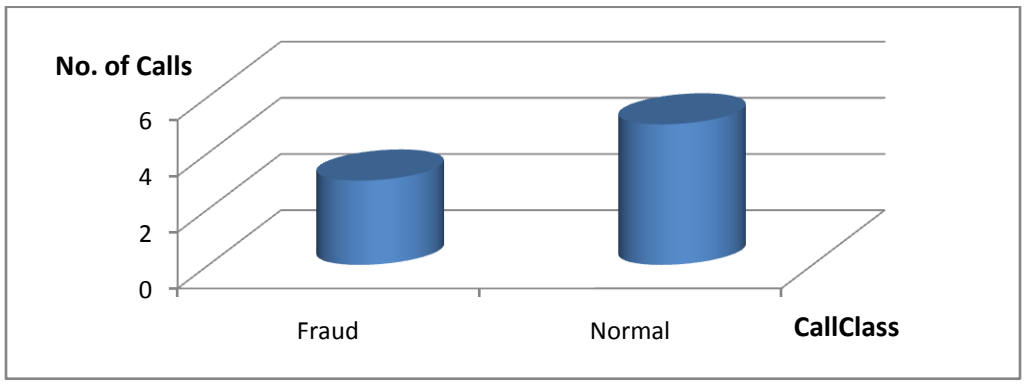



Figure 4.7: Call classification

### Method 2 Data Analysis

Data collected in Technique 1 & 2 was used to classify a call (independent variable) as fraudulent or normal call (dependent variable). Threshold was set on how long a call takes.

Table 4.9: Call classification

| Source DN | Destination DN | Call type | Total call duration(mins) | Call class |
|---|---|---|---|---|
| 2067 | 90722659753 | national | 1 | normal |
| 2081 | 9011120211111 | international | 30 | fraud |
| 2048 | 90086790808283 | | 20 | fraud |
| 2222 | 2081 | local | 20 | normal |
| 2034 | 90202336711 | | 60 | fraud |
| 2340 | 900358722659753 | international | 1 | normal |
| 2081 | 9073905723 | national | 3 | normal |

The table above and the figure below show call data with fraud cases. The toll fraud detection model starts by measuring call setup time (which relates to destination distance) and compare with site set thresholds. If the condition is satisfied, the call is deemed normal and connected. The second stage is the model to access call duration. The sample CDR data given above show three fraud calls detected. These are 2nd, 3rd and 5th calls, and the reason of marking the fraud is because the unusual long duration considering that they are long distance calls and cost a lot. The criteria used here is the set threshold for minimum long distance call cost/duration (quota) allocated per extension number. If the threshold exceeds the call is marked fraud and disconnected and alerts sent.

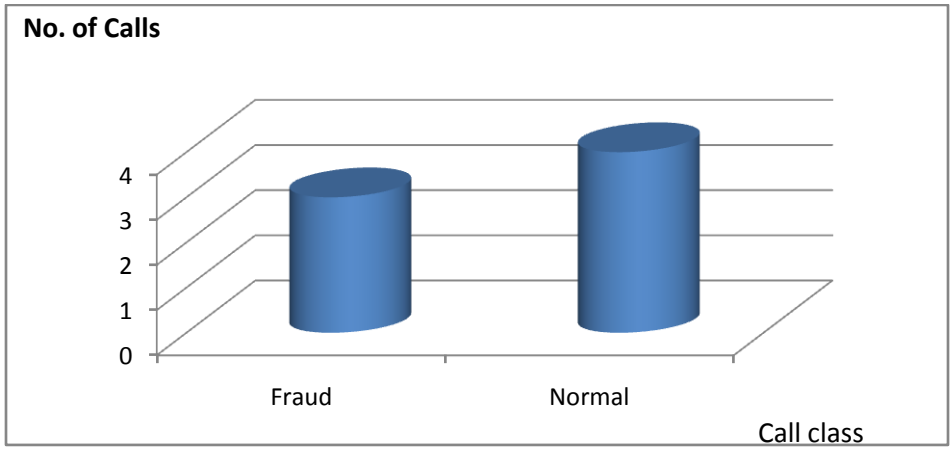



Figure 4.8: Call class

### 4.4.4 Challenges in Call Testing

We sought permission to use a university working VoIP PBX. The international call testing was not allowed because security reasons.

The timing of the calls was done using a Smartphone stopwatch.

**Classification of calls**

Telephone calls were categorized as normal and fraudulent. The techniques used were call setup delay timing (CSDT), call duration and cost monitoring. If a set threshold exceeds the model is to disconnect the call.

Overall, the experimental results show have sown that the proposed approach (CSDT) can be effective in detection and prevention of toll fraud in VoIP PBX infrastructure, overcoming limitations existing methods which are often offline.

## 4.5 Final Model

The data analyzed yielded the empirical values of call setup delay time (CSDT). The measured time is set between 1 and 3 seconds for local calls, 3 to 6 seconds for national calls and 6 to 9 seconds for international calls. For international calls, range 6 to 9 had been deducted from ITU-T recommendation.

The technique is implemented using the minimum hardware and software thus; it forms a Model VoIP PBX infrastructure.

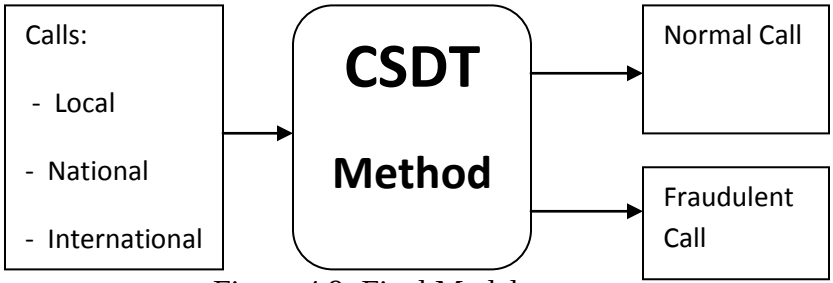



Figure 4.9: Final Model

## 4.6 Discussion of results

The purpose of discussion is to interpret and describe the significance of study findings in relation to what is exiting.

### 4.6.1 Discussion of objective one results

The goal of this objective was to find out the existing methods of fighting toll fraud in VoIP infrastructure. A literature review was done using ***literature based research methodology.*** A ***systematic literature review method*** was designed where the target population was all

sources where data was found. The source materials are such as latest publications, journals, text books, face to face discussions, etc. A number of data bases were accessed to download most of the data collected. These data bases are such as Google Scholar, Google Search Engine, Springer, IEEE Explorer, etc.

The search criterions were limited to the relevant literature review of before year 2013, while the latest literature review (state of art) of up to year 2017. The two categories form two research groups of the target population. The literature before year 2013 gave most of the research study's background information. The state of art (SOA) review identified four major works (methods), given on Table 2.1 of Chapter Two. This made our sample size to of 4 (four). Guber et al, (2013) used Honeynet method to detect toll fraud and recommended further study on **signaling** protocol. Wiens et al, (2014) approach on toll fraud in VoIP networks was on user profiling and the study recommended further work on call **destination**. The Wiens et al, (2015) studied the problem using technique of destination profiling and recommended use of **labeled data**. Final sample is that of Koiser (2016) who used a method of artificial neural networks to detect toll fraud in VoIP infrastructure. Koiser demonstrated way of classifying calls as genuine or fraudulent.

Detailed analysis and criticism resulted to identification of a research gap in fighting call toll fraud in VoIP environment. That is to say, we synthesized signaling, destination and labeled data information. Our method CSDT (call setup delay time) is a technique we have used to measure **signaling time**. This time is utilized to determine the fixed distance between the called (**destination**) and the calling parties. The time (distances) we measured for local, national and international calls make the labeled data (thresholds) that is used in the CSDT method to compare with actual time measured. So signaling time, origin-destination distances formed our basis for our method for detecting and preventing toll fraud in VoIP PBX infrastructure.

### 4.6.2 Discussion of objective two results

VoIP fraud generally involves 3rd party making frequent long distance, long duration toll calls at expense of a corporate business. VoIP callers would be classified as either fraudulent or normal based on the following attributes:

- Source number
- Destination number
- Call setup delay time (CSDT)
- Total call duration/cost, etc.

The main goal of the research study was to design and deploy a VoIP PBX model that classifies VoIP calls as either normal or fraudulent using call setup delay timing (CSDT). Local and national call test samples were simulated for the study. Calls from PBX extensions were timed, i.e. time between the 1st dialed digit and the 1st ringback signal.

The following are some of previous research works (methods):

1. Honeynet
2. User CDR profiling
3. Local outlier factor (LOF)
4. Artificial neural networks (ANNs)

Most of these techniques rely on offline data. Data is collected and analyzed to classify normal and fraud calls. Too much data is required for a good analysis hence detection of call fraud. The Honeynet and ANNs methods use many resources to setup trap for detection and prevention of toll call fraud. The methods seem to cost more on overheads.

The LOF uses probabilistic (random) approach to classify toll fraud. This means there mighty be possibilities of false positives and false negatives in fraud detection.

When we compare our technique with the mentioned methods; the CSDT seems a better alternative. It collects data for analyzes in real-time mode, analyze and categorize calls there by then. Also once a fraud call is detected, it is disconnected immediately.

The study model has a good consistence of call connection timing. The research established that the CSDT is successful technologies that classify legal and fraud calls in a telephone exchange.

The implementation of CSDT method will be of great step in fighting toll fraud in VoIP networks.

### 4.6.3 Discussion of objective three results

The toll fraud detection and prevention model was tested under controlled conditions to examine the hypotheses that the call setup time increases as the distance between the caller (sender) and the called (receiver) parties increases. It showed valid and consisted results.

# CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Summary

This study was intended detect and prevent toll call fraud in VoIP PBX, a crime that can deplete business, especially SMEs. The call theft, especially long distance calls and premium number calls must be eliminated.

According to the research findings, the impact of toll fraud is financial losses that a business incurs by paying calls made by fraudulently. The call thieves, after hijacking the PBX make expensive calls (and resale the calls) of which it is the business that pays for the services which it never used. It is a loss, embarrassing, suspicion of users' calling behavior, etc.

The study found out that call setup delay time is almost 3 seconds for local calls and almost 5 seconds for national calls as recommended by ITU (1999). The tests showed good consistency of the measured values. It was found that internal calls (within the premises) the delay time is less than 3 seconds. These are agents within distances of less than a kilometer. Similar testing done on agents at distances of about 20 kilometers (local calls) from the exchange (PBX) the expected call setup delay time is around 3 seconds, almost same as extension to extension calls.

As for national calls, coverage of distance hundreds of kilometers, a longer call setup delay time of 5 seconds is expected. The tests measured around 6 seconds for national calls. This was done timing call setup time for mobile calls. The long distance (international) calls cover thousands of kilometers, thus long call setup delay time of 8 seconds is expected (ITU, 1999).

This call setup delay time and the signal propagation speed in a particular medium (copper/fiber/air) convert to distances between the calling agent, the called agent, and the PBX. The fixed distance between the source/destination agents is very important in the toll fraud detection and prevention. During initial installation of IP PBX, a site distance measurements standard distances of local, national and international destinations should be taken to make a site mapping table in the system. The table will be the master reference (training) data to compare with actual caller distance for filtering legal and fraudulent calls.

It was found that the call setup delay time (signaling) of DP (dial pulse) telephone was longer compared to that of DTMF (dual tone multi-frequency) telephone. Most of current phones are DTMF phones, so the study concentrated on multi-frequency voice devices/soft phones.

The study also found that monitoring call billing behavior in the call accounting subsystem can detect and prevent toll fraud. A fraud call may pass the first fraud detection, thus why there is need to monitor connected calls. Figure 4.17, Figure4.18and Figure 4.x shows graphs of call cost, call duration and call class respectively. The data analyzed are CDR from a field working IP PBX of an institution. The CDR user profiles analyzed show that thresholds (limits) set will detect fraudulent call billing. If threshold is attained/exceeded alerts will be triggered.

## 5.2 Recommendations

The VoIP PBX is very important in enterprise business tool for communication. Voice (speech) is the best form of real-time communication. And more to this, video phone, a service supported by IP PBX, makes the communication seem, feel, test, etc. as it is a face to face natural dialogue. The security of the PBX is of great importance and mainly when it comes to hijacking of the whole exchange by cyber criminals. Our alternative method we didn't finish the testing because of challenges in interconnections to the PSTN. We recommend for the tests, otherwise, we deduced conclusion using the consisted results of the internal calls from the VoIP PBX model.

## 5.3 Suggestion for further study

To come up with an application software for fraud detection and prevention for this study (call setup delay time). The software will ensure real-time detection and action trigger when a fraud is detected. Also a proper timing measurement of the delay time is required. Our study used timing of automatic call redialing using a stopwatch to capture the timing. There was consistence of the results, but an in-built call setup delay time measuring subsystem is required for the precise measurements, hence avoid false positives or negatives.

## 5.4 The Research Study contributions

The CSDT (call setup delay time) method is an alternative technique in detecting and preventing toll fraud in VoIP infrastructure, and is of great significance to the field of telephony (telephone call processing) and other related institutions.. There are many VoIP stakeholders who are to benefit from the CSDT method. The beneficiaries are such corporate business who relies on PBX for voice communication in supporting business, Telephone Service Providers (TPSs), Communication Authorities, ISPs, Academic Institutions (for further research), etc. So we can say the CSDT (call setup delay time) is a contribution of knowledge in the field of Telecommunications and ICT.

## REFERENCES


Africa Cyber Security Report.(2016).

Agundez, R., Penya, Y., & Bringas, P. (2010). Fraud detection for Voice over IP services on next generation networks. In International Security Theory and Practices and Security and Privacy of Pervasive Systems and Smart Devices. Springer, Berlin, German.

Alves, R., et al. "Dicovering telecom fraud situations through mining anomalous behavior patterns," in Proceedings of the DMBA Workshop on the 12th ACM SIGKDD, 2006.

Best VoIP Providers 2017. Available from: getvoip.com/. Accessed on 12/11/2016.

Bordens, K.S., & Abbott, B.B. (2011). Research Design and Methods. (8thEdition). McGraw Hill. New York.

Burge, P., Shawe, J.,Cooke, C., Moreau, Y., Preneel, B., and Stoermann, C. (1997). Fraud detection and management in mobile telecommunications networks, in European Conference in Security and Detection.

Burge, P & Shawe, J. (1997). Detecting Cellular Fraud Using Adaptive Prototypes. In Proceedings AAA1-97 Workshop on AI Approaches to Fraud Detection and Risk Management.AAA1 Press.

Breuniq, M., Kriegel, H., Ng, R. and Sander, J. "LOF identifying density-based local outliers," SIGMOD Record Vol.29, no. 2, 2000.

Catherine, d., (2002). Practical Research Methods.Available from: www.modares.ac.ir/ Accessed on 5/11/2016.

Chandola V., Banjee A., and Kumar (2009). " Anomaly detection: A survey," ACM Comput.Surv. vol 41. No 3.p. 15: 1-15:58. 2009

Chapuis, R. and Joel, A. (2003). "Technology and Engineering 100 years of Telephone Switching." Available from: https://books.google.com/books. Accessed on 15/2/2017.

Cisco. CiscoVoIP Networking Design/ IP Telephony. Available from:

Cisconetworkingcenter.blogspot.co.ke. Accessed on 27/07/2016.

Cross-bar switching systems. Available from: course.sdu.edu.cn.Accessed on 11/11/2016.

Dempster et al. Trixbox Made Easy. Packet Publishing, UK. . Available from: trixboxmadeeasy. Asterisk.ru. Accessed on 15/11/2016.

Eyers, T., et al (2000). Predicting Internet Telephony Call Setup Delay.

Facet, T., & Provost, F. (1997). Adaptive Fraud Detection. Data Mining and Knowledge Discovery, vl 21, no 3, 1997.

Geneiatakis D. et al. Survey of Security Vulnerabilities in Session Initiation Protocol.Available from:.Accessed on 19/10/2016.

Grosser, H., Britos, P. and Garcia-Martinez, R. "Detecting fraud in mobile telephony using neural networks," in Proceedingsof the 18th International Conference of Innovations in Applied Artificial Intellegence, Bari, Italy, Springer-Verlag, 2005.

Gruber, M., Frakhauser, F., Taber, S., Schanes, C and Grechnig, T., 2012. Security Status of VoIP Based on the Observation of Real-World Attacks on Honeynet.

Guo, F., Sui, A. and Shi, L. "Billing attack detection and prevention in mobile communication network," in IEEE 13th International Conference on Communication Technology, 2011.

Hilas, C. and Mastorocostas, P. "An approach of supervised and unsupervised learning approaches to telecommunications fraud detection." Knowledge-Based Systems, Vol 21,no. 7, 2008.

Hoffstadt, D., Marold, A., & Rathgeb E.P. (2012). Analysis of SIP-Based Threats Using a VoIP Honeynet System.

Hoffstadt, D., et al. (2014) "A comprehensive framework for detecting and preventing VoIP fraud and misuse," in International Conference on Computing, Networking and Communications (ICNC), 2014.

Hollmen, J., Tresp, V. and Simula, O. "A Self-Organizing Map for clustering probabilistic models," in Ninth International Conference on Artificial Neural Networks (ICANN), vol. 2, 1999.

Hung P.C.K. Security Issues in VoIP Applications. University of Ontario Institute of Technology, Canada.

Hung P.C.K. et al. Through the Looking Glass: Security Issues in VoIP Applications. University of Ontario Institute of Technology.

History of PBX. . Available from:bebusinessed.com. Accessed on 11/11/2016.

Cisco Press, October 16, 2006.

ITSPA (Internet Services Providers' Association – Recommendations for secure deployment of an IP PBX – Version 2, November 2013. Available from: www.itspa.org. Accessed on 19/10/2016.

Kang, B., Kim,D. and Kang, S. “Extended KNN imputation based LOF prediction algorithm for real-time business process monitoring method,” The Journal of Society for e-Business Studies, vol.15,pp 303-317, 2010.

Kapourniotis,T., Digiuklas,T., Polyzos, G. and Alefragkis, P. “Scam and fraud detection in VoIP networks: Analysis and countermeasures using user profiling,” in 50th FITCE Congress, 2011.

Kenya Cyber Security Report 2016.Available from: www.serianu.com/news. Accessed on 4/9/2016.

Keromytic A.D. Voice-over-IP Security: Research and Practice. Colombia University Published in IEEE Security and Privacy (Volume 8 Issue 2).

Kim, K., Kim, T., Cho, N. and Kim, M. (2015). “Toll Fraud Detection of VoIP Service Networks in Ubiquitous Computing Environments,” in International Journal of Distributed Sensor Networks, 2015.

Kim, S., Cho, N.,Lee, Y. et al. Application of density-based outlier detection to database activity monitoring, information Systems Frontiers, vol. 15, no. 1, 2013.

Koiser, N., 2016. Toll Fraud in VoIP Networks Using Artificial Neural Networks. Available from: erepository.uonbi.ac.ke. Accessed on 3/3/2016.

Kothari, C., 2004. Research Methodology: Methods and Techniques. (2nd Edition). New Age International Publishers. Available from: www.modares.ac.ir/uploads. Accessed on 28/10/2017.

Kubler, S., Massoth, M., Wiens, A. and Wiens, T. “Toll fraud detection in Voice over IP networks using communication behavior patterns on unlabeled data,” in The Fourteenth International Conference on Networks (ICN 2015) IARIA, 2015, in press.

Kumar, R., (2005). Research Design and Methodology.Available from: Books.google.com/. Accessed on 1/10/2016.

Market Research Store. Global VoIP Services Market Poised to Surge. . Available from: globenewswire.com/news. Accessed on 15/11/2016.

Mobile switching centre. Available from: www.ccs.neu.edu/home...../cellularnetworks. Accessed on 15/02/2017.

Moreau, Y., Verreist, H. and Vandewalie, J. “ Detection of Mobile Phone Fraud Using Supervised Neural Networks: A first Prototype,” in Proceedings of the 7th International Conference on Artificial Neural Networks, Springer-Verlag, 1997.

Nassar, M., Noccolini, S., State, R. and Ewald, T. “Holistic VoIP intrusion detection and prevention system,” in proceedings of the 1st International Conference on Principles, Systems and Applications of IP Telecommunications, New York City, New York, ACM, 2007.

Olszewski, D., Kacprzyk, J. and Zadrozny,S. “Employing Self-Organizing Map for fraud detection,” in the 12th International Conference on Artificial Intelligence and Soft Computing (ICAISC 2013), 2013.

Patrick Cairns.Frost & Sullivan's analysis of the South African IP PBX market. patrick.cairns@frost.com

Pindrop Call Centre Fraud Report (2017). Available from: www.pindrop.com/. Accessed on 18/11/2016.

Rasol M., et al, 2016. An improved Secure SIP Registration Mechanisms to Avoid VoIP Threats. International Journal of Cloud Applications and Cloud Computing.

Rosset, S., Murad, U., Neumann, F., Idan, Y. and Pinkas, G. “Discovery of fraud rules for telecommunications-challenges and solutions,” in Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Diego, Calif, USA, August 1999.

Stored Program Control. Available from: www.daenotes.com. Accessed on 25/11/2016.

Taniguchi, M., Haft, M., Hollmen, J. and Tresp, V. “Fraud detection in communication networks using neural and probabilistic methods,” in Proceedings of the 1998 IEEE International conference on Acoustics, Speech and Signal Processing, vol2, 1998.

Telephone Tribute. “Switches, PBX, and Central Offices. Available from: www.telephonetribute.com/switches.hmtl. Accessed on 9/02/2017.

TELUS Communications Company (2017). Available from: www.telus.com/news. Accessed on 15/02/2017.

The Communications Fraud Control Association (CFCA) for Telecommunications Fraud Survey year 2015.

The Theory of Electro-Mechanical Switching. . Available from: www.seg.co.uk/teleco/automatic1.htm. Accessed on 10/11/2016.

Voice scrambler. Crypto Museum. Available from: www.cryptomuseum.com/crypto/voice.htm. Accessed on 15/02/2017.

Voxox VoIP Technology (2015). Available from: www.voxox.com/. Accessed on 15/03/2017.

Wiens, A., Kubler,S., Wiens, T. and Michael, M. “Improvement of user profiling, call destination profiling and behavior pattern recognition approaches for telephony toll fraud,” in International Journal on Advances in Security, vol 8 no 1&2, year 2015.

Wiens, A,, Wiens, T. and Massoth, M. “A new approach for detecting toll fraud in Vo IP networks, “in the Tenth Advanced International Conference on Telecommunications (AICT 2014) IARIA, 2014.

Wiens, A,, Wiens, T. and Massoth, M. “Approach on fraud detection in Voice over IP Networks using call destination profiling based on analysis of recent attacks on FRITZ!Box units,” in the Sixth Interantional Conference on Emerging Network Intelligence (EMERGING 2014) IARIA, 2014.

Voxox. (2015). PBX Fraud: Stay Informed Against Fraudulent Calls.” VoIP Technology. Available from: assist.voxox.com. Accessed on 21/04/2017.

## Appendix I

## CDR data from a local university

User 2133

| | Date | Time | Destination | Duration | Cost (Ksh.) |
|---|---|---|---|---|---|
| 1 | 31/7/2017 | 11:55:45 | 0715255393 | 05.18 | 19.55 |
| 2 | 31/7/2017 | 12:10:51 | 0715255393 | 02.20 | 08.61 |
| 3 | 31/7/2017 | 12:37:21 | 0715255393 | 11.00 | 41.13 |
| 4 | 31/7/2017 | 12:50:00 | 0715255393 | 05.23 | 19.86 |
| 5 | 31/7/2017 | 14:45:54 | 0715255393 | 00.35 | 02.15 |
| 6 | 01/8/2017 | 11:03:12 | 0715255393 | 01.36 | 05.90 |
| 6 | 01/8/2017 | 11:53:35 | 0715255393 | 00.51 | 03.14 |

User 2134

| | Date | Time | Destination | Duration | Cost (Ksh.) |
|---|---|---|---|---|---|
| 1 | 31/7/2017 | 09:53:58 | 020271710 | 04.59 | 17.34 |
| 2 | 31/7/2017 | 09:57:57 | 0721935911 | 00.51 | 02.14 |
| 3 | 31/7/2017 | 09:59:25 | 0725208204 | 01.02 | 03.81 |
| 4 | 31/7/2017 | 11:10:14 | 0722333253 | 00.24 | 01.48 |
| 5 | 31/7/2017 | 11:53:42 | 0202989000 | 02.09 | 07.48 |
| 6 | 31/7/2017 | 11:57:59 | 0720260995 | 02.09 | 07.93 |
| 7 | 01/8/2017 | 09:37:00 | 0725247285 | 00.59 | 03.63 |
| 8 | 01/8/2017 | 09:49:25 | 0774889316 | 03.55 | 13.63 |
| 9 | 01/8/2017 | 10:44:31 | 0770254098 | 01.28 | 05.10 |
| 10 | 01/8/2017 | 11:17:47 | 0731400200 | 11.37 | 40.43 |

| | | | | | |
|---|---|---|---|---|---|
| 11 | 01/8/2017 | 11:33:41 | 0728500120 | 03.00 | 11.07 |
| 12 | 01/8/2017 | 11:45:29 | 0725247285 | 00.41 | 02.52 |
| 13 | 01/8/2017 | 12:13:26 | 0723349323 | 00.31 | 01.91 |
| 14 | 01/8/2017 | 12:24:27 | 0712690313 | 00.25 | 01.54 |
| 15 | 01/8/2017 | 14:49:42 | 0712690313 | 00.14 | 00.86 |
| 16 | 01/8/2017 | 15:01:47 | 0712690313 | 00.25 | 01.54 |
| 17 | 01/8/2017 | 15:01:10 | 0712690313 | 01.32 | 5.66 |

User 2413

| | **Date** | **Time** | **Destination** | **Duration** | **Cost (Ksh.)** |
|---|---|---|---|---|---|
| 1 | 31/7/2017 | 13:05:10 | 0726075831 | 00.21 | 01.29 |
| 2 | 01/8/2017 | 11:19:45 | 0722629310 | 01.18 | 04.80 |
| 3 | 01/8/2017 | 15:24:25 | 0719262034 | 00.46 | 02.83 |
| 4 | 01/8/2017 | 15:26:19 | 0720995207 | 00.58 | 03.57 |
| 5 | 01/8/2017 | 15:32:01 | 0718222425 | 01.58 | 07.25 |
| 6 | 01/8/2017 | 15:40:16 | 0724201810 | 00.51 | 03.14 |
| 6 | 01/8/2017 | 15:42:45 | 0719591979 | 01.37 | 05.96 |

User 2005

| | **Date** | **Time** | **Destination** | **Duration** | **Cost (Ksh.)** |
|---|---|---|---|---|---|
| 1 | 31/7/2017 | 09:22:20 | 0720780742 | 00.41 | 02.52 |
| 2 | 31/7/2017 | 11:19:17 | 079883255 | 02.09 | 7.93 |
| 3 | 31/7/2017 | 13:07:54 | 075074218 | 00.11 | 00.68 |
| 4 | 31/7/2017 | 13:09:09 | 0704515694 | 00.33 | 02.03 |

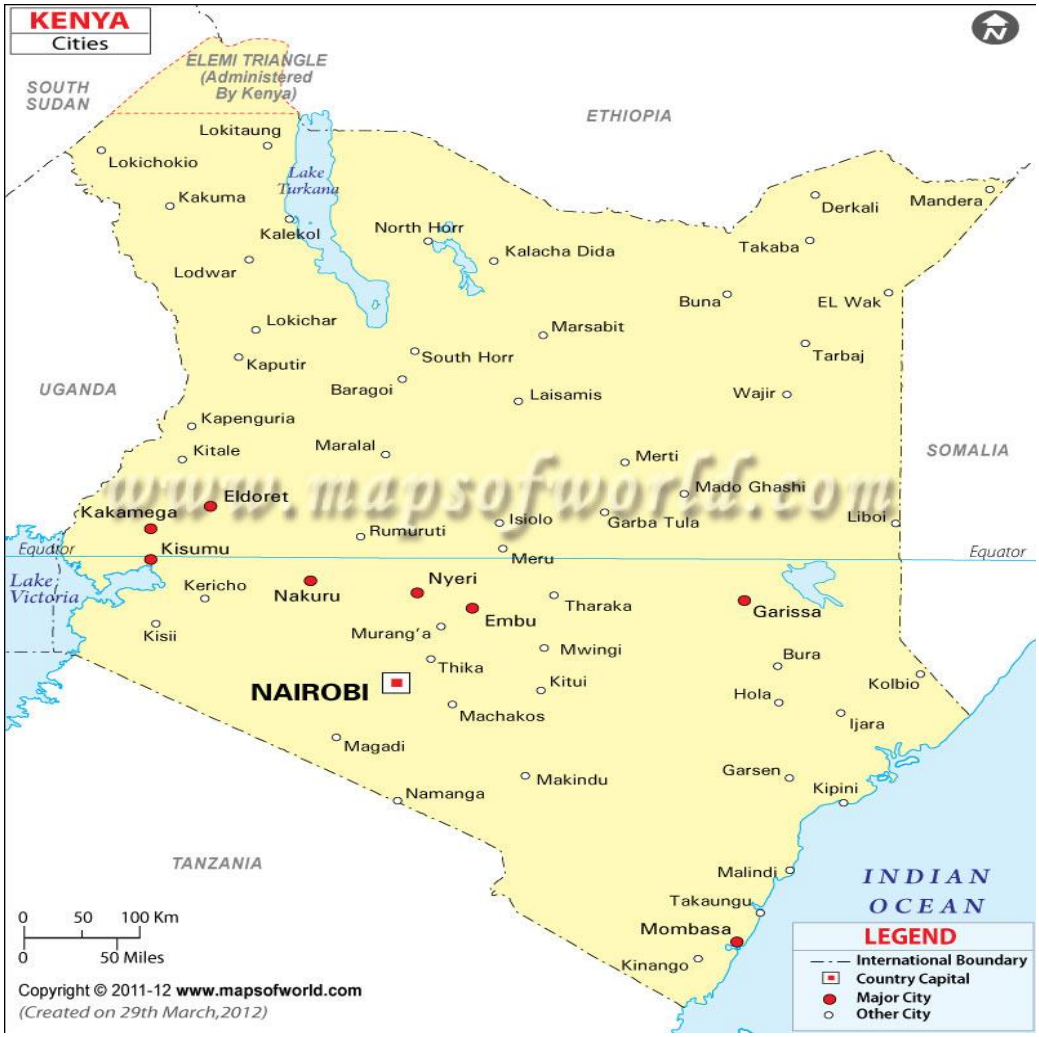
| | | | | | |
|---|---|---|---|---|---|
| 5 | 31/7/2017 | 13:09:53 | 0704515694 | 02.42 | 09.96 |
| 6 | 31/7/2017 | 13:19:18 | 0705545815 | 00.22 | 01.35 |
| 7 | 31/7/2017 | 14:12:20 | 0724391479 | 02.54 | 10.70 |
| 8 | 31/7/2017 | 14:16:20 | 0724391479 | 00.43 | 02.64 |
| 9 | 31/7/2017 | 14:16:10 | 0797109283 | 02.17 | 08.42 |
| 10 | 31/7/2017 | 14:17:33 | 0797109283 | 00.18 | 01.11 |
| 11 | 31/7/2017 | 14:23:03 | 0797109283 | 00.23 | 01.41 |
| 12 | 31/7/2017 | 14:27:13 | 0725473007 | 00.27 | 01.66 |
| 13 | 31/7/2017 | 14:43:53 | 0725473079 | 02.11 | 08.05 |
| 14 | 01/8/2017 | 11:59:56 | 0716906361 | 01.14 | 04.55 |
| 15 | 01/8/2017 | 12:11:44 | 0716906361 | 02.08 | 07.87 |
| 16 | 01/8/2017 | 16:18:59 | 36088378 | 01.26 | 04.90 |
| 17 | 01/8/2017 | 16:27:01 | 0728398124 | 00.15 | 00.92 |

User 2049

| | **Date** | **Time** | **Destination** | **Duration** | **Cost (Ksh.)** |
|---|---|---|---|---|---|
| 1 | 31/7/2017 | 17:17:08 | 0789556659 | 01.04 | 03.71 |
| 2 | 31/7/2017 | 17:20:00 | 0717220350 | 01.05 | 04.00 |
| 3 | 31/7/2017 | 17:44:45 | 0720342251 | 02.58 | 10.94 |
| 4 | 31/7/2017 | 18:22:42 | 0723873741 | 07.01 | 25.88 |
| 5 | 31/7/2017 | 18:55:34 | 0724246830 | 12.59 | 47.89 |
| 6 | 31/7/2017 | 19:25:44 | 0719181720 | 05.06 | 18.81 |
| 7 | 31/7/2017 | 19:34:29 | 0726042174 | 31.17 | 115.40 |
| 8 | 01/8/2017 | 09:49:30 | 0726356094 | 03.49 | 14.08 |

| | | | | | |
|---|---|---|---|---|---|
| 9 | 01/8/2017 | 10:08:06 | 0724835138 | 08.54 | 32.83 |
| 10 | 01/8/2017 | 10:57:48 | 0727888679 | 05.06 | 18.81 |
| 11 | 01/8/2017 | 11:18:04 | 0720176103 | 00.25 | 01.54 |
| 12 | 01/8/2017 | 12:01:01 | 0726042174 | 00.50 | 03.07 |
| 13 | 01/8/2017 | 12:02:17 | 0726042174 | 01.32 | 05.65 |
| 14 | 01/8/2017 | 15:11:21 | 0726446770 | 01.58 | 07.25 |
| 15 | 01/8/2017 | 17:48:39 | 0720176103 | 03.03 | 11.25 |
| 16 | 01/8/2017 | 17:54:21 | 0720176103 | 03.03 | 11.25 |
| 17 | 01/8/2017 | 18:07:33 | 0720176103 | 03.03 | 11.25 |

## Appendix II


KENYA
Cities
ELEMI TRIANGLE
(Administered
By Kenya)
SOUTH
SUDAN
ETHIOPIA
Lokitaung
Lokichokio
Lake
Turkana
Kakuma
Kalekol
North Horr
Kalacha Dida
Derkali
Mandera
Takaba
Lodwar
Buna
EL Wak
Lokichar
Marsabit
Tarbaj
Kaputir
South Horr
UGANDA
Baragoi
Laisamis
Wajir
Kapenguria
Kitale
Maralal
Merti
SOMALIA
Mado Ghashi
Eldoret
Kakamega
Isiolo
Garba Tula
Liboi
Rumuruti
Kisumu
Meru
Equator
Equator
Lake
Victoria
Kericho
Nakuru
Nyeri
Tharaka
Embu
Garissa
Kisii
Murang'a
Mwingi
Bura
Thika
NAIROBI
Kitui
Hola
Kolbio
Machakos
Ijara
Magadi
Garsen
Makindu
Kipini
Namanga
TANZANIA
Malindi
INDIAN
OCEAN
Takaungu
Mombasa
LEGEND
Kinango
International Boundary
Country Capital
Major City
Other City
0 50 100 Km
0 50 Miles
Copyright © 2011-12 www.mapsofworld.com
(Created on 29th March,2012)
www.mapsofworld.com


WORLD MAP
N
Prime Meridian
Greenland (Denmark)
Alaska (USA)
Canada
United States of America
Mexico
Russia
China
India
Kazakhstan
Mongolia
Australia
Indonesia
Brazil
Peru
Bolivia
Paraguay
Chile
Argentina
Uruguay
Colombia
Venezuela
Ecuador
Guyana
Arctic Circle
Tropic of Cancer
Equator
Tropic of Capricorn
NORTH ATLANTIC OCEAN
NORTH PACIFIC OCEAN
SOUTH ATLANTIC OCEAN
SOUTH PACIFIC OCEAN
INDIAN OCEAN
SOUTHERN OCEAN
ANTARCTICA
1. Netherlands
2. Belgium
3. Luxembourg
4. Switzerland
5. Slovenia
6. Croatia
7. Bosnia and Herzegovina
8. Czech Republic
9. Slovakia
10. Austria
11. Hungary
12. Serbia
13. Moldova
14. Macedonia
15. Albania
16. Cyprus
17. Lebanon
18. Guinea-Bissau
19. Guinea
20. Ghana
21. Togo
22. Benin
23. Cameroon
24. Equatorial Guinea
25. Rwanda
26. Cambodia
27. Panama
28. Malawi
29. Liechtenstein
30. Montenegro
31. Kosovo
32. Palestinian Territories
33. St. Vincent and the Grenadines
1260Km 0 1260 2520 3780 5040Km
Copyright © 2013 www.mapsofworld.com
(Updated on 8th March, 2013)

**Abstract**

The document is a project work on study of call toll fraud. It presents a method of detection and prevention/mitigation of toll fraud in VoIP PBX infrastructure. The method is based on measuring call setup delay time, of which the measured time is used to calculate the fixed distance between the calling and called directory numbers. Distance cannot be spoofed like IP addresses or directory numbers, hence an alternative method to detect toll fraud. A lookup table built during initial installation provides radial distances (times) for local, national and international calls' destinations. A comparison between measured and threshold set values classifies normal and fraudulent calls. The document also provides a method for monitoring long duration calls. In this approach when a set threshold is exceeded, that call is suspicious, and is classified as a fraudulent call. In both techniques, a fraud call is disconnected and the same reported to the administrator.

Cyber criminals use brute force or else to get the PBX credentials. Once access allowed, they hijack the system and make it their call center. This crime is currently making businesses loss billions of money to fraudulent callers who make/re-sale calls without paying. VoIP is a new technology and its proper use has many benefits on national development of a country. The importance of VoIP services motivated the us to study the mentioned methods above.

## DEDICATION

This research work is dedicated to my mother Nthenya Muli, wife Mutheu and children Muli, Kyama, Mumo and Nduva and to all those who appreciate the flexible services offered by VoIP systems.

**ACKNOWLEDGEMENT**

As the first priority, thanks to the Mighty God for the health and the drive to go through the demanding task.

I pass my regards to all my lecturers, and mainly the faculty projects co-coordinator, Dr. S. Mwendia and my two supervisors Dr. R. Rimiru and Mr. C. Onsomu for their guidance, corrections and encouragement that enabled me to complete this study project.

My thanks go to Multimedia University for helping me in data collection for the research.

Thanks to the following; my wife Mutheu for financial support, encouragement and prayers, my brother Muema for many supports he gave and last but not least the MMU ICT support staff they were handy during the testing phase.

**Table of contents:**

**List of figures:**

**List of tables:**

## LIST OF ABBREVIATIONS AND ACRONYMS

| WORD | PHRASE |
|---|---|
| VoIP | Voice over Internet Protocol |
| IP | Internet Protocol |
| PBX | Private Branch eXchange |
| SIP | Session Initiation Protocol |
| ANN | Artificial Neural Network |
| NN | Neural Network |
| PC | Personal Computer |
| CDR | Call Detail Record |
| WAN | Wide Area Network |
| LAN | Local Area Network |
| PSTN | Public Switched Telephone Network |
| USD | US Dollar |
| CFCA | Communications Fraud Control Association |
| PMBX | Private Manual Branch eXchange |
| SxS | Step by step |
| CPU | Central Processing Unit |
| MSC | Mobile Switching Centre |
| LE | Local Exchange |
| CLG | CalLinG party |
| CLD | CalLeD party |
| PSD | Post Selection Delay |
| DN | Directory Number |
| ISDN | Integrated Services Digital Network |

| | |
|---|---|
| SVR | Server |
| TP | Twisted Pair |
| RF | Radio Frequency |
| i/p | Input |
| o/p | Output |
| AI | Artificial Intelligence |
| IETF | Internet Engineering Task Force |
| 1G | 1st Generation Mobile Voice Switching System |