Fraud Detection in International Calls Using Fuzzy Logic

Osama Mohamed Elrajubi, Hussein Marah, Abdulla A. Abouda

Abstract

Telecommunications fraud is a problem that affects operators and telecommunication companies all around the world and one of the most known frauds is bypass fraud which is used in International call, in order to avoid access charges and this causes severe losses to operators and this is the type taken up in our research.

In this research a new fraud detection system has been proposed. The proposed system is depends on user profiling and using fuzzy logic for make the decision. A real database for a telecommunication company (Almadar Company for Telecommunication, Libya) was studied and five characteristics (fields) are used as detection patterns in the proposed system. The five characteristics are Subscribers Mobility, IN/OUT ratio, Cell Users density, Irregular calls, and Using voice service only. The proposed system was designed and implemented by language C++, Microsoft Visual Studio .NET. As well as, we have tried to use Fuzzy Logic Toolbox of MatLAB for designing the membership functions and defining fuzzy rules in the system.

Key Words: Bypass fraud ،user profiling, fuzzy logic, fraud detection, International / Mobile Calls, SIM Box Detector (SBD).

1. Introduction

Generally, the problem of the fraudulent use of mobile phones is a common thing to communication service providers. By definition, fraud in communication networks can be defined as the illegal access to the network and the use of its services with no intention to pay service charges or making money by using these services, while fraud detection is referred to as a try to detect illegal usage of a communication network.

In addition to financial losses, fraud may cause danger, loss of service, loss of customer confidence, hurting reputation of network operators and may threaten national security of the country. So, user profiling and understanding the behavior of customers is necessary to developing effective fraud detection model. [1][3]

1.1 Motivation

Following the huge growth in telecommunication networks in the last years, that becomes important part in our lives, the service providers face a new challenge, fraud. That spreading so fast with millions of dollars profits around the world. The fraud with all different types and specially bypass fraud (SIM boxes) causes great losses in total revenue for the operators and can be a threat on national security of the affected country. So, operators know that should put limits for this problem and protect its customers and itself against these fraudsters.

1.2 Problem description

Telecommunication operators worldwide have lost a significant amount of revenue from interconnection bypass of international calls in mobile networks, mainly due to increased usage of GSM VoIP Gateways - often called SIM Boxes - and technology advances in GSM gateways and VoIP technologies. GSM VoIP Gateways are telecommunication devices that enable calls from fixed, mobile or Internet telephones to be routed through VoIP directly into a relevant GSM network to be like domestic call. And this leads to loss the international terminations revenue. [3]

1.3 Objectives of the work

The objective of this project is to building a new fraud detection system in international calls (design and implementation). The Fraud detection system must be efficient, and accurate in detection process.

First of all, in this work we will studying and understand types of fraud in telecommunication companies and focused on fraud in international calls, as well as understanding types of solutions which are used to face this problem.

The rest of this report is structured as follows: Section 2 presents Types of Frauds while section 3 focused on Bypass Fraud. The basic of Fuzzy Logic is defined in Section 4. Approached are used in Fraud detection systems are showed in section 5 whilst the design of the proposed system is presents in Section 6. A conclusion and a future Work are given in Section 7 and section 8.

2. Types of fraud

With the development of the telecommunication and technology, and the big size of telecom market that found very attractive to fraudsters, so traditional types of fraud has been replaced with more complex ways of frauds that was spread too fast in the world.

Fraud in Traditional Telecommunication Networks represent any type of frauds in telecommunication companies which can be done without using of VOIP (Voice Over IP) technique. As example of Fraud in Traditional Telecommunication Networks: Subscription fraud, which is the signing up for a service using fake or stolen identification, with no commitment of paying the bill. And there are many other types such as: Subscription fraud, Premium Rate Service (PRS) fraud, Internal fraud, Dealer fraud, Roaming fraud, Calling card fraud, PBX fraud, Fraud in PSTN and Bypass Fraud. This research focused on Bypass Fraud which is the most type of fraud is used in International Calls. Bypass Fraud is a type of fraud can be done by VOIP (Voice Over IP) technique.

3. 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.

Bypass Fraud is use several of least cost call termination techniques like SIM Boxes, Leakey (hacked) PBXs etc, to bypass the legal call interconnection and diverting international incoming calls to 'on' or 'off' network GSM/CDMA/Fixed calls through the use of VoIP or Satellite gateway, thus avoiding revenue for international calls termination which operators and government regulators are entitled to. See to Figure 1.

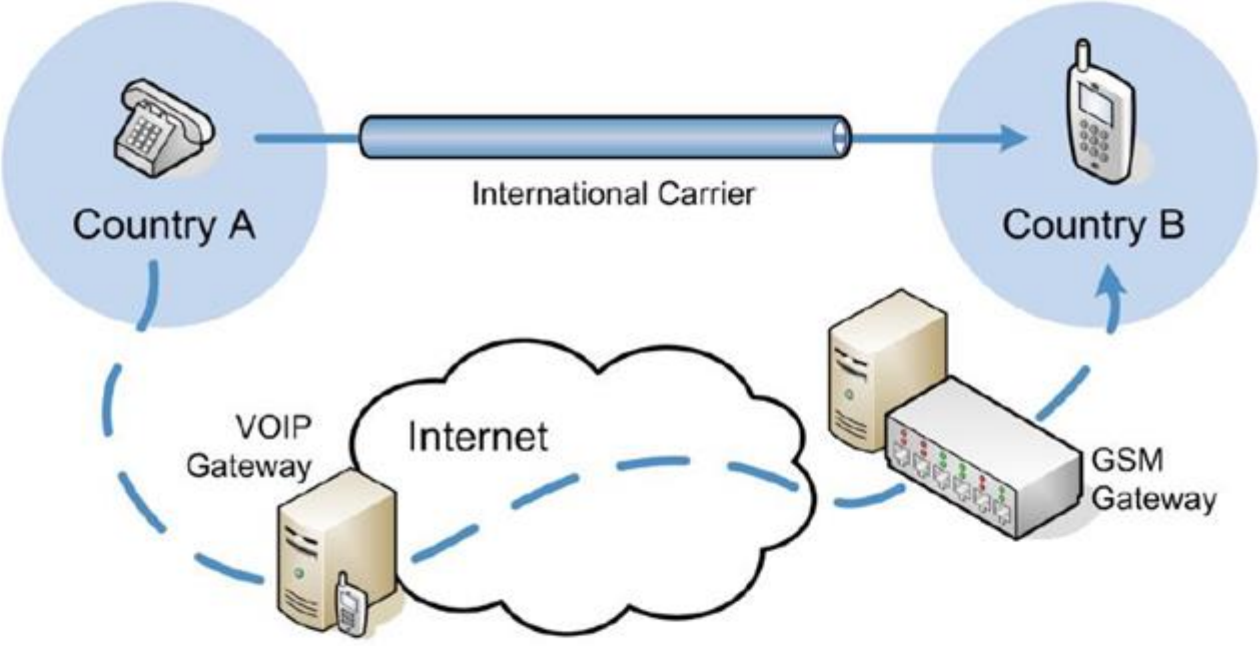


Figure 1: Basic Bypass Structure

The fraudsters manage Bypass fraud to earn through the international to national terminating charge margins of the calls.

- **On-net Bypass:** For an operator, if a fraudster uses its own network's connections to terminate the bypassed calls, it is defined as ON-NET bypass.
- **Off-net Bypass:** if the fraudster uses competitor's connections or any other network for termination, it is defined as OFF-NET Bypass.

As the ON-NET calls are expected to provide the least national calling rates, some modern Bypass equipment (SIM Boxes, computer programs etc.) scan the terminating party numbers and originate calls only from those connections which belong to the same operator's network as the terminating party for maximum profit.

But in the regions where even OFF-NET calls rates are equal with ON-NET, there can be national calls originated even from the OFF-NET connections to conduct Bypass fraud. [3][2]

4. Fuzzy Logic

In this research, fuzzy logic was studied and how we can used it to make the decision (determine which sim cards are used as for fraud). So, In this paragraph, the basics of fuzzy logic technique will be clarified.

It is important to know the difference between fuzzy logic and probability. Both operate over the same numeric range, and have similar values: 0.0 representing False (or non-membership), and 1.0 representing True (or full-membership). However, there is a difference between the two statements: The probabilistic approach yields the naturallanguage statement ”There is a 50% chance that x is low (for example in subset A)”, while the fuzzy terminology corresponds to ”x’s degree of membership within the subset A (low) is 0.50” [4].

"A fuzzy set A is a subset of the universe of discourse X that admits partial memberships. The fuzzy set A is defined as an ordered pair A = {x, μ(x)} , where x Є X, and 0<= μA(x)<= 1" [5].

The membership function μA(x) describes the degree to which the object x belongs to the set A. μA(x) = 0 represents no membership, and μA(x) = 1 represents full membership.

There are several types of membership functions such as Triangular membership function, Trapezoidal Membership Function, Gaussian Membership Function and Generalized Bell Membership Function. In our system in this research, Triangular membership function is used [5].

Fuzzy Rules

A fuzzy system is a collection of membership functions and rules that are used to reason about data [5]. A fuzzy rule can be defined as a conditional statement in the following form [6]:

IF x is A

THEN y is B

where x and y are linguistic variables; and A and B are linguistic values determined by fuzzy sets on the universe of discourses X and Y, respectively. In classical rule-based systems, if the rule antecedent (the condition) is true, then the consequent (the result) is also true. In fuzzy systems, where the condition is a fuzzy statement, all rules are corrected to a defined range, i.e they are partially correct. If the antecedent is true to some degree of membership, then the result is also true to same degree.

For example, rules in a fuzzy form can be represented as:

Rule: 1

IF "number of calls at night" is little

THEN The Result of test is "Sim card is legal"

Rule: 2

IF " number of calls at night " is much

THEN The Result of test is " Sim card is fraud "

In Rule:1 the condition is fuzzy. So the result of condition (Sim card is legal) is correct in a ratio which is equal to the membership of the variable " number of calls at night" in the set " little". Also In Rule:2 the condition is fuzzy, so the result of condition (Sim card is fraud) is correct in a ratio which is equal to the membership of the variable "number of calls at night" in the set "much".

The linguistic variable " number of calls at night" also has the range (the universe of discourse), but this range may include fuzzy sets, such as small, medium and large. The universe of discourse of the linguistic variable "Result of test" can include fuzzy sets such as very correct, correct and incorrect. Thus fuzzy rules relate to fuzzy sets. Fuzzy expert systems combine the rules and consequently cut the number of rules by at least 90 per cent [6].

5. Fraud detection systems

There are three common approaches using in fraud detection systems as following [1]:

1. Call generation analytics; By analyses the traffic which is actually received to determine if this international traffic is on-net traffic or traffic from another local network.

2. Call database analytics; By analyses call records for every sim card call of telecommunication company. The following Criteria can be used in this approach In Detection Process.

- Count and ratio comparison of calls for outgoing/incoming calls and off-net/on-net calls.
- Exclusion of calls to “allowed numbers” (e.g. customer service)
- Diversity of calls, including total diversity and on-net/off-net diversity
- Usage of non-voice services (SMS, GPRS sessions, etc.)
- Mobility, Use of only one cell site.
- Calls during irregular hours
- Suspicious cells, including automatic suspicious cell locator. [3]

3. Hybrid analysis; by using the both approaches, Call database analytics and call generation analytics, in order to develop more efficiently detection system. [1]

6. The proposed System

A real database for a telecommunication company (Almadar Company for Telecommunication, Libya) was obtained. Then, studying and understanding all fields in the database, this database contain 65 fields of CDR (CDR is a collection of data related to a call). After that, determining the important fields which can be useful in detection process.

In this research, the proposed system is using Call database analytics approach. As well as, the proposed system is used only 11 fields from 65 fields, to implementing a detection system based on five detection patterns. The five detection patterns are as following:

1. No or Low Mobility:

Mobility pattern is represent the number of calls for subscribers with no mobility.

2. Ratio of Incoming to Outgoing Calls:

Ratio of Incoming to Outgoing Calls is represent the number of outgoing calls for subscribers with no incoming calls.

3. Use voice-only service (NoUsage of SMS, MMS, GPRS, etc.):

Use voice-only service is represent the number of outgoing voice calls for subscribers with no non-voice service (SMS, MMS, GPRS sessions, etc.).

4. Suspicious Activity in Close Proximity (Hot Cell Site Calling):

Cells Suspicious Activity is represent the number of subscribers that they make voice calls within the same cell.

5. Calls during Irregular Hours, Unusual Night long Calls

This pattern is represent the number of subscribers who they had make irregular voice calls at night.

In our system, the triangular membership function is used for all detection patterns, by knowing the (Max, Min) values of these patterns, which extracted from database and calculating the membership value using this equation:

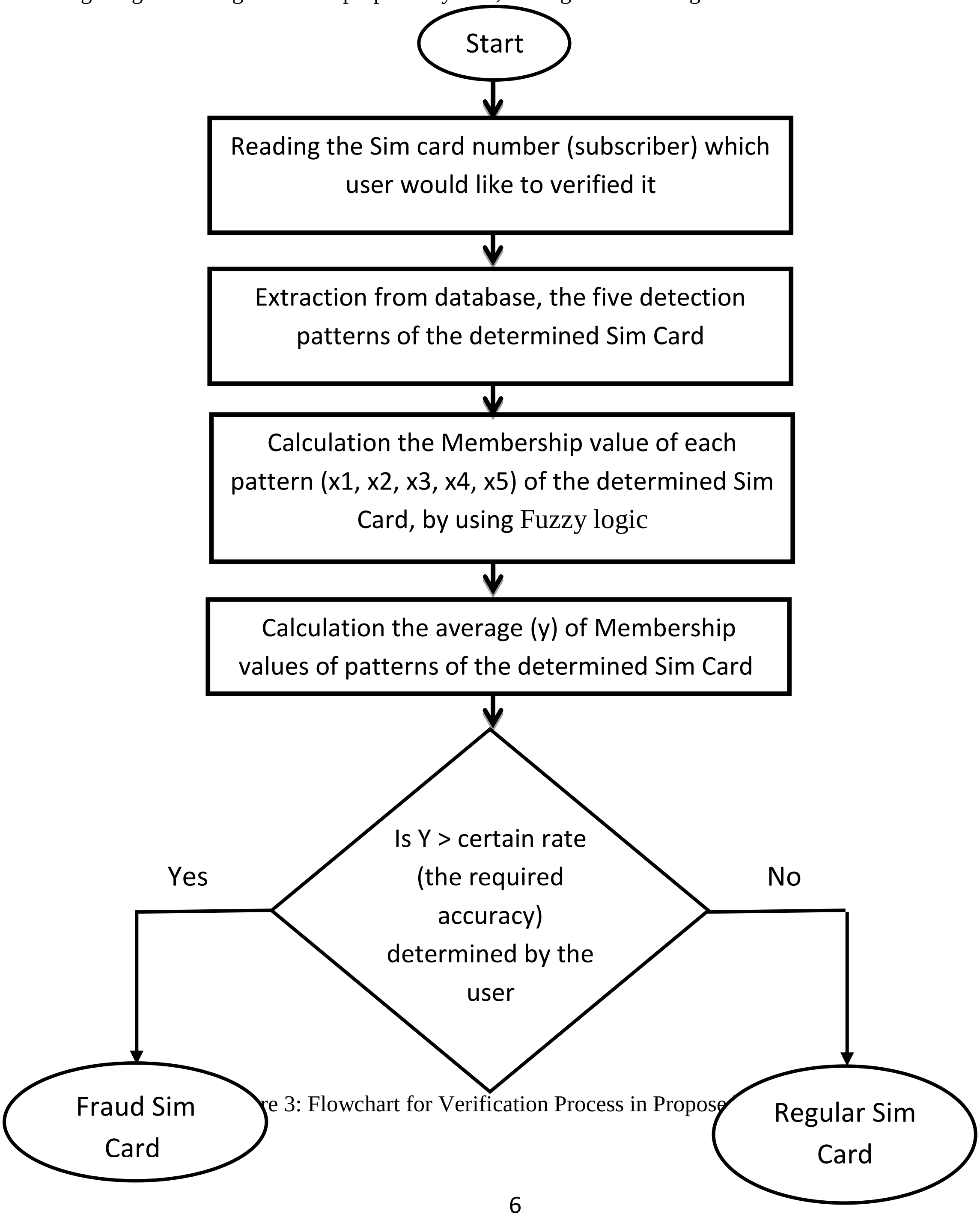
$$X = \frac{Value - Min}{Max - Min}$$

Where X is the membership value of detection pattern for certain subscriber, Min is the minimum value for certain detection pattern, Max is the maximum value for certain detection pattern.

And if some subscriber has all detection patterns with high membership function value by using the next equation, it can be considered as potential case of fraud.

$$y = \frac{x1 + x2 + x3 + x4 + x5}{5}$$

Where y is sum of MF for all detection patterns $(x1, x2, x3, x4, x5)$ divided by number of these patterns. Figure 2 showing the general diagram of the proposed system, and figure 3 showing the follow chart of verification process.


Start
Reading the Sim card number (subscriber) which user would like to verified it
Extraction from database, the five detection patterns of the determined Sim Card
Calculation the Membership value of each pattern (x1, x2, x3, x4, x5) of the determined Sim Card, by using Fuzzy logic
Calculation the average (y) of Membership values of patterns of the determined Sim Card
Is Y > certain rate (the required accuracy) determined by the user
Yes
No
Fraud Sim Card
Regular Sim Card

re 3: Flowchart for Verification Process in Propose

7. System Interface (GUI)

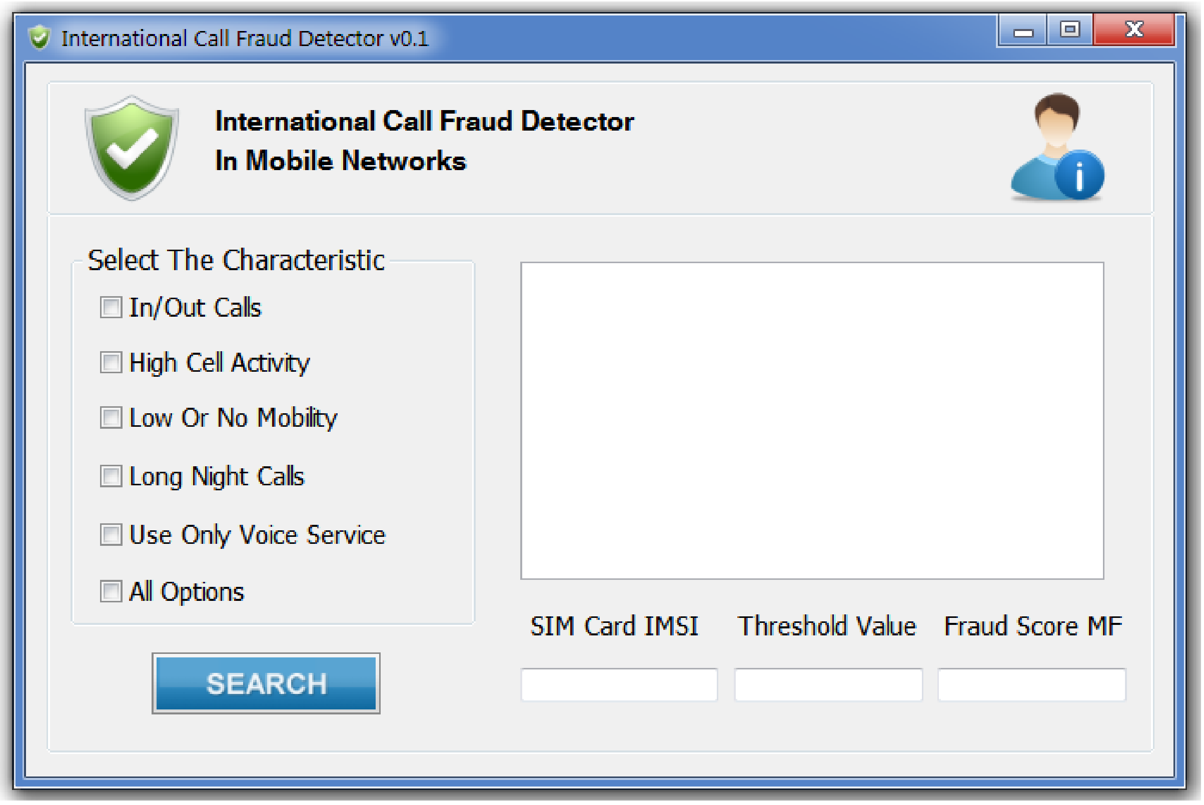


Figure (5-15) System Interface (GUI)

In this interface the user select one of these options (patterns) and click on the search button to find the SIMs that have this characteristics based on FM of this characteristic (pattern).as in the following two figures.

Conclusions

- The results of such system depend on database that will be used in the system.
- The results of fuzzy logic system depends on membership function (MF) for every detection pattern, the MF it can increased or decreased depending on the importance or effectiveness of the detection pattern.
- The results of every detection patterns or all patterns should be tested by the telecommunication company by using test call to confirm the fraud happening in this SIMs.
- In such systems is so important to getting documents or reports of data subscribers (SIMs) who had did fraud, and also normal subscribes, to study these data, compare between this data and figuring out how the fraudsters think and act to help us in the construction of the system and design detection process, to minimize the fault rate and maximize the detection rate.

Future work

- For improving the performance and accuracy of the system, it is proposed to use more detection patterns.

- The fraud detection system is so complicated and hard to deal with it easily, so its needs to a good trained work team in this field, as well the huge amount of data that needs to supercomputers with high specification to processing this data and dealing with it.

References

1. YacineRebahi, Jens Fiedler, FabricioGouveia. SCAMSTOP: Scams and Fraud Detection in Voice over IP Networks,FraunhoferFokus, INRIA, TEIMES, PDMFC, Telio, VozTelecom.8-13, 2010.
2. TransNexus. Introduction toVoIP Fraud. 2012 http://www.transnexus.com/index.php/whitepapers?id=430
3. Subex Inc. White Paper Bypass Fraud- Are you getting it right?., 2010http://www.subex.com/
4. Epsilon.Nought Radar Remote Sensing. http://epsilon.nought.de

(Last access, June, 2009).

5. Hussain Abuaishia, “Content-Based Image Indexing and Retrieval using Fuzzy Color Signature”, Unpublished master’s thesis, Academy of Graduate Studies, Tripoli - Libya, 2007.
6. Michelle Negnvitsky, Arabization: Sorour Ali Ibrahim Sorour, "Artificial Intelligence guide intelligent systems", Saudi Arabia, Al- for Publishing, 2004.