Czech Technical University in Prague

Faculty of Electrical Engineering

Master’s Thesis

2015 **Nikulin Nikita**

Czech Technical University in Prague

Faculty of Electrical Engineering

Department of Telecommunications Engineering

Billing system’s extension for fraud detection

2015 **Author:** **Nikulin Nikita**

Supervisor: **Ing. Pavel Troller, CSc.**

STATEMENT

I declare that this master's thesis has been completed by myself with the assistance of supervisor and consultant and only given literature has been. Further, I declare that I have no objection to borrowing or disclosure of my master's thesis or part of it.

Date: 11. 5. 2015

_____

Signature

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra telekomunikační techniky

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Nikulin Nikita**

Studijní program: Komunikace, multimédia a elektronika
Obor: Komunikační systémy

Název tématu: **Rozšíření účtovacího systému pro detekci fraudu**

Pokyny pro vypracování:

Create a set of scripts or programs for scanning billing data and searching for patterns typical for fraudulent traffic. Criteria (thresholds, database of suspect numbers etc.) will be entered over a simple graphical user interface, probably web-based. Verify the functionality by running the scripts over billing data sets with and without known fraudulent activity.

Seznam odborné literatury:

[1] Hunter,J.M., Thiebaud,M.: Telecommunications Billing Systems. Amazon 2002, ISBN 0071408576

[2] Yarberry, V.A.Jr, DiMarsico, B.. Phelps, T.: Telecommunications Cost Management. Amazon 2002, ISBN 0849311012

Vedoucí: Ing. Pavel Troller, CSc.

Platnost zadání: do konce letního semestru 2015/2016

L.S.

V Praze dne 13. 11. 2014

SUMMARY

This final project deals with creating of automated program for scanning billing data and searching typical and untypical fraudulent traffic. Scripts are created in PHP using MySQL database environment. First chapter of this project deals with general information about fraud and methods of its detecting. Second chapter deals with created system's extension including examples of scanning and fraud detecting.

Index Terms:

Fraud detecting, detection of fraud, anti-fraud system

PLAN

INTRODUCTION

Fraud is a billion-dollar business and it is increasing every year. The PwC global economic crime survey suggests that close to 30% of companies worldwide have reported being victims of fraud in the past year.

Fraud involves one or more persons who intentionally act secretly to deprive another of something of value, for their own benefit. Fraud is as old as humanity itself and can take an unlimited variety of different forms. However, in recent years, the development of new technologies has also provided further ways in which criminals may commit fraud. In addition to that, business reengineering, reorganization or downsizing may weaken or eliminate control, while new information systems may present additional opportunities to commit fraud.

Taking into account the rapid pace of development of IP-based technologies and the increasing demands on the quality of communication the development of software tools for scanning billing data and detecting fraudulent traffic is an urgent task.

1. ANALYTICAL PART

1.1. FRAUD DETECTION

Traditional methods of data analysis have long been used to detect fraud. [1]They require complex and time-consuming investigations that deal with different domains of knowledge like financial, economics, business practices and law. Fraud often consists of many instances or incidents involving repeated transgressions using the same method. Fraud instances can be similar in content and appearance but usually are not identical.

The first industries to use data analysis techniques to prevent fraud were the telephony companies, the insurance companies and the banks (Decker 1998). One early example of successful implementation of data analysis techniques in the banking industry is the FICO Falcon fraud assessment system, which is based on a neural network shell.

Retail industries also suffer from fraud at POS. Some supermarkets have started to make use of digitized closed-circuit television (CCTV) together with POS data of most susceptible transactions to fraud.

Internet transactions have recently raised big concerns, with some research showing that internet transaction fraud is 12 times higher than in-store fraud.

Fraud that involves cell phones, insurance claims, tax return claims, credit card transactions etc. represent significant problems for governments and businesses, but yet detecting and preventing fraud is not a simple task. Fraud is an adaptive crime, so it needs special methods of intelligent data analysis to detect and prevent it. These methods exists in the areas of Knowledge Discovery in Databases (KDD), Data Mining, Machine Learning and Statistics. They offer applicable and successful solutions in different areas of fraud crimes.

Techniques used for fraud detection fall into two primary classes: statistical techniques and artificial intelligence. Examples of statistical data analysis techniques are:

- Data preprocessing techniques for detection, validation, error correction, and filling up of missing or incorrect data.
- Calculation of various statistical parameters such as averages, quantiles, performance metrics, probability distributions, and so on. For example, the averages may include average length of call, average number of calls per month and average delays in bill payment.
- Models and probability distributions of various business activities either in terms of various parameters or probability distributions.
- Computing user profiles.
- Time-series analysis of time-dependent data.
- Clustering and classification to find patterns and associations among groups of data.
- Matching algorithms to detect anomalies in the behavior of transactions or users as compared to previously known models and profiles. Techniques are also needed to eliminate false alarms, estimate risks, and predict future of current transactions or users.

[1] WikiPedia

Some forensic accountants specialize in forensic analytics which is the procurement and analysis of electronic data to reconstruct, detect, or otherwise support a claim of financial fraud. The main steps in forensic analytics are (a) data collection, (b) data preparation, (c) data analysis, and (d) reporting. For example, forensic analytics may be used to review an employee's purchasing card activity to assess whether any of the purchases were diverted or divertible for personal use. Forensic analytics might be used to review the invoicing activity for a vendor to identify fictitious vendors, and these techniques might also be used by a franchisor to detect fraudulent or erroneous sales reports by the franchisee in a franchising environment.

Fraud management is a knowledge-intensive activity. The main AI techniques used for fraud management include:

- Data mining to classify, cluster, and segment the data and automatically find associations and rules in the data that may signify interesting patterns, including those related to fraud.
- Expert systems to encode expertise for detecting fraud in the form of rules.
- Pattern recognition to detect approximate classes, clusters, or patterns of suspicious behavior either automatically (unsupervised) or to match given inputs.
- Machine learning techniques to automatically identify characteristics of fraud.
- Neural networks that can learn suspicious patterns from samples and used later to detect them.
- Other techniques such as link analysis, Bayesian networks, decision theory, land sequence matching are also used for fraud detection.

1.2. VOIP FRAUD

Voice over Internet protocol (VoIP) services transmit telephone calls over high-speed Internet connections rather than over traditional land-based telephone lines.[2] They do not usually travel directly from a caller to a recipient's computer but rather through computers belonging to several layers of intermediary VoIP service providers, or wholesalers.

[3]VoIP fraud is considered to be the use of a VoIP telecommunications network with the intention of avoiding payment. In that sense, the payment may be incorrect, entirely lacking, or the attempt to force another party to pay. We will consider both illegal activities and those activities which, though technically legal, may still hurt telecommunications companies by taking advantage of systems and vulnerabilities.

VoIP fraud can affect any organization which uses or sells VoIP services. In most cases, the fraud target is an enterprise. Most enterprises never realize that they have been hacked, refuse to pay the fraudulent charges and threaten to switch to a different service provider. The SIP service provider has little leverage over its international long distance vendors and is left to cover the bill.

However, in some cases, service providers will demand the enterprise pay for fraudulent charges. This was the case in a 2009 when Michael Smith, a small business owner in Massachusetts, found that someone had hacked into his PBX to make $900,000 worth of calls

[2] TransNexus

[3] The State of Phone Fraud: 1H 2012.

to Somalia. AT&T attempted to sue Smith for $1.15 million to recoup the cost of the calls and interest. Though AT&T eventually dropped the charges, a spokeswoman for the company maintained that they had been entitled by law to collect the amounts owed, and that Smith should have put more safeguards in place to protect his phone system.[4]

VoIP fraud can and does occur in any industry. Certain industries, such as banking, tend to attract more fraud than others. A recent study from Pindrop Security found that nine out of the top ten banks, and 34 of the top 50 banks had been victims of call fraud.[5]

VoIP fraud comes from all over the globe. Traditionally, Africa has been a "Hot Continent" from telecom fraud, because the termination costs are very high and regulation is not as stringent as in other parts of the world. However, a 2011 study from the Communications Fraud Control Association (CFCA) found that the top 5 countries from which fraud originates are the United States, India, the United Kingdom, Pakistan, and the Philippines. The top five fraud terminating countries were Cuba, Somalia, Sierra Leone, Zimbabwe, and Latvia.[6]

VoIP fraud is a significant and growing problem in the telecommunications industry. Because fraudsters often attack during weekends, fraud events often go undetected for many hours. A single fraud event can easily cost a company between three and fifty thousand dollars. In many cases, this number can be even larger. A 2009 attack on an Australian company's VoIP PBX resulted in 11,000 international calls in just 46 hours, leaving the SIP provider with a bill in excess of $120,000.[7] A 2011 weekend episode in South Africa resulted in a bill of over $12,000 and another in the US cost victims more than $1.4 million.[8]

Experts have trouble estimating an aggregated global yearly loss, because calculations are often based on subjective and individual standards. However, most experts agree that total loss is somewhere between 3 and 10 percent of income. This translates to a total global losses of somewhere between 30 and 50 billion dollars per year.[9] The CFCA's 2011 report put the number at $40.1 billion dollars lost.

This is a problem that is only increasing. According to the CFCA report, phone fraud is growing at a rate of 29% per year. As the popularity of VoIP continues to grow, the problem of VoIP fraud will become an increasing threat to the industry.

1.3. TYPES OF VOIP FRAUD

Fraudsters have come up with a myriad of techniques for exploiting the VoIP industry. Some are stolen from traditional phone fraud tactics, others from computer hacking techniques, and still others exploit VoIP specific equipment and software. These are just a few examples of the many types of VoIP fraud and fraudsters are constantly coming up with new ways to attack. In addition, most fraudsters will use some sort of combination of the techniques listed below.

---

[4] AT&T Agrees to Drop $1.15M Suit against Phone Hacking Victim

[5] The State of Phone Fraud: 1H 2012.

[6] CFCA

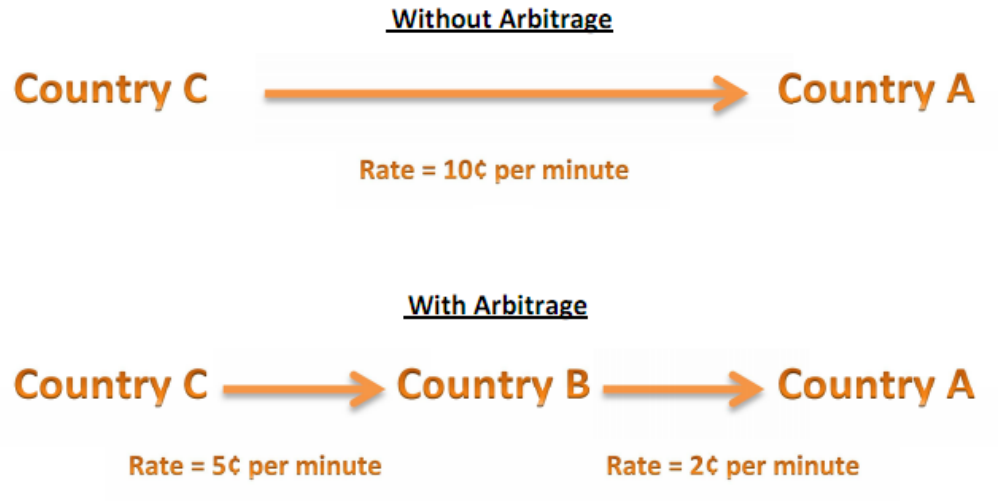
[7] Winterford

[8] "VoIP Hacking on the Increase."

[9] The Importance of Quantifying Fraud.

Arbitrage

Arbitrage is simply the exploitation of the differences in settlement rates between countries, the complexity of services and rates, and the multi-operator environment of VoIP. For example, if country A has much lower settlement rates with country B than with country C, it might be cheaper for country C to send its traffic for country A via country B. One of the first larger arbitrage routes was for traffic between Australia and the US, which was cheaper if sent via New Zealand and Canada.



Pic 1. Arbitrage

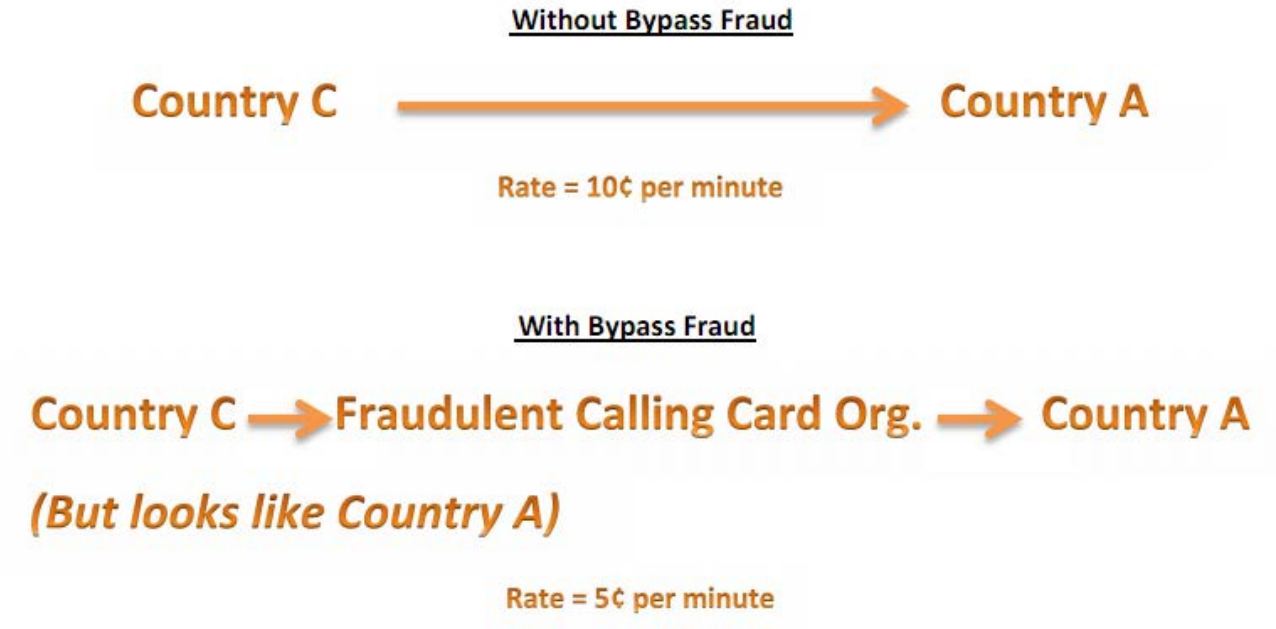
This is an example of a type of fraud that falls into a legal gray area. Though it may seem relatively tame, arbitrage can create problems for telecom companies. Unscrupulous providers may artificially inflate traffic where arbitrage is profitable. Telcos often withhold payments to Service Providers if this type of activity is suspected.

Buffer Overflow

Some VoIP fraud relies on methods typically used for computer fraud. In this case, fraudsters use buffer overflow errors in handling INVITE or SIP (Session Initiation Protocol) packets. The flaw might be used to crash applications or run arbitrary code. This is an issue that has was a particular problem for Asterisk users in the past, but has since been remedied.

Bypass Fraud

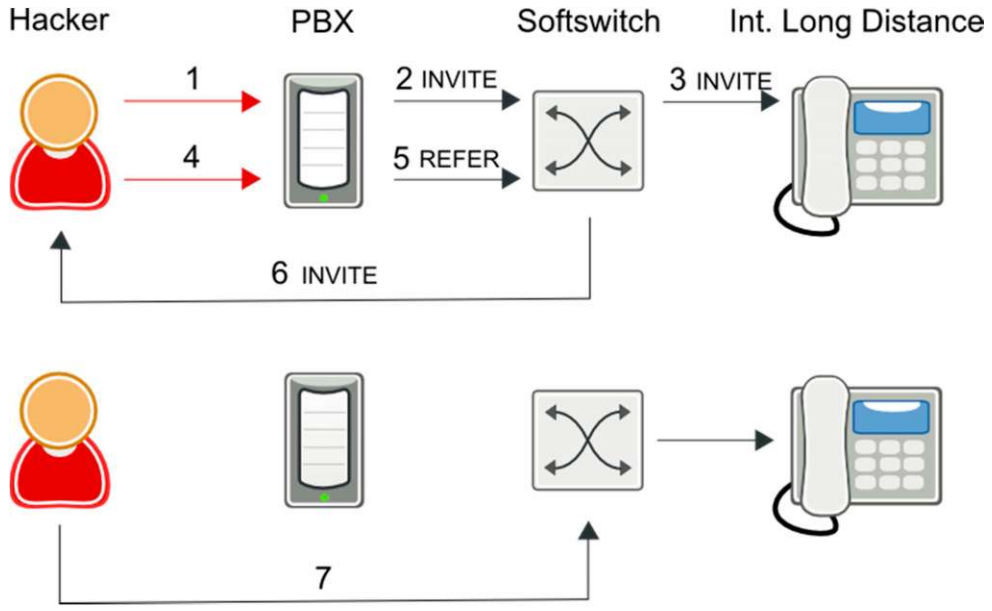
Bypass fraud is the unauthorized insertion of traffic onto another carrier's network. You may also find this type of fraud referred to as Interconnect fraud, GSM Gateway fraud, or SIM Boxing. This scenario requires that the fraudsters have access to advanced technology, which is capable of making international calls appear to be cheaper, domestic calls, effectively "bypassing" the normal payment system for international calling. The fraudsters will typically sell long distance calling cards overseas. When customers call the number on the cards, operators are able to switch the call to make it seem like a domestic call.



Pic 2. Bypass fraud

Call Transfer Fraud

One type of fraud that has been a particular issue for softswitch users is call transfer fraud. In this scenario, the fraudster hacks into a PBX and uses that PBX's services to make free long distance calls. By instructing the compromised PBX to transfer the call to the hacker's own phone service, subscribers to the fraudster's phone service can speak to their international destinations through the hacked soft switch, and the soft switch operator cannot bill the hacker's subscriber. Those familiar with three way calling will recognize the inspiration for this type of fraud. See the chart below for a detailed explanation.



Pic 3. Call Transfer Fraud

1. Hacker phone service hacks unsuspecting PBX to make a call to make international calls
2. PBX sends SIP INVITE to soft switch
3. Softswitch routes call to international carrier
4. Hacker instructs PBX to blind transfer call to Hacker Phone Service
5. PBX sends SIP REFER to soft switch to blind transfer call to Hacker Phone Service
6. Softswitch sends SIP INVITE to Hacker Phone Service
7. Hacker's Subscriber speaks to international destination through soft switch.

Most soft switches has no way of tracking a call once it is transferred out of the network, so fraudsters can generate a significant amount of traffic and revenue for themselves before being caught.

Domestic & International Revenue Share Fraud

Revenue share fraudulent activities are those which abuse carrier interconnect agreements. Cooperation is the key to this type of fraud. The fraudster's goal is to pair up with a destination that can charge high rates, and then inflate traffic to his numbers at little or no cost to himself. These types of schemes can occur within a country, or across international borders. Though they may not be technically illegal, they are often also paired with PBX hacking or other forms of fraud that generate illegal and artificial traffic.

Fraud rings in the Philippines and other developing countries have been known to have human "labor pools" that perform password guessing and make manual calls originating from hacked PBXs. Sometimes the "labor pool" can even be generated by scams on social media, SPAM mail, or post cards that advertise free vacations or important information just by calling a phone number.

Traffic Pumping or Switch Access Stimulation

Because of telephone regulations, long distance carriers must pay access fees to local exchange carriers for calls to those carriers' local subscribers. Rural carriers may charge substantially higher access fees than urban carriers. In order to increase their incoming call volume and revenue, some rural carriers partner with telephone service providers to route their calls through the rural carrier. These often include phone sex or free conference call providers, who expect a high volume of incoming calls. A similar scenario occurs internationally, with fraudsters making setting up conference servers in third world countries and making deals with the local (often state owned) telephone company.

To qualify as access stimulation, a fraudster must:

1. Have a revenue share agreement between the terminating carrier that stimulates demand
2. Have a 3 to 1 increase in interstate terminating to originating traffic or 100% traffic growth in a month year over year.

In this case, authorities can step in and force the terminating carrier to re-file their access tariff with the public utilities commission within 45 days. Because of the extra traffic, the

terminating carrier would no longer be eligible for the high access fee, and will likely be dropped from the revenue sharing agreement.

CNAM Revenue pumping or Dip pumping

Every call to a U.S. telephone number that has Caller ID (CID) enabled requires that the terminating phone company perform a lookup in one of several national databases that contain all the U.S. subscriber names and numbers. This database lookup is called a CNAM dip. When this database is being 'dipped' the originating local phone company gets compensated by the terminating phone company - this is commonly referred to as a CNAM dip fee (or simply a dip fee). This compensation happens for every call where the calling party name is displayed to the called party - even if the call is not answered.

CNAM Revenue Pumping, or DIP Pumping is similar to regular traffic pumping. The fraudsters again, partner with the parties who typically receive CNAM revenue, and generate traffic to conference servers with those numbers, sharing the revenue created. This traffic is generated wither by high volume call services (such as free conference calls and phone sex services) or by hacking into vulnerable PBX systems.

Premium Rate Services

Premium-rate telephone numbers are telephone numbers for telephone calls during which certain services are provided, and for which prices higher than normal are charged. Unlike a

normal call, part of the call charge is paid to the service provider, thus enabling businesses to be funded via the calls. Like the other revenue sharing schemes, fraudsters will pair up with a premium rate service to generate traffic to the number, via legal or illegal methods.

False Answer Supervision

When a dialed phone number is not in service, the calling party will hear a brief recording telling them so. There is no answer supervision or connection between the calling and called party. Since the call never connects, it is an incomplete call and should not be billed. However, fraudsters use false answer supervision to make these calls appear as completed calls which may be billed. Perhaps the fraudster has published rates for terminating calls without any intention of actually completing the calls. Here, service providers will route calls through the fraudster, who, instead of terminating the call, will play a not in service message and then bill the service provider for more than 10 seconds of calling. This type of fraud hurts the originating service provider both by costing money, and by hurting their reputation. The key indicators of this type of fraud are:

- Short phone calls
- Calling party hangs up nearly 100% of the time
- High answer seizure ratio
-

Location Routing Number (LRN) Fraud

Location Routing Number Fraud or LRN fraud works based on the desire of some service providers to avoid extra charges from LRN "dips." Most providers will run an LRN dip to

determine the correct LRN for a dialed number. However, many service providers will not perform an LRN dip if the LRN is already in the SIP message. Fraudsters take advantage of this by inserting fake LRNs into their calls. For example, they may insert the LRN for a relatively cheap terminating destination, when the call is actually going to a high cost rural destination. The service provider will then bill the fraudster for the cheaper call, but will have to eat the cost of the expensive rural call. In some cases, this can be up to 5x the price they billed the fraudster.

PBX Hacking

One of the most typical and certainly one of the most insidious flavors or VoIP fraud involves hacking into a Private Branch Exchange (PBX). Fraudsters who can exploit the vulnerabilities of the IP PBX are able to generate a significant amount of traffic. PBX hacking is the common technique used to perpetrate the Domestic and International Revenue Share Fraud and Call Transfer Fraud schemes listed above.

PBXs can be compromised in-band (over the phone circuit) by someone guessing the numeric pin on an extension or by finding an extension with a default pin. There are four common configuration mistakes in VoIP security architecture that can increase the risk of PBX hacking:[10]

1. Weak user authentication and access control
2. Relying solely on Session Border Controllers (SBCs) to provide security
3. Inadequate virtual LAN separation and control
4. Inadequate use of encryption

Phreaking

Phreaking is simply unauthorized access and control of a phone. Phreaked PBXs may be networked together to create a fraudulent telecom network for providing services. Customers of phreaked services often inclue call shops, call centers, calls sold through the internet, and traffic pumping schemes.

Roaming Fraud

Roaming fraud is the use of a wireless operator's services outside of the user's "home country" where there is no intention of paying for calls made. Typically, this type of fraud is committed by someone who is also using subscription fraud (obtaining a service subscription using a false identity). Because of the delay in the transference of call detail records between roaming partners, it can take days or even weeks for the home network to receive the call records and notice the fraud.

SIPVicious

Though its name suggests otherwise, the SIPVicious program is a mainstream auditing too for VoIP systems. Unfortunately, fraudsters have also found a way to use it for attacks. The attacks are apparently aimed at taking control of VoIP servers to place unauthorized calls. The attacks use a known Trojan, jqs.exe, and connect to command and control servers to

[10] Boone

receive instructions on downloading instructions as well as the SIPVicious tool from a .cc domain. After installation, SIPVicious is run and scans for SIP devices on the compromised computer's network and launches brute force attacks to guess the administrative password on those systems. This creates a base from which attackers can make VoIP calls from the victim's phone or VoIP infrastructure. Those calls might be used to rack up charges on premium rate numbers controlled by the attackers, or as part of voice phishing (vishing) scams that target unwitting consumers.

Shell Companies

Recent VoIP scams have involved the creation of "shell companies." These companies have no operations, and their sole purpose is to induce companies that sell VoIP services to extend service on credit to the shell companies. When the victim providers sell VoIP services to the shell companies on credit, the conspirators "bust out" the accounts by causing the companies to use substantially more VoIP services than the companies had been approved to buy in such a short period of time. Fraudsters typically do this over weekends and holidays so that the providers do not notice. When the invoices for the services came due, the fraudsters send fake wire transfer confirmations via e-mail or submit small payments to keep the victim providers from cutting off service.

Subscription Fraud

Subscription fraud is simple the use of service with no intent to pay. Often this type of fraud is associated with other crimes, such as identity theft. The true impact of subscription fraud often goes unrecognized because providers mistake it for bad debt. The key difference between the two is intention. Fraudsters gain access to a service in one of three ways:

1. Giving valid details, but disappearing without paying the bill
2. Using false details
3. Using the identity of another person (identity theft)

Once recognized as a 'bona fide' customer, fraudsters then have access to a network and are able to carry out revenue generating schemes that can seriously damage reputation and bottom-line profits.

Toll Fraud

Toll fraud is a scheme where fraudsters break into a company's VoIP network and sell long distance minutes. Fraudsters hack companies' PBXs, and then control the hacked PBXs to make long distance calls. To make money, the fraudsters will sell long distance calls, and then use the hacked PBXs to complete the calls.

Unallocated Number Fraud

Some fraudsters simply add unallocated phone numbers to their rate deck. They can then simulate traffic to this number, which will generally be routed to the fraudulent company, as they are the only ones that will "terminate" the call, based on the rate decks. According to recent estimates, 32% of called numbers used for fraud are unallocated numbers.

1.4. TECHNIQUES FOR FIGHTING VOIP FRAUD

The most prevalent threats to VoIP deployments today are rooted in hacking of the underlying and supporting infrastructure. The major IP PBX vendors can be secured, but security has to be considered during deployments. It's important to consider your existing network security before adding VoIP components. A VoIP security assessment and penetration test will help identify vulnerabilities.

Analyzing Call Detail Records

The simplest way of tacking VoIP fraud is by analyzing Call Detail Records (CDRs). VoIP providers should put into place a system that detects the most common symptoms of fraud and PBX hacking, traffic spikes, unusual call patterns, etc. These systems work best when they are reported in near real time. In addition, there must be a system for monitoring this analysis, especially overnight, on weekends, and over holidays. The best analysis solutions will include alerts or will be integrated with your routing system to temporarily block suspicious calls.

Call Blocking

Some providers are taking an even more aggressive stance against fraud and completely blocking destination countries with reportedly high incidences of fraud, or simply blocking calls to fraud related phone numbers. Obviously, this method risks denying legitimate traffic, and can also become a time consuming project, keeping up with the latest news on countries and numbers that have been reported.

Call Routing

Other service providers have chosen to route fraud countries only through operators who perform a validation on the call before completing it. Though this method may cut down on fraud, it is also time consuming and not cost effective.

Legal Action

Though it is tempting to rely on legal action as a solution to VoIP fraud, it is not the ideal option. Recent studies have shown that companies choose not to report fraud cases to law enforcement due to a perceived lack of interest and understand from authorities. Though 89.1% of communications companies refer at least one case to law enforcement per year, few fraudsters are ultimately caught and punished.[11]

That being said, there have been recent arrests over VoIP fraud. In May 2012, two men who stole more than $4.4 million from VoIP providers were sentenced to three years in prison and will be required to pay back the millions of dollars.[12] In another recent case in Australia, fraudsters were sentenced to 40 years in federal prison and owe $18 million in restitution.[13]

---

[11] 2011 Global Fraud Loss Survey

[12] U.S. Attorney's Office. District of New Jersey

[13] Wilonsky

Session Border Controllers

Session Border Controllers, which work in front of an IP connected PBX can be a deterrent to fraud. The SBC can detect and stop attempts to guess user credentials or unauthorized attempts to route traffic. Some SBCs can analyze call patters and dynamically learn your normal traffic patterns, sending warnings when something deviates from the norm. Session Border Controller cannot, however, detect Traffic Pumping attacks.

1.5. FRAUD DETECTION SOLUTIONS

Alaris Anti Fraud System

The principle is that each network node that is able to provide relevant data to the fraud management system is set to stream this data to the servers – by means of OS native utilities (like rsyslog) or with the help of the agent – that is installed on the monitored nodes[14]. The data does not need to be transformed into any unified form, it can be in the original raw format.

The servers parse the data, index it, apply search algorithms to detect potential fraud and send notifications to the client personnel in case of fraud detection events.

The search algorithms consist of the following principle stages:

- identification of most common behavior templates (of an end-user, of a partner, of a destination of the calls, etc.)
- correction of the detected profiles by the end user
- detection of non-standard behavior of objects (or on the contrary – detection of the behavioral template of a fraud agent)
- alarming the client on the detected fraud events

One of the most important things to mention is that the search algorithm is able to combine events that are not linked to each other from the human point of view. For example: user self-registration date, location, credit card type can be linked with the calls pattern generated by this user, with the type of equipment used to generate calls, with the time frame during which the calls are made.

Search Formula Sample

```
sourcetype=CDR* type=CLI | stats count AS Views, count(eval(action="purchase")) as Purchases | eval percentage=round(100-(Purchases/Views*100)) | count percentage AS "% Difference"
```

Such semi-self learning approach can be used to automatically detect new fraud events.

For situations when it is unknown what kind of fraud the network awaits – the system acts the opposite way: all existing traffic is categorized and structured. Any behavior that deviates from the known non-fraud patterns is marked as potentially fraudulent. The end user then

[14] Alaris Anti Fraud System

decides whether a particular event is fraud or not. Based of that the system automatically creates a “fraud” template and all new data is checked against this pattern as well.

TransNexus System

TransNexus has developed a number of solutions to detect fraud in VoIP networks. [15]NexOSS, in addition to its already industry-leading least cost routing features, effectively eliminates the problem of traffic pumping fraud for VoIP providers. The solution is to include smart monitoring features that sense when there is an unusual spike in call traffic to a specific destination. When a suspicious spike occurs, the NexOSS system simply and automatically puts a temporary block on the route, ensuring that fraud losses are kept to an absolute minimum without interrupting legitimate calls.

TransNexus solutions analyze CDRs or RADIUS records, and can identify fraud by IP address, or by group or user id.

2. PRACTICAL PART

2.1. JUSTIFICATION OF CREATING THE SYSTEM’S EXTENSION

VoIP is about convergence, saving money and resources. However, these types of systems also create more inroads for attack. As VoIP has become more accessible and popular, security threats have grown as well. The most prevalent threats to today's VoIP deployments are rooted in traditional data networking and PSTN (public switched telephone network) attacks.

Today, VoIP devices are the primary tools used by fraudsters and 46% of fraudulent calls were made from VoIP phones. That is why creating systems for scanning billing data and detecting fraudulent traffic are on a priority basis.

2.2. MINIMAL REQUIREMENTS FOR HARDWARE AND USED SOFTWARE

Here are the minimal requirements for software are shown:

- CPU 1GHz
- RAM 512Mb
- OC Linux Red Hat Enterprise 3 update 4, OC Windows XP or higher,

Recommended:

- CPU 2x Xeon 3000
- RAM 1024M
- OC Linux Red Hat Enterprise 3 update 4, OC Windows 7

[15] TransNexus

For creating the system following components have been used: XAMPP v3.2.1 which includes Apache and MySQL modules, Dev-PHP v2.6.1 which includes module PHP v5.2.1 SQLyog Community – MySQL GUI v12.09 for creating MySQL databases, Google Chrome v42.0.2 web browser and Bootstrap framework. Here are justifications of choosing particular software are provided:

XAMPP

XAMPP is a free and open source cross-platform web server solution stack package, consisting mainly of the Apache HTTP Server, MySQL database, and interpreters for scripts written in the PHP and Perl programming languages. XAMPP also provides support for creating and manipulating databases in MySQL and SQLite among others.

Once XAMPP is installed, it is possible to treat a localhost like a remote host by connecting using an FTP client. Using a program like FileZilla has many advantages when installing a content management system (CMS) like Joomla or WordPress. It is also possible to connect to localhost via FTP with an HTML editor.

PHP

PHP is free software released under the PHP License[16]. It is a server-side scripting language designed for web development but also used as a general-purpose programming language. PHP code can be simply mixed with HTML code, or it can be used in combination with various templating engines and web frameworks. PHP code is usually processed by a PHP interpreter, which is usually implemented as a web server's native module or a Common Gateway Interface (CGI) executable. After the PHP code is interpreted and executed, the web server sends resulting output to its client, usually in form of a part of the generated web page; for example, PHP code can generate a web page's HTML code, an image, or some other data. PHP has also evolved to include a command-line interface (CLI) capability and can be used in standalone graphical applications. PHP includes various free and open-source libraries in its source distribution, or uses them in resulting PHP binary builds. PHP is fundamentally an Internet-aware system with built-in modules for accessing File Transfer Protocol (FTP) servers and many database servers, including PostgreSQL, MySQL, Microsoft SQL Server and SQLite (which is an embedded database), LDAP servers, and others. Numerous functions familiar to C programmers, such as those in the stdio family, are available in standard PHP builds. There are two primary ways for adding support for PHP to a web server – as a native web server module, or as a CGI executable. PHP has a direct module interface called Server Application Programming Interface (SAPI), which is supported by many web servers including Apache HTTP Server, Microsoft IIS, Netscape (now defunct) and iPlanet.

[16] PHP

MySQL

MySQL the world's second [17]most widely used relational database management system (RDBMS) and most widely used open-source RDBMS. MySQL is a popular choice of database for use in web applications, and is a central component of the widely used LAMP open source web application software stack (and other 'AMP' stacks). LAMP is an acronym for "Linux, Apache, MySQL, Perl/PHP/Python." Free-software-open source projects that require a full-featured database management system often use MySQL.. It is a relational database management system (RDBMS), and ships with no GUI tools to administer MySQL databases or manage data contained within the databases. Users may use the included command line tools, or use MySQL "front-ends", desktop software and web applications that create and manage MySQL databases, build database structures, back up data, inspect status, and work with data records. MySQL ships with many command line tools, from which the main interface is 'mysql' client. MySQL works on many system platforms, including AIX, BSDi, FreeBSD, HP-UX, eComStation, i5/OS, IRIX, Linux, OS X, Microsoft Windows and others. MySQL can also be run on cloud computing platforms such as Amazon EC2.

Apache HTTP Server

The Apache HTTP Server, colloquially called Apache is the world's most widely used web server software. This software is available for a wide variety of operating systems, including Windows, OS X, Linux, Unix, FreeBSD and others. Apache supports a variety of features, many implemented as compiled modules which extend the core functionality. These can range from server-side programming language support to authentication schemes. Some common language interfaces support Perl, Python, Tcl, and PHP. Popular authentication modules include mod_access, mod_auth, mod_digest, and mod_auth_digest, the successor to mod_digest. A sample of other features include Secure Sockets Layer and Transport Layer Security support (mod_ssl), a proxy module (mod_proxy), a URL rewriter (mod_rewrite), custom log files (mod_log_config), and filtering support (mod_include and mod_ext_filter). Virtual hosting allows one Apache installation to serve many different Web sites. For example, one machine with one Apache installation could simultaneously serve www.example.com, www.example.org, test47.test-server.example.edu, etc. Apache features configurable error messages, DBMS-based authentication databases, and content negotiation. It is also supported by several graphical user interfaces (GUIs). It supports password authentication and digital certificate authentication. Because the source code is freely available, anyone can adapt the server for specific needs, and there is a large public library of Apache add-ons.

Google Chrome

Google Chrome is a freeware web browser developed by Google. As of January 2015, StatCounter estimates that Google Chrome has a 51% worldwide usage share of web browsers as a desktop browser. It is also the most popular browser on all the other platforms it supports, mobile, tablets (except for the iPad where the Safari browser is preinstalled, and because of that popularity Safari is most popular on all tablets combined) or any combinations of platforms, such as mobile ("smartphones") plus tablets.

[17] MySQL

Bootstrap (front-end framework)

Bootstrap is a free and open-source collection of tools for creating websites and web applications. It contains HTML- and CSS-based design templates for typography, forms, buttons, navigation and other interface components, as well as optional JavaScript extensions. The bootstrap framework aims to ease web development. Bootstrap is a front end, that is an interface between the user and the server-side code which resides on the "back end" or server. And it is a web application framework, that is a software framework which is designed to support the development of dynamic websites and web applications. Bootstrap is compatible with the latest versions of the Google Chrome, Firefox, Internet Explorer, Opera, and Safari browsers, although some of these browsers are not supported on all platforms.Since version 2.0 it also supports responsive web design. This means the layout of web pages adjusts dynamically, taking into account the characteristics of the device used (desktop, tablet, mobile phone). Bootstrap provides a set of stylesheets that provide basic style definitions for all key HTML components. These provide a uniform, modern appearance for formatting text, tables and form elements.

2.3. USER MANUAL

The system consist of three main sections and couple of subsections.
1) Database's management
2) Manual section
2.1) Filter "Threshold"
2.2) Filter "Iteration"
3) Automatic section
3.1) Automatic filters management

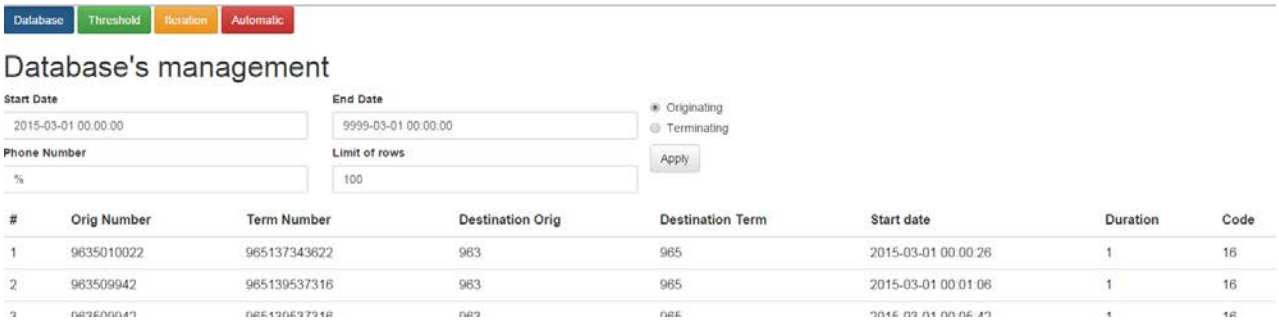
3.2) Detailed report

3.3) Automatic send out of the report to user's email address

Section "Database's management" belongs to the button "Database" and represents data from given database . Data can be filtered using parameters "Start Date", "Stop Date" and "Phone Number" for both originating (orig) and terminating (term) numbers. Changing "Limit of rows" can be set represented rows of data.

Important information:

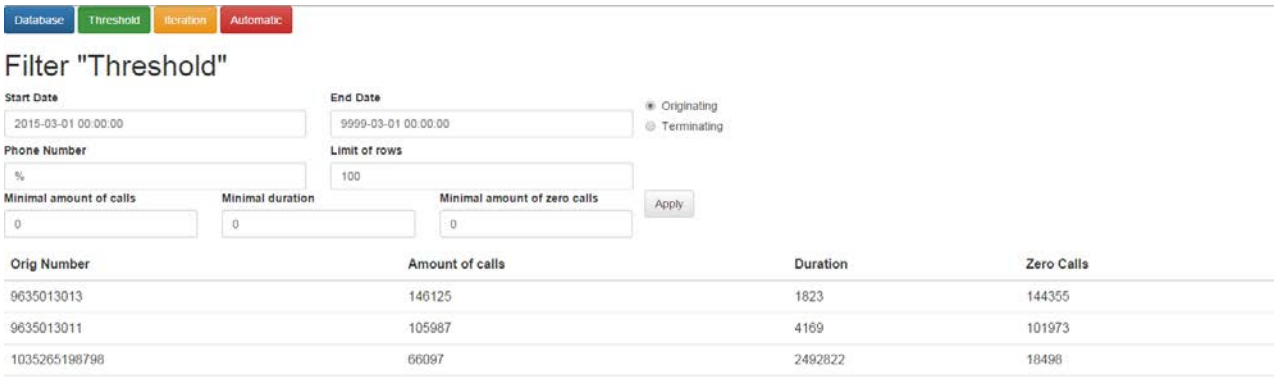
1) All fields of parameters should be always filled
2) Parameters "Start Date" and "Stop Date" should be always set in format YYYY-MM-DD HH:MM:SS
3) One of the parameters "Originating"/ "Terminating" should be always chosen if they are belong to particular section.



Pic 4. Database's management

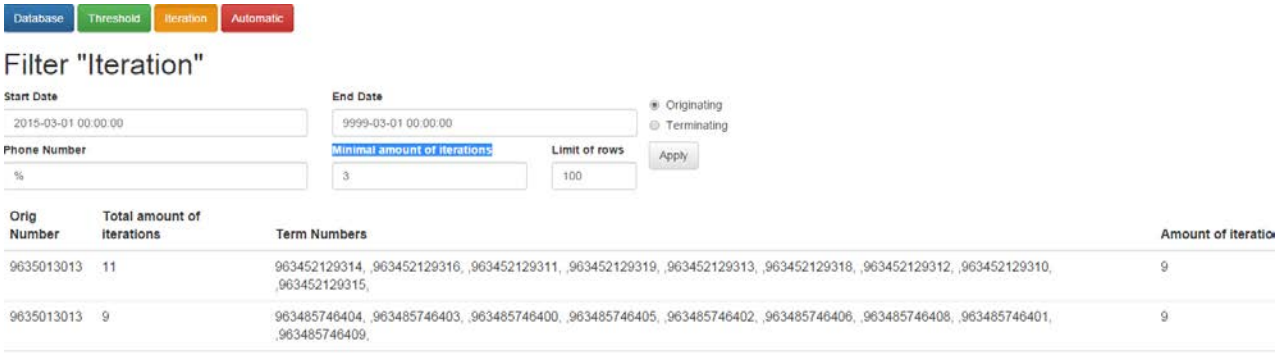
Manual section was created for manual checking data from the database using particular parameters.

Filter "Threshold" belongs to the button "Threshold" and can be used for filtering data from the database by parameters "Total amount of calls", "Total duration" and "Minimal amount of zero calls" for both originating (orig) and terminating (term) numbers.



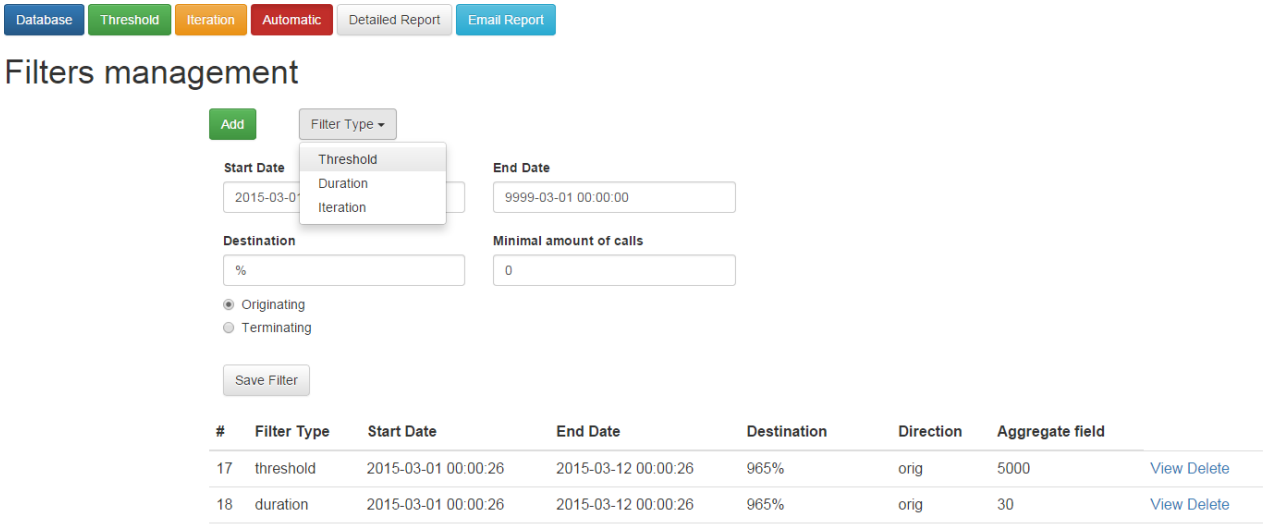
Pic 5. Filter "Threshold"

Filter "Iteration" belongs to the button "Iteration" and created for scanning database for specific events when caller tries to call to any numbers in series. In other words orig number "9635013013" calls to "963452129314", then to "963452129316", then to "963452129311" and this is a symptom of fraud. Parameter "Minimal amount of iterations" regulates amount for such numbers.



Pic 6. Filter "Iteration"

Section "Automatic" belongs to the button "Automatic" and allows user to create set of filters for automatic procedure of checking database according to the set filters and its parameters. Every filter in set has its own parameters for set destinations.



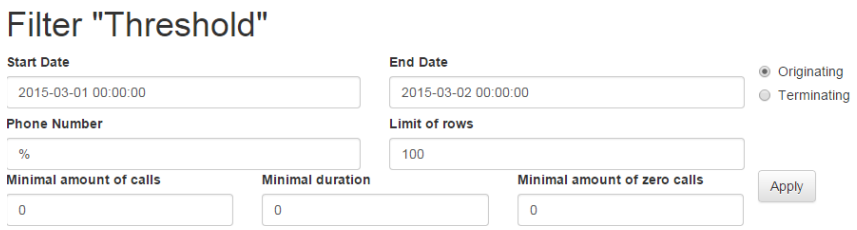
Pic 7. Section "Automatic"

There are 2 more buttons belong to section "Automatic" – button "Detailed report" which represents detailed report of filtered database according to set filters and button "Email Report" which sends report to user's email address.

Function "Email Report" was specially created for demonstrating the project and will be replaced by cronjob when script will be connected to real database in real-time regime. It will allow to check database accordingly with set of filters and send report to user's email every time when filters are actuated which means that fraudulent activity is detected.

2.4. ANALYZING GIVEN DATABASE AND DETECTING FRAUD

Analyzing given database using manual filter “Threshold” for period of 1 day below data is received:



Orig Number	Amount of calls	Duration	Zero Calls
963509923	189	26	163
963509932	28	21	7
1035265198798	19	707	11

Pic 8. Analysis using filter “Threshold” for period of 1 day

As it shown above the numbers 963509923 and 963509932 has big amount of calling attempts and more them 80% of them are zero calls. It is a symptom of fraud as well as wrong routing or non-working destination. In any case this information is important for the user and should be sent to him. In “Automatic” section for Syria destinations for period of one day we set filter “Threshold” and will get reports about such activity in future automatically.

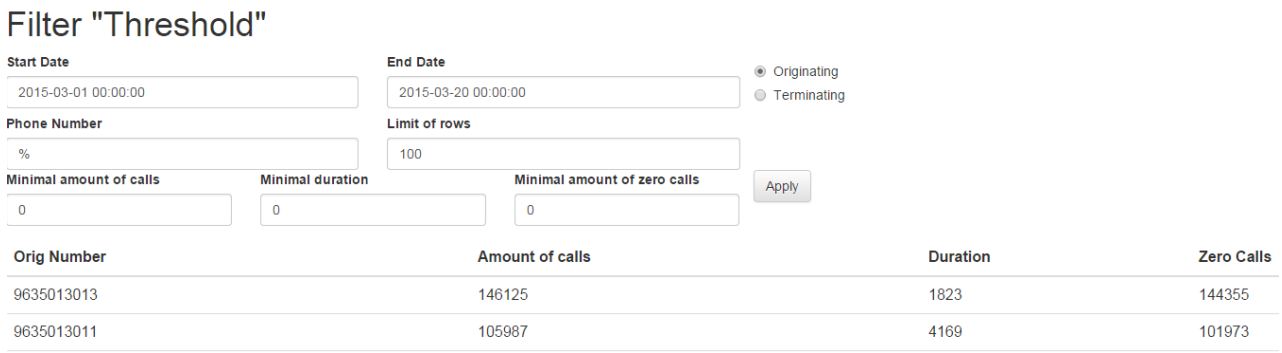
Next analyzing using manual filter “Threshold” for destination Syria for weekend period by parameter “Duration” below data is received:

Orig Number	Amount of calls	Duration	Zero Calls
963509923	431	67	364
963509942	71	66	5
963509932	69	62	7
963509924	22	7	15
963451783323	21	770	4
963456302134	21	517	9
963457286134	19	770	5
963456302853	18	1172	7

Pic 9. Analysis using filter “Threshold” for Syria for weekend

As it shown above callers from Syria are very active during weekend period. And it is a symptom of fraud and should be investigated by user. In “Automatic” section for Syria destinations for weekend period we set filter “Duration”.

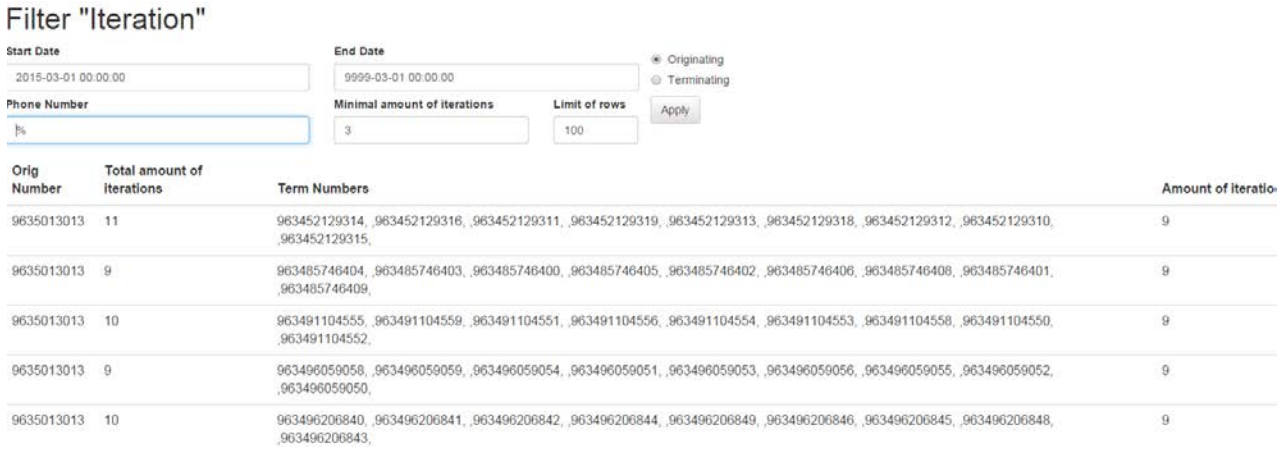
Next analyzing all destinations using manual filter “Threshold” for whole period below data is received:



Pic 10. Analysis using filter "Threshold" for whole period

As it shown above 2 numbers 9635013013 and 9635013011 has big amount of calling attempts and more them 80% of them are zero calls. Will set filter for these 2 numbers in section "Automatic".

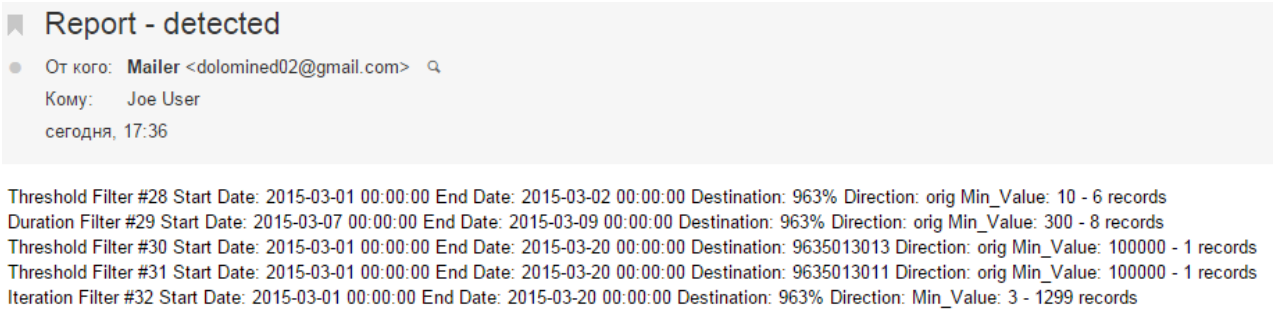
Next analyzing all destinations using manual filter "Iteration" for whole period below data is received:



Pic 11. Analysis using filter "Iteration" for whole period

As it shown above there were a lot of calls from Syria to numbers in series and it is fraud. Let us set filter "Iteration" for destination Syria in "Automatic" section to get to know if it again happen in future.

Using set of filters according to above parameters user will get following report on his email:



Pic 12. Automatic report about detected fraud

CONCLUSIONS

VoIP providers and enterprises must work together to ensure their networks are secure from every angle. By securing networks and analyzing traffic for signs of fraud, VoIP providers can minimize their fraud risks.

VoIP fraud is, and will remain, a lucrative criminal business. As VoIP continues to grow in popularity, schemes for beating the system will continue to become more complex and powerful.

To prevent any fraudulent activity current system extension has been created including possibility of filtering data taking into account all necessary parameters, creating detailed reports in accordance with set filters and possibility of automatic signaling to user about any fraudulent activity detecting. System has been tested using autonomous database but can be easily implemented in working billing system. Possibility of improving system extension for making it automatically change routings and block destinations according to results of automatic analysis was implemented as well.

REFERENCES

[1] Wikipedia, the free encyclopedia. Data analysis techniques for fraud detection <*http://en.wikipedia.org/wiki/Data_analysis_techniques_for_fraud_detection*>.

[2] TransNexus Solutions. Fraud Detection <*http://transnexus.com/solutions/fraud-detection/*>.

[3] The State of Phone Fraud: 1H 2012. Rep. Pindrop Security, 2 Aug. 2012. Web. 10 Sept. 2012. <*http://www.cfca.org/pdf/whitepapers/wp_pindrop_phonefraudreport.pdf*>.

[4] AT&T Agrees to Drop $1.15M Suit against Phone Hacking Victim. CFCA, 10 July 2012. Web. 10 Sept. 2012. <*http://www.cfca.org/fraudalert.php ?id=5232*>.

[5] The State of Phone Fraud: 1H 2012. Rep. Pindrop Security, 2 Aug. 2012. Web. 10 Sept. 2012. <*http://www.cfca.org/pdf/whitepapers/wp_pindrop_phonefraudreport.pdf*>.

[6] CFCA. Communication Fraud Control Association (CFCA) Announces Results of Worldwide Telecom Fraud Survey. CFCA. CFCA, 4 Oct. 2011. Web. 10 Sept. 2012. <*http://www.cfca.org/pdf/survey/CFCA2011GlobalFraudLossSurvey-pressrelease.pdf*>.

[7] Winterford, Brett. "Perth Firms Phreakedby VoIP Hackers."ITnews. N.p., 12 Apr. 2011. Web. 10Sept. 2012. <*http://www.itnews.com.au/News/254255,perth-firms-phreaked-by-voip-hackers.aspx*>.

[8] "VoIP Hacking on the Increase." My Broadband. N.p., 9 May 2011. Web. 10 Sept.2012. <*http://mybroadband.co.za/news/telecoms/20140-voip-hacking-on-the-increase.html*>.

[9] AT&T Agrees to Drop $1.15M Suit against Phone Hacking Victim. CFCA, 10 July 2012. Web. 10 Sept. 2012. <*http://www.cfca.org/fraudalert.php ?id=5232*>.

[10] Boone, Adam. "Toll Fraud Is Alive and Well." Computerworld. N.p., 2 Oct. 2009. Web. 10 Sept. 2012. <*http://www.computerworld.com/s/article/9138803/Toll_fraud_is_alive_and_well*>.

[11] 2011 Global Fraud Loss Survey. Rep. CFCA, 4 Oct. 2011. Web. 10 Sept. 2012. <*http://www.cfca.org/pdf/survey/CFCAGlobalFraudLossSurvey2011.pdf*>.

[12] U.S. Attorney's Office. District of New Jersey. Two Fraudulent Telephone Service Wholesalers Sentenced to Prison for $4.4 Million VoIP Fraud Scheme. FBI. N.p., 15 May 2012. Web. 10 Sept. 2012. <*http://www.fbi.gov/newark/press-releases/2012/two-fraudulent-telephone-service-wholesalers- sentenced-to-prison-for-4.4-million-voip-fraud-scheme*>.

[13] Wilonsky, Robert. "Three Ringleaders of Local "cybercrime Conspiracy" given Long Prison Sentences, Forced to Pay Back Many Millions." Web log post. Crime Blog. Dallas News, 25 May 2012. Web. 10 Sept. 2012.

<*http://crimeblog.dallasnews.com/2012/05/three-ringleaders-of-local-cybercrime-conspiracy-given-long-prison-sentences-forced-to-pay-back-many-millions.html*>.

[14] Alaris Solutions. Alaris Anti Fraud System

<*http://alarislabs.com/solutions/anti-fraud/*>.

[15] TransNexus Solutions. Fraud Detection

<*http://transnexus.com/solutions/fraud-detection/*>.

[16] PHP

<*https://en.wikipedia.org/wiki/PHP* >.

[17] MySQL

<*https://en.wikipedia.org/wiki/MySQL* >.