Mebone Fraud

Summary

Today, it is really important to earn and retain the confidence of our customers. The banking business is certainly no exception. Bank fraud not only imposes large losses on banks, but equally important, it decreases customer confidence.

New technologies have definitely become a key competitive factor in banking services. Fraud has increases simultaneously with the growth of electronic payments and electronic payment systems. Criminals are acutely aware of the new opportunities they have as a result of this substantial growth, and financial institutions must adapt their security systems to the new challenges.

Anti-fraud implementations are difficult, especially when there are multiple legacy systems to integrate. But they are very successful when they enable financial services companies to achieve a full return on investment in one year or less after deployment.

This document presents a brief overview of the current state of fraud and introduces our unique approach for stopping the losses caused by fraudulent threats.

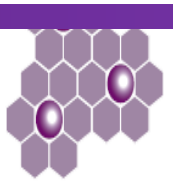
Anti-Fraud Systems definition

What elements make the fraud problem a difficult task?

Data volumes. Only a small percentage of the processed transactions are fraudulent. The total number of transactions is typically very large, which means the system must analyze large volumes of data in order to detect a very small number of illegal operations.

Fraud Patterns change constantly. Criminals are continually looking for new ways to break the security mechanism. Different methods have been used; Terminal manipulation, skimming carding operations, phishing, hacking into data bases, PIN guessing, etc.

MOTO transaction increase. There is also a substantial increase in the number of people across the globe who have access to electronic payments. Chinese banks have multiplied the number of cards they has issued by a factor 13 in the last two years alone. Spanish and Eastern European banks similarly expect a growth rate of ten percent. .



Market Scope:

Where money or value is exchanged, fraud detection is important. Thus, the financial services sector has been an early adopter of fraud detection systems, but other sectors, such as healthcare, insurance, government benefits, telecom and retail, are starting to beef up their fraud detection systems to ward off increasingly sophisticated attacks. Some of these new scenarios are:

- Insurance claims fraud
- Identity-theft
- Account takeover-related fraud

Conservatively speaking, global fraud costs across sectors are likely to total well more than $300 billion, making an investment in anti-fraud systems a worthwhile undertaking.

Technology limits. Because very high volumes of data have to be analyzed, some specific technologies must be used. These technologies can be expensive and some of them are not going to solve the problem. Additionally, these solutions must be scalable.

Payment Processors. Although there are some standards in acquiring operations, from the fraud-detection point of view, the systems have to be flexible in order to adapt to different ways to present and analyze the payment information. Diverse data sources cause extra integration time and costs for this kind of software.

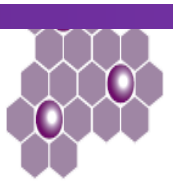
Regulatory Changes. There are a lot of regulatory changes impacting the patterns of fraud: EMV, claim mechanisms, clearing formats, etc. Those changes affect the fraud detection systems if they are unable to adapt quickly.

Technical Requirements

Some of the problems an anti-fraud system should solve indicate the prerequisites that they should fulfil.

Fast adaptation to fraud change. It is necessary that the system learns as fast as possible from the new transactions. It is highly advisable that the system also yields comprehensible and actionable information to fraud analysts.

New fraud scenarios require the testing of new variables and models for learning. Unfortunately, some systems spend too much time (months) to activate new models for detecting new patterns of fraud. If the rate of change is high, those systems will be obsolete before they are put into production.



It is important that the system allows the easy inclusion of new information. In turn, the system should not have to be re-educated each time a new variable is included.

Proper database. The database engine should be able to handle millions of rows and it must store data as efficiently as possible. Additionally the database has to be scalable.

Quick analysis. Anti-fraud software must be able to analyze the transaction and produce a result as fast as an authorization system (real-time with sub second response time).

It is crucial to obtain a value that tells us if the transaction is legitimate or not. This value can also be used into the authorization process.

Diverse Data Sources. Anti-fraud systems have to be able to get the information from different locations and formats. It is important to have enough flexibility to integrate different data sources.

Further, erroneous data, redundant information, non-mandatory fields, etc… must be ruled out during the learning process in order to avoid artefacts.

Friendly use. It is important the system to be as manageable as possible for the fraud analyst. To reach this objective it is necessary to allow anti-fraud professionals to use their own language when they are defining detection rules.

These are automatic methods (none directly dependant on the analyst) and they are used when the analyst does not have an intuitive vision of the cause of fraud. It is also applied when the rule's mechanism approach is too complex in order to cover all the cases.

Required ?	ANN	BBN	Rules
Analyst Knowhow	-	✓	✓
High speed of learning	✓	✓	N/A
High speed of analysis	✓	✓	N/A
Feedback /Transparency	-	✓	✓
Critical reaction against changes	-	✓	N/A
Fraud detection	✓	✓	✓

ANN: Artificial Neural Network
BBN: Bayesian Belief Network

Technology

This paragraph will review the systems based on a mixture of Bayesians Networks and Rules. It will be interesting to define which will be the technological requirements for accomplishing the fraud detection.

Database.

Database used for fraud detection should achieve some specific objectives:

- It should be able to store huge amounts of data.
- The engine should be fast and efficient.
- It should be able to manage data aggregation.
- It must optimize data allocation.
- It must make efficient use of indexes.

There is a great variety of database types but only vertical databases fulfil all of the requirements listed above.

Rules and Bayesian Networks.

Rules are looking for transactions that match the criteria contained in the rule. There is an engine in charge of the searching. The characteristics of the searching are well supported by vertical data base engines.

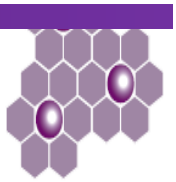
Additionally the query language must be accessible to the analysts. Rule building tools using natural language make the system easier to use.

On the other hand, a Bayesian Network adapts perfectly to fraud detection. Bayesian algorithms are based on probabilistic calculus. When appropriate variables are used and the network (composed by nodes/variables) is properly adjusted and trained, we get grater probability for detecting an operation to be fraudulent.

Bayesian Networks are able to learn during their training with transactions and their status (legitimate – illegitimate).

Rules and Bayesian filters must be initially configured by the analysts and in both cases, the ability to use a natural/business language, based on aliases, is a great help.

Probabilistic calculus finds the ratios for each variable value used for fraudulent and legitimate transactions. Vertical data bases are very efficient in this kind of aggregations. Therefore, these technologies (Bayesian algorithms & Vertical data bases) complement each other very well.



Every fraud detection system has to yield basic information about its detection quality. Two parameters are especially important:

- Percentage of detected fraud.
- Percentage of false positives.

False positive indicates the number of transactions that were erroneously marked as illegitimate. The system is obviously as good as its ability to detect fraud, but when the percentage of false positives increases, its quality decreases drastically. The anti-fraud job is seriously affected by the increment of false positives. The system is indicating that legitimate transactions are fraudulent, so, they must be reviewed when in reality they are legitimate and the fraud analysts will be wasting precious time investigating them.

Both dimensions have to be tuned properly not only for the complete network or rules set but also for each particular node or rule.

When a transaction is alerted as illegitimate, the analysts have to decide what kind of action to take. The system should be easily integrated with the back-office tools (black lists update, card blocking, international lists update,...)

Finally, good reporting capabilities are required within an anti-fraud system. It is very important for the operator to be constantly informed about the fraud evolution.

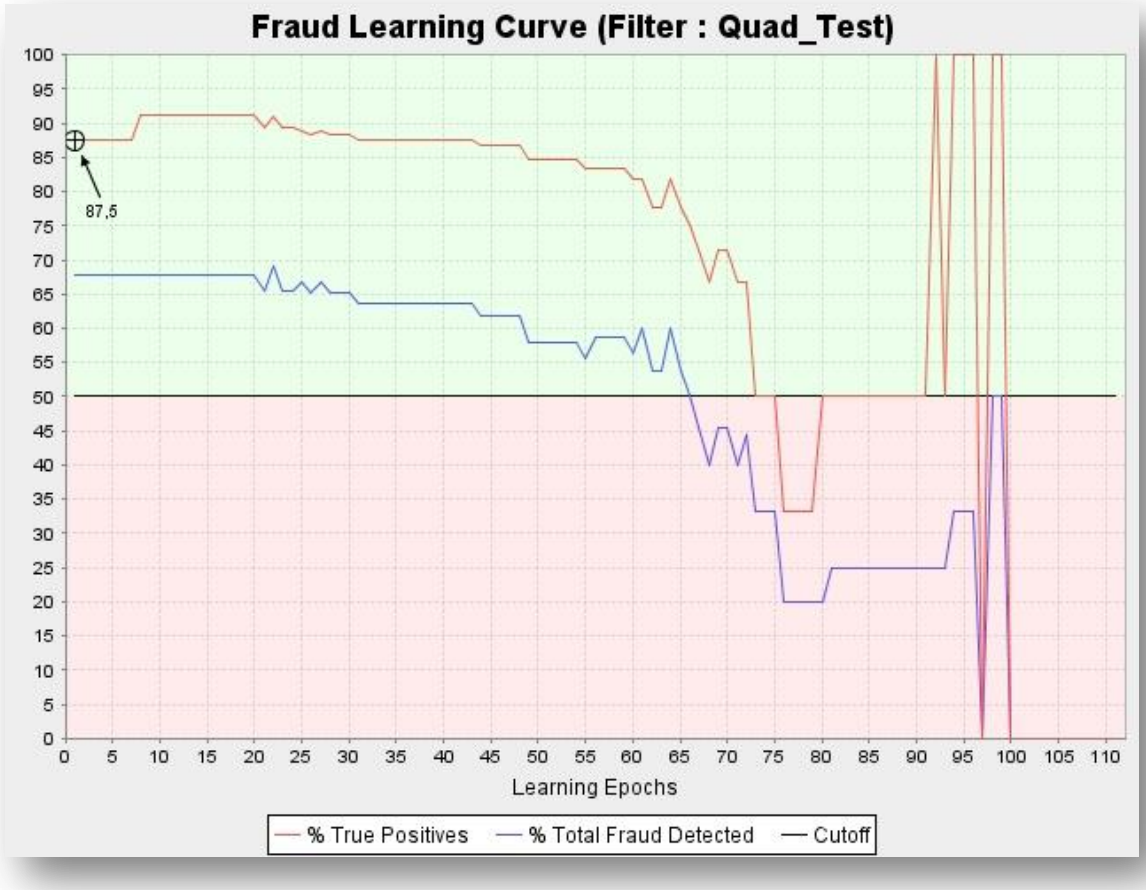
Mebone® Fraud

Mebone® Fraud is an anti-fraud tool of Evendor Engineering. Technologically it is based on:

- Vertical Data Base
- Bayesian filters (BBN)
- Natural language Rules
- Evendor Mebone® ETL

There are different vertical data bases available in the market. This kind of data storage is most suitable for the type of analysis that the system has to do. The rules and Bayesian filters use the data engine properties to achieve a high performance level.

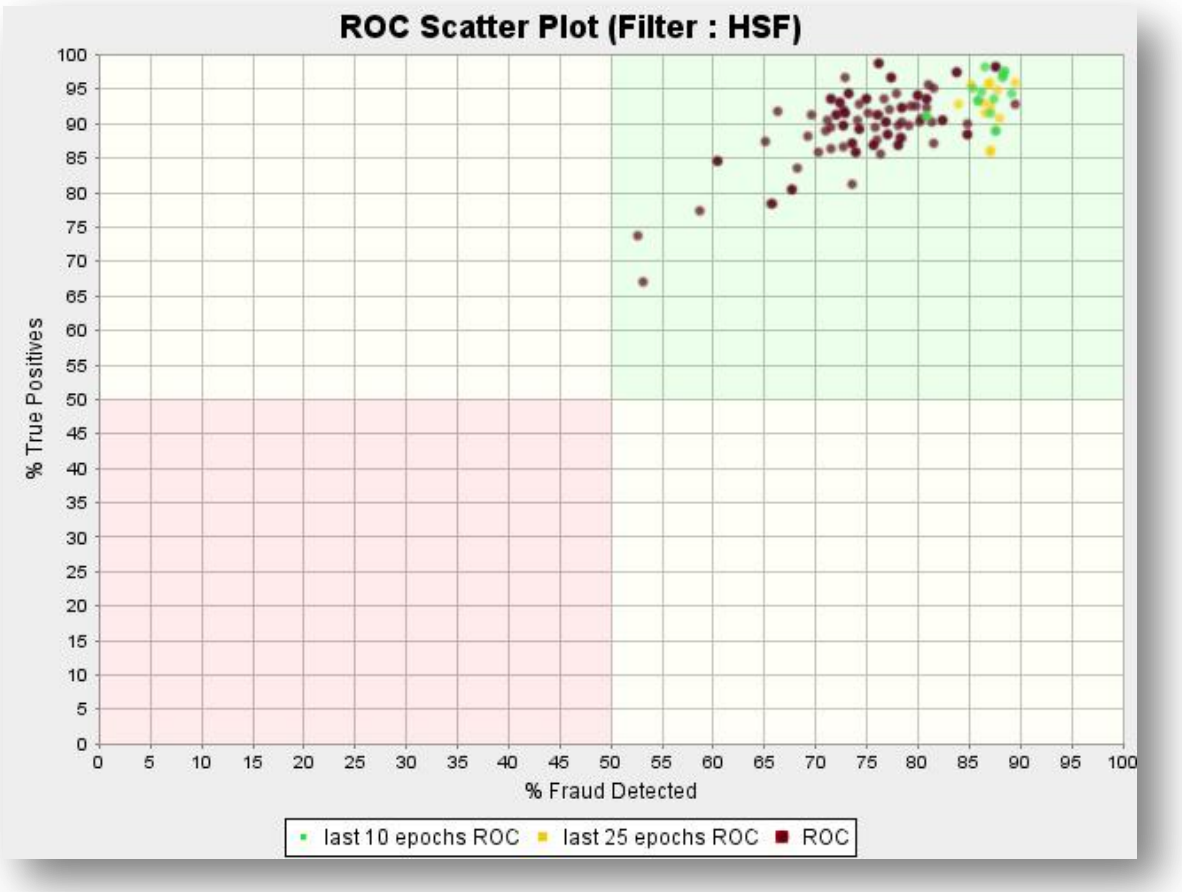
The following graph shows (right to Left) the evolution of a Bayesian filter exposed to eleven learning sessions.



Bayesian network evolution during learning sessions

A Bayesian filter is composed of different variables i.e. nodes. Its quality for detecting fraudulent transactions and avoiding false positives improves over time.

It is interesting to analyze the ratio between identified fraud and false positives. This is done in the ROC graphics that demonstrates the Bayesian filter quality.

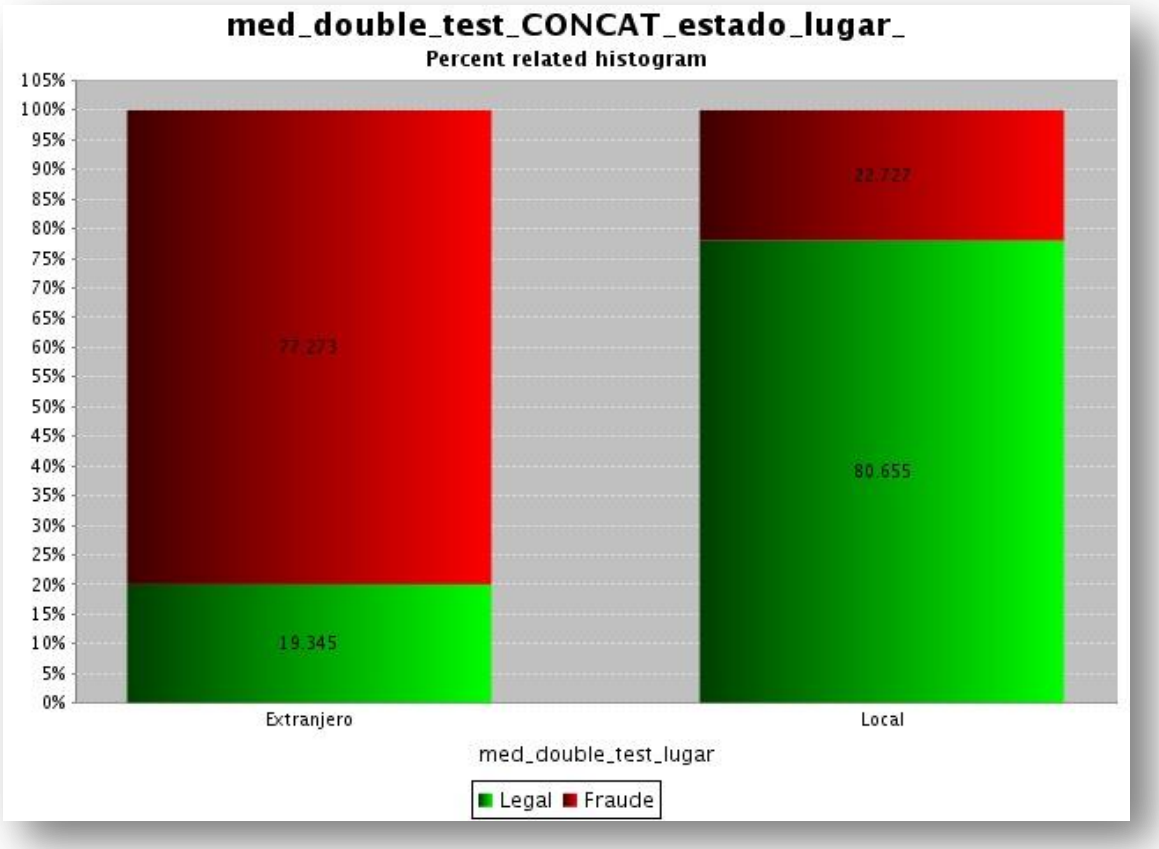


ROC Graphic. Green zone represents a high quality filters inside.

Fraud elements are constantly changing and the points in the ROC plot are also moving. That's the reason this information is very useful as it enables analysts to modify variables/nodes used by a Bayesian filter as soon as a change in ROC or Fraud learning curve is detected.

Additionally a bad choice of variables is easily seen on a ROC plot. It allows taking actions in order to channel the fraud system to the right direction.

These and other graphics illustrate to the analysts the changes that are happening in the fraud environment allowing them to act as soon as possible.



Graphic showing the quality of the variable 'location of the operation' (local or non-local operation). The difference between both columns indicates if the variable is appropriated for the Bayesian filter.

The fraud detection system is regularly reporting its evolution. The Rules and Bayesian network are under on-line control. Its fraud change adaptation, quality and efficiency are watched making use of these reports.

A set of tools is available in the system for building different rules, activating and deactivating them, as well as helping users define new Bayesian networks or modify those already created.

The learning period for the Bayesian network is so efficient that if the training data is available, it can learn what it needs to in a matter of weeks.

The system is especially designed for a complete integration with the data sources and the back-office services that are commonly required. Its ETL[(*)] layer is responsible for this fast and advantageous integration.

Vertical Data Bases

Vertical data bases are also known as columnar, column-oriented or column-based databases. Their engine is specially designed for data analysis. Counting, data aggregation and binary searching capacities are strengths of these types of data bases. All of those properties are very useful in the analysis and detection of fraudulent operations.

Mebone® Fraud uses all of these properties to achieve very high performance.

Conclusions

Different sectors and markets are claiming more efficient solutions for fighting emerging fraud. Some of the most demanded characteristics for this kind of product are:

- High performance data analysis.
- Self-learning capacities.
- Transparency – no black boxes.
- Real-time fraud detection.
- Fast adaptation to fraud pattern changes.
- Efficient Case Management tools.

Mebone® Fraud is able to cover the most of those necessities. Some of its strengths are:

- High performance forensics analysis using vertical database engines.
- Self-learning mathematically predictive model based on Bayesian Networks.
- Natural/Business language facilities for building rules and networks.
- No black box. Probabilities of illegal operation are justified.
- Real-time fraud detection (more than 100.000 trans. processed per second).
- Low false positive ratio. On-line configurable positive threshold.
- Fast adaptation to fraud's changes (without retrained periods - continuous self-learning).
- Fast-low cost installation and Integration (external systems, legacy systems, source data) Thanks to its own data mapping strategy Mebone®ETL.
- Pre-configured Ruleset for Compromise Point of Purchase (CPP) detection.

Bayesian filters and rules engines work together in order to achieve the high efficiency of Mebone® Fraud.