Ai.Fraud

Fraud Platform ▫ SimBox Detection ▫ & Usage threshold/black list monitoring

2012-06-11
Definitions

- Fraud prevention is the **process** of methodically **analyzing** customer **behavior** and systematically detecting any out-of-ordinary **patterns** that could be a potential **loss** to the operator revenue.

- In criminal law, a fraud is an **intentional deception** made for personal gain or to damage another entity.
Ai.Fraud Functionalities

General fraud detection
- Customer aggregation
- Counter collection and management
- Threshold management on 3 levels
- Alarm raising and resolving
- Alarm scoring
- Black listing
- Reporting on various aspects

SimBox detection
- SIM card data collection and aggregation
- Online & Offline rules
- Real-time and scheduled filtering
- Integration with MPS
- Rule based scoring
- Integration with CRM & CBS
- Reporting on agent success rates

Risk Scoring Legal Module
- Scoring of postpaids based on paid amount history
- Risk scoring of new postpaids (TBD)
- Legal module for police enquires

Forecasting on future behavior
- Next invoice value
- Next 30 days of spendings
- Next 30 days of roaming spendings
- Forecasts on 3 levels
General fraud detection

- Based on two types of convergent counters (spanning GSM, fix, internet & digital TV)
  - Daily counters – one profile for every day
  - Monthly counters – one profile for every month
- Counters cover spending, GSM & Internet data usage, on-net, national, international roaming traffic usage and spending, VAS service activations, monthly bill, refills, etc.
- Counters are managed on three levels, subscriber, customer and fraud customer group (same company, but registered as distinct customers)
- Agent can define thresholds on any of the three levels with filtering options based on tariff, counter, billing limit info, subscription type etc.
- Agent can define “user defined rules” – complex expressions spanning mix of daily and monthly counters, simple comparison and complex statistical operations

<table>
<thead>
<tr>
<th>Subscription Type</th>
<th>Tariff</th>
<th>Counter</th>
<th>Threshold</th>
<th>Total limit from</th>
<th>Total limit to</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSM</td>
<td>Postpaid 400</td>
<td>UNBILLED_AMOUNT</td>
<td>200</td>
<td>0</td>
<td>400</td>
</tr>
<tr>
<td>GSM</td>
<td>Postpaid 700</td>
<td>UNBILLED_AMOUNT</td>
<td>250</td>
<td>0</td>
<td>300</td>
</tr>
<tr>
<td>GSM</td>
<td>iBusiness 1000</td>
<td>UNBILLED_AMOUNT</td>
<td>400</td>
<td>0</td>
<td>500</td>
</tr>
</tbody>
</table>
General fraud detection – Customer Aggregation

The aggregation of the customer accounts is process of joining two or more customers into a single entity, one the fraud system it’s called fraud customer group. The purpose of the aggregation is to track more than one customer as a single customer.

- Aggregate based on MB and TAX NUMBER
- Aggregate based on the personal or company contact data
- Aggregate based on the company data
- Manual aggregation by fraud agent

The platform scores all the matches and creates so called “Account Join Proposal” list

By joining two or more contracts the fraud group is created. All the counters are also maintained for the group just as for subscriber and individual company customer
General fraud detection – Thresholds and UDRs

- Thresholds and User Defined Rules are evaluated online, together with data integration step. With every file imported (charging transactions, call detail records, refill records) all thresholds and rules are executed.

- If threshold or rule is evaluated positively the alarm is raised.

- Fraud agent has the freedom to define their own alarm types

- Alarm is delivered directly to CRM as a pop-up, e-mail or SMS (by configuration)

- Raised alarms are grouped based on the subscription, customer or fraud group instance level.

- Before alarm resolving, agent can execute additional set of rules and then decide to what to do with alarm.
General fraud detection – Black listing

- Fraud platform supports various black lists, such as:
  - Black listed IMEI numbers
  - Black listed Cells and BSS stations
  - Black listed destination numbers and patterns
  - Black listed Point Of Sales
  - Black listed A party numbers and patterns
  - Black listed geographical locations
  - Black listed address patterns

- When subscriber profile attribute is found alarm is risen automatically
General Fraud Detection – Alarm resolving

- When alarm is risen, the agent must resolve it, one has following options:
  - Ignore alarm until specific date
  - Ask system to remind agent after specific date
  - Automatically inform subscriber / customer KAM about risen alarm
  - Set new credit limit in the billing system
  - Suspend subscriber / customer services
  - Deactivate subscriber / customer account

- All agent actions and every data changed by the agent is logged by the system

- All agent actions are enabled via GUI

- Agent privileges are granted based on the role and rights system
SimBox Detection

- According to different sources the total damage done by the commercial SIM boxes is around 5% of the total revenue of telecoms. If we break down negative effects by category it becomes clear why damage is so high:
  - Revenue loss due to call redirection – roaming calls are intercepted, redirected and terminated as if they are done in home network
  - Revenue loss due to service inaccessibility – due to the poor line quality there is a high rate of dropped calls
  - Revenue loss through missing callbacks – the call redirection is done with strange or missing call line identifier making it impossible for the called party to return call later
  - Image loss due to bad quality
  - The interconnection between carriers and local SIM boxes are done using highly compressed IP connection resulting in loss of voice quality and call setup is extended.
SimBox Detection – How it works?

The detection of SimBox Sim cards is based on behavior pattern detection. The detection logic is encapsulated by the automatic online and on demand filtering rules.

Filters are configured by the fraud agents based on their experience and recommendation by the system. The output of the filtering module is scored result list...
SimBox Detection – Filters

- The detection is performed by running more than 45 detection filters on every subscriber SIM profile.

- Example filters are:
  - First call destination and duration
  - Where scratch cards are bought (geographically) and their value
  - Ratio of MO and MT calls and duration
  - When was the SIM card installed
  - When and how the tariff is changed
  - Number of consecutive calls without MT
  - Number of distinct parties called
  - Ratio between onnet and national calls
  - Money transfer statistics
  - Whether SIM card properties are black listed (cell, bss, imei, geography)
  - etc.
SimBox Detection - Scoring

For every SIM card profile that is being analyzed, the system is creating so-called “Scored result list”. Every MSISDN gets its semaphore showing which filters passed (green) and which have failed (red).

Based on the score, the agent can quickly decide what to do with suspicious profiles.

<table>
<thead>
<tr>
<th>MSISDN</th>
<th>Score</th>
<th>Other lists</th>
<th>Lock/Resolve</th>
<th>Details</th>
<th>Locate</th>
</tr>
</thead>
<tbody>
<tr>
<td>38649713821</td>
<td>37.6%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649724866</td>
<td>37.5%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649725184</td>
<td>37.5%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649119979</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649152304</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649173708</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649416284</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649437258</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649550655</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
<tr>
<td>38649573541</td>
<td>25%</td>
<td></td>
<td>Resolve</td>
<td>Details</td>
<td>Locate</td>
</tr>
</tbody>
</table>
SimBox Detection - Resolving

- When fraud agents get scored result list he can decide what to do with profile:
  - Ignore the number for specific period (used for operator test numbers, operator own employees, VIP’s, business numbers, etc.)
  - Set reminder to check the profile later on.
  - Suspend the services
  - Terminate the SIM card

- All agent actions and every data changed by the agent is logged by the system

- All agent actions are enabled via GUI

- Agent privileges are granted based on the role and rights system
SimBox Detection – Geo searching

- When agent finds a single SIM card that is installed in the SimBox device he can initiate automatic search for rest of the SIM cards. The search is radius based.

- The system will automatically filter all found SIM cards in the vicinity with the same filtering rules.

- With this functionality agents can be extremely successful in detecting SimBox “devices” and can put whole site out of operation.

MSISDN Location - entry

MSISDN: 3684992166

Position: 42° 29' 42N, 021° 06' 51E, Ellipsoid Area, angle=52, semiaxis=442, semiaxis=222
SimBox Detection – Success Factors

The following statistics is accurate data about the number of terminated SIM cards by fraud agents in period of 24 months.

<table>
<thead>
<tr>
<th>DATE</th>
<th>SIMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6.2009</td>
<td>19,557</td>
</tr>
<tr>
<td>1.7.2009</td>
<td>35,437</td>
</tr>
<tr>
<td>1.8.2009</td>
<td>15,366</td>
</tr>
<tr>
<td>1.9.2009</td>
<td>3,282</td>
</tr>
<tr>
<td>1.10.2009</td>
<td>5,048</td>
</tr>
<tr>
<td>1.11.2009</td>
<td>3,882</td>
</tr>
<tr>
<td>1.12.2009</td>
<td>2,442</td>
</tr>
<tr>
<td>1.1.2010</td>
<td>1,733</td>
</tr>
<tr>
<td>1.2.2010</td>
<td>2,069</td>
</tr>
<tr>
<td>1.3.2010</td>
<td>1,595</td>
</tr>
<tr>
<td>1.4.2010</td>
<td>468</td>
</tr>
<tr>
<td>1.5.2010</td>
<td>420</td>
</tr>
<tr>
<td>1.6.2010</td>
<td>124</td>
</tr>
<tr>
<td>1.7.2010</td>
<td>146</td>
</tr>
<tr>
<td>1.8.2010</td>
<td>505</td>
</tr>
<tr>
<td>1.9.2010</td>
<td>390</td>
</tr>
<tr>
<td>1.10.2010</td>
<td>201</td>
</tr>
<tr>
<td>1.11.2010</td>
<td>115</td>
</tr>
<tr>
<td>1.12.2010</td>
<td>92</td>
</tr>
<tr>
<td>1.1.2011</td>
<td>66</td>
</tr>
<tr>
<td>1.2.2011</td>
<td>1</td>
</tr>
<tr>
<td>1.3.2011</td>
<td>2</td>
</tr>
<tr>
<td>1.4.2011</td>
<td>4</td>
</tr>
<tr>
<td>1.5.2011</td>
<td>3</td>
</tr>
</tbody>
</table>
Risk Scoring

- Risk Scoring is a statistical model based on invoiced and paid amount history. The model goal is to quantify risk attached to postpaid subscriber of becoming a bad payer.

- Model rules are based on:
  - Deltas between invoiced and last payment
  - Dunning actions (scheduled or performed)
  - Risk score of other contracts of the same customer
  - etc.

- Model can be combined with external info (e.g., Bank credit profile like FICO).

- For new postpaid customers, the Risk scoring depends on:
  - Whether person has employment
  - Whether he is returner
  - Geography area
  - Marital status
  - Age
  - etc.
Legal Report

One place where agent can retrieve all known profile data, usage and movement history.

The export document is PDF with
- Transaction raw information
- Transaction summary
- Geography locations (cell based)
- Google map images of every location (cell based)
- All profile data

The legal report supports all type of the subscriptions: GSM, FIX & Internet

Report can be executed ad-hoc or scheduled and delivered to agent email when finished.
General Time series predictor model is used for:

- Forecast spending day by day for every subscriber in next 30 days
- Forecast next invoice value for postpaid subscribers / business customers
- Forecast roaming expenses based on history roaming records for subscribers in roaming
- Detect in front - bill shocks
- List out-of-ordinary spending – and check them in traditional way

The Fraud platform is extended with a Data Mining system.
SimBox Detecting

SimBox detection Neural Network:
- Based on 40+ filters best explaining profile variables
- Trained neural network and decision tree model
- The output of the model checked by the agents in traditional way
- Very early detection of “potential” SimBox profiles, event before their behavior is checked in traditional way.

Attributes:
- MO/MT ratio, SMS/Call ration, International/National ratio, CLIR indicator, any attribute on black list, average duration between top-up, ratio of distinct calls, average duration of MO and MT calls, call pattern tolerance, money transfer value, etc.
Fraud Architecture

**CRM Portal (Information Delivery)**
- Interactive Forms
  - Common Reporting
  - SimBox Detection
  - General Fraud

**Information & Process Integration**
- Linux Bash Parsers
- Oracle External Tables
- Linux OS

**Data store, Data Mining and Data analysis**
- Counters
- Alarms
- Thresholds
- UDRs
- Filtering
- Online Rules
- Profiles
- ...
- General Forecast
- SimNet

**Fraud DB**
**FRAUD DM**
Fraud Architecture – Data integration

The collection of source data is done via:

- Bash scripts and Oracle External data table definitions + Oracle SqlPlus
- DB Links with source databases

Every source data set has its own stage table where data is cleaned and transformed before moving into fraud data tables.

All the collection procedures are scheduled either on database or OS level (Linux crontab)

- Subscriber data is collected every 15 minutes for delta, every night full sync. is occurring
- Invoices and payments are collected every 3 days and every 15 minutes respectively
- CDR’s, ISUP, Fix CDR’s and Charged transactions are collected in real time, as soon as they come from mediation
- Vouchers and refills are collected every one hour
- Transaction data from shops (sold SIM and scratch cards) are collected every 1 hour
VCCE is CRM extension solution designed in modular approach meaning that Fraud modules can operate without rest of the system (except user, role management.

- Fraud GUI developed in PHP
- VCCE is Web 2.0 application
- GUI is simple and intuitive to use
Fraud Architecture – Data store and Data mining

- Detection algorithm, rules and filters are running on the Oracle 11g database level.
- The systems stores all the transaction data for up to 6 months (even 12 months if storage and HW permits)
- Database is well documented and available to Operator development department
- Data mining models are running on separate server (Microsoft Windows Server)
- Data mining software used is Microsoft Analysis Service
Fraud Integration - Overall

- Fraud platform is integrated with every (directly or indirectly) transactional system.
Ai.Fraud – What makes it different?

Our platform is:

- Fully convergent platform with support for all type of CSP services
- Open – we share all the algorithms and models internals
- Successful – proven record in SimBox detection
- Supported – we are continuing with development, it’s not a one time project
- Extensible – we adapt the platform to the problem not problem to the platform