TANGO Native Anti-Fraud Features

Tango embeds an anti-fraud service that has been successfully implemented by several large French banks for many years. This service can be provided as an independent Tango environment application.

List of features:

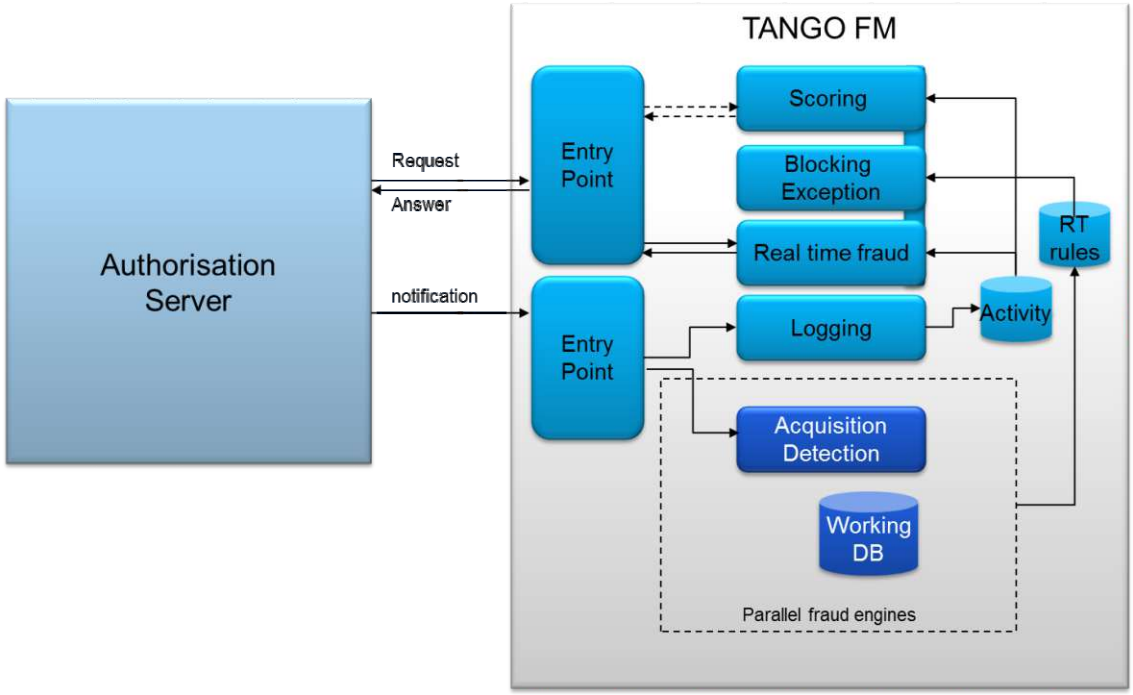
- The capacity to integrate flows from various sources: (ISO 8583 messages, JSON Rest API, XLM, WEB services, Files, ISO 20022 messaging) and interconnect to switches using API, ISO 8583 messages, WEB services, JSON Rest API
- A powerful rule based Boolean language and real time engine allowing Markovian anti-fraud processing
- A complementary authorization engine allowing non-Markovian anti-fraud processing
- A GUI to set up rules and monitor alerts

Markovian real time engine works as follows:

- Risk Management activity is notified by Servers and reacts on a real-time basis
- Reactions are:
    - Create Alerts
    - Notify Blocking rules to server
    - Notify Score rules (or score modifications) to server for cards, merchants, countries…
    - Risk Management can process complementary authorization requests from the Server and answers:
        - Scoring value
        - Authorization code (refused for fraud suspicion)
- Acquire information
- Detect fraud attack or suspect behavioral
- Product:
    - alerts
    - orders to blocking service
- Three parallel services
    - Counters management and threshold crossing
        - Using an open Bool algebra
        - Operators: +, -, AND, OR, EQUAL, In list, NOT, Not in list.
        - Ex: suspected fraud in Portugal for a list of Merchant
    - Rules
        - Pre-configured or added in a rules engine
        - EX: suspected fraud in Portugal for a list of Merchant
    - Execution of Scoring / Blocking rules for transaction refusals
        - Delivers a suspicion score or a direct answer (refuse)
        - Parameterized by user or by Acquisition/detection service
        - Using memory filters for a direct quick answer

-Exception

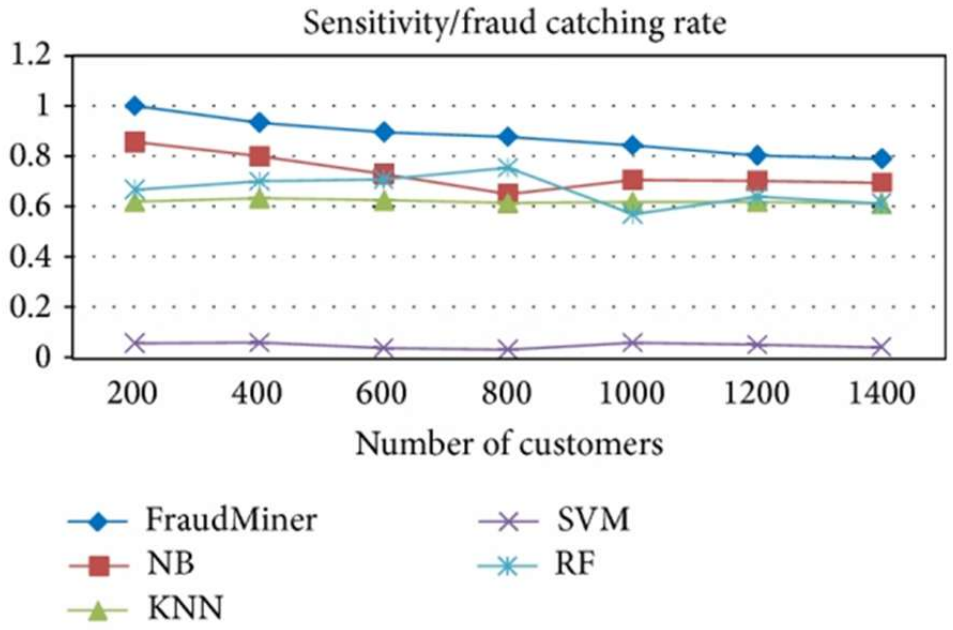
- "Since only one financial transaction of a thousand is invalid, no prediction success less than 99.9% is acceptable" (R. Brause)
- This function is there to review the result of the Scoring / Blocking in case the card (or merchant…..) belongs to a specific list or combination
- This feature allows to by-pass Scoring / Blocking rules



Lusis R&D Activity in Anti-fraud Space

Lusis has led several research projects with Universities and research institutes in the anti-fraud space. This research activity has enabled Lusis to build both a consulting practice and a library of conceptual tools and software modules than can be used in specific projects.

For each application, standard binary classification performances are measured (http://www.damienfrancois.be/blog/files/modelperfcheatsheet.pdf). The key variables to establish the accuracy of a fraud detection system are the sensitivity and the false alarm rate. Sensitivity represents the portion of actual positives which are predicted positives. In credit card fraud detection, sensitivity denotes the fraud detection rate and is defined as TP/P (with P: all positives, number of fraud transactions; negatives, TP: true positives).



False alarm rate represents the portion of actual negatives which are predicted as positives and is defined as FP / N (with FP: false positives, number of legal transactions predicted as fraud; and N: negatives, number of legal transactions).

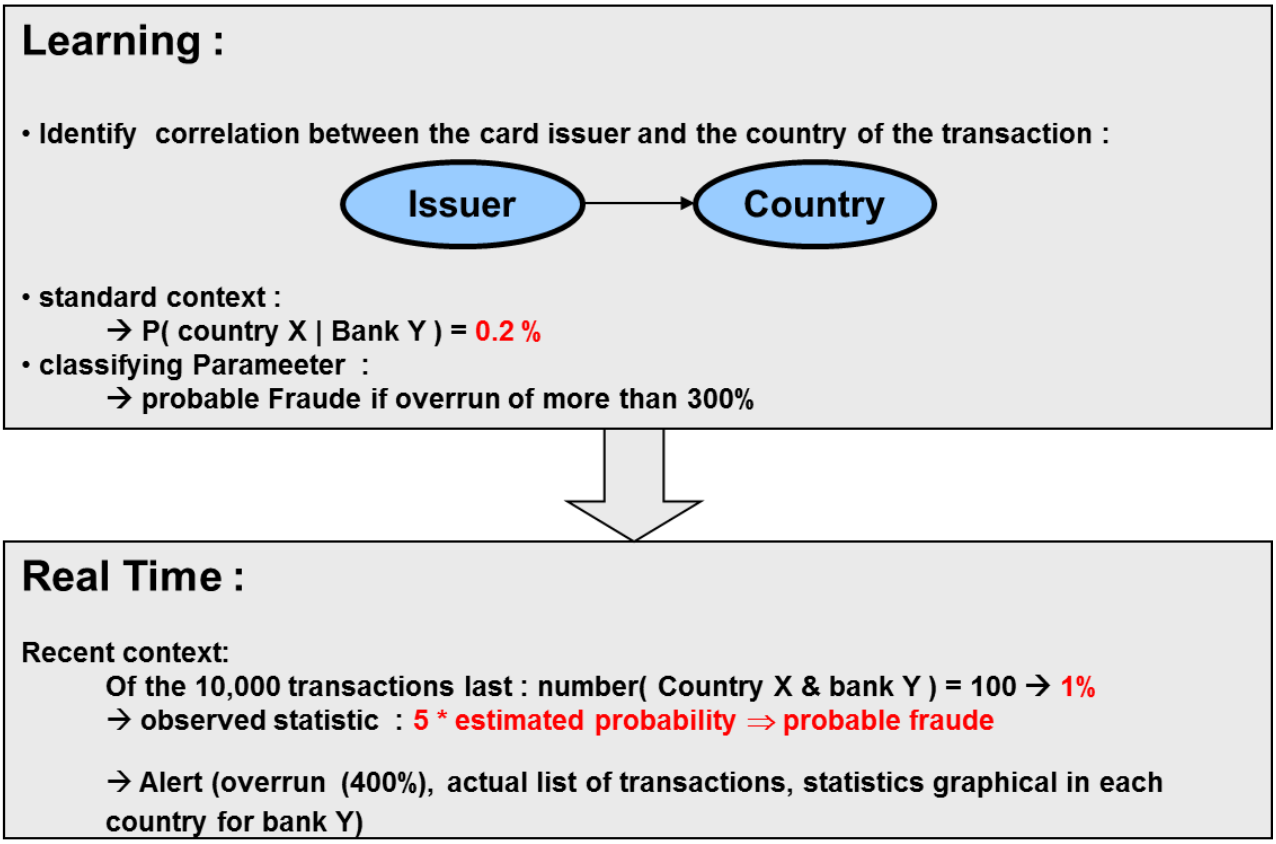
This must be calculated on real datasets that represent the clients' actual business.

The domains worked on can be classified in 3 areas:

- Probabilistic or para-probabilistic models:

    - Bayesian networks

Lusis completed its' first project on Bayesian networks applied to Fraud in 2003. Bayesian networks are a very straight forward application of probability (and conditional probability) theory. For the purpose of fraud detection, two Bayesian networks are constructed to describe user behavior. First, a Bayesian network is constructed to model behavior under the assumption that the user is fraudulent (F) and another model under the assumption the user is a legitimate (NF). The 'fraud net' is set up by using expert knowledge. The 'user net' is set up by using data from non-fraudulent users. During operation the user net is adapted to a specific user based on emerging data. By inserting evidence in these networks and propagating it through the network, the probability of the measurement x less than two above mentioned hypotheses is obtained. This means, it gives judgments to what degree observed user behavior meets typical fraudulent or non-fraudulent behavior. Bayesian approach incorporates both, expert knowledge and learning.



- Dempster/Shaffer theory

Lusis has completed various projects on this very innovative and difficult theory. Dempster–Shafer theory is a generalization of the Bayesian theory of subjective probability. Belief functions base degrees of belief (or confidence, or trust) for one question on the probabilities of a related question. The degrees of belief may or may not have the mathematical properties of probability; how much they differ depends on how closely the two questions are related. Put another way, it is a way of representing epistemic plausibility but it can yield answers that contradict those arrived at using probability theory.

Dempster–Shafer theory is based on two ideas: obtaining degrees of belief for one question from subjective probabilities for a related question, and Dempster's rule for combining such degrees of belief (similar to Bayes Theorem) when they are based on independent items of evidence. In essence, the degree of belief in a proposition depends primarily upon the number of answers (to the related questions) containing the proposition, and the subjective probability of each answer. Also contributing are the rules of combination that reflect general assumptions about the data.

In this formalism, a degree of belief is represented as a belief function rather than a Bayesian probability distribution. Probability values are assigned to sets of possibilities rather than single events: their appeal rests on the fact they naturally encode evidence in favor of propositions.

- FDS: fusion approach using Dempster-Shafer theory and Bayesian learning

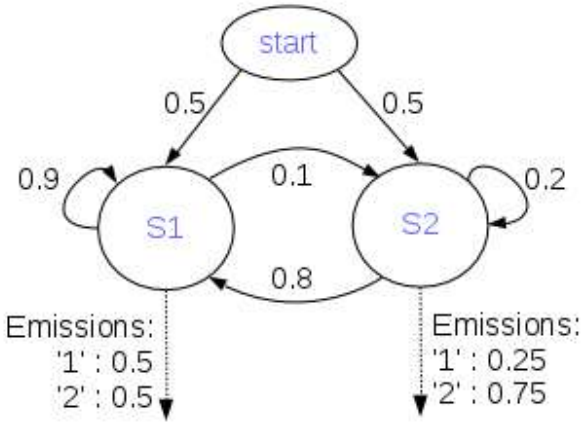
Dempster–Shafer theory and Bayesian learning is a hybrid approach for credit card fraud detection which combines evidence from current as well as past behavior. Every cardholder has a certain type of shopping behavior, which establishes an activity profile for them. This approach proposes a fraud detection system using information fusion and Bayesian learning to counter credit card fraud.

The FDS system consists of four components, namely, rule-based filter, Dempster–Shafer adder, transaction history database and Bayesian learner.

In the rule-based component, the suspicion level of each incoming transaction based on the extent of its deviation from good pattern is determined. Dempster–Shafer's theory is used to combine multiple such evidences and an initial belief is computed. Then the initial belief values are combined to obtain an overall belief by applying Dempster–Shafer theory. The transaction is classified as suspicious or not suspicious depending on this initial belief. Once a transaction is found to be suspicious, belief is further strengthened or weakened according to its similarity with a fraudulent or genuine transaction.

- Hidden Markovian model and Dynamic Bayesian Network

A Hidden Markov Model is a double embedded stochastic process used to model a much more complicated stochastic processes than a traditional Markov model. It can be seen as a “double-deck” random process where we see only the final results:



If an incoming credit card transaction is not accepted by the trained Hidden Markov Model with sufficiently high probability, it is considered to be a fraudulent transaction.

The HMM advantage is that it reduces the tedious work of employees but produces a high false alarm as well as high false positive. Therefore, it is only used as a benchmark.

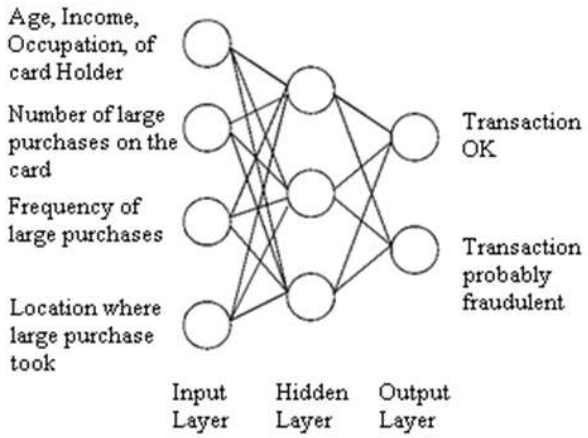
- Neural networks and GANN (genetic algorithm neural networks)

- Neural networks and GANN (genetic algorithm neural networks)

An artificial neural network consists of an interconnected group of artificial neurons. Neural network based fraud detection is based totally on the human brain working principal. As human brains learn through past experience and use this knowledge or experience in making decisions to solve daily life problems - the same technique is applied with credit card fraud detection technology. When a particular consumer uses their credit card, there is a fixed pattern of credit card use, established by the way consumer uses their credit card.

Using the last one or two years of data, the neural network is trained about the pattern of credit card use by a particular consumer. As shown in the figure the neural network is trained on information specific to various categories about the card holder such as occupation and income. Occupation may fall into one category, while in another category information about purchase amount is placed. This information includes the number of large purchases, frequency of large purchases, location of purchase etc. - within a fixed time period.

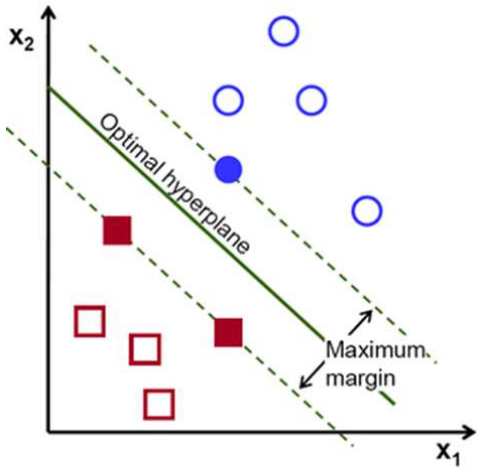
In spite of the pattern of credit card use, neural networks are also trained about the various credit card frauds previously faced by a particular bank. Based on the pattern of credit card use, neural networks make use of prediction algorithms on pattern data to classify whether a particular transaction is fraudulent or genuine. When a credit card is being used by an unauthorized user, the neural network based fraud detection system checks for the pattern used by the fraudster and matches this with the pattern of the original card holder (on which the neural network has been trained). If the pattern matches, the neural network declares the transaction ok.



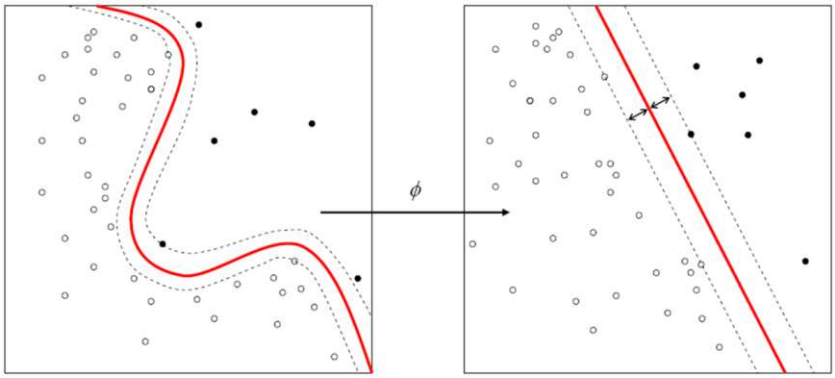
Genetic algorithms are used to accelerate the learning process. Genetic algorithms are evolutionary algorithms which obtain better solutions as time progresses. When a card is copied or stolen or lost and captured by fraudsters it is usually used until its available limit is depleted. Thus, rather than the number of correctly classified transactions, a solution which minimizes the total available limit on cards subject to fraud is more prominent. It aims in minimizing the false alerts using a genetic algorithm where a set of interval valued parameters are optimized.

- Statistical and para-statistical models
    - Support vector machines

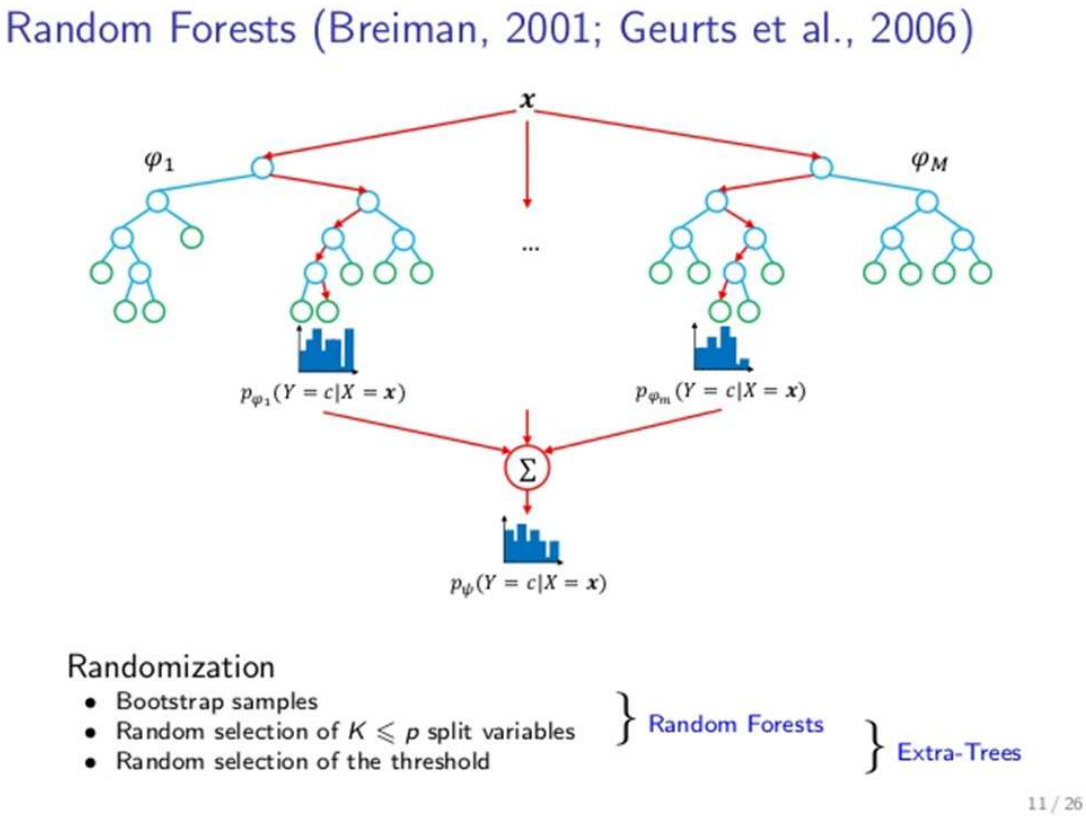
Support vector machines (SVM) are statistical learning techniques that are very successful in a variety of classification tasks. Support vector machines are based on the conception of decision planes which define decision boundaries. A decision plane is one that separates sets of different classes. SVM classification algorithms construct a hyperplane as the decision plane which separates the samples into the two classes—positive and negative.



The strength of SVMs comes from two main properties: kernel representation and margin optimization. Kernels, such as a radial basis function (RBF) kernel, can be used to learn complex regions. This algorithm finds a special kind of linear model, the maximum margin hyperplane, and it classifies all training instances correctly by separating them into correct classes through a hyperplane. The maximum margin hyperplane is the one that gives the greatest separation between the classes. The instances that are nearest to the maximum margin hyperplane are called support vectors. A transformation is required to put the data in a linear space (transformation from input space to feature space).



There is always at least one support vector for each class, and often there are more. In credit card fraud detection, for each test instance, it determines if the test instance falls within the learned region. Then if a test instance falls within the learned region, it is declared as normal; else it is declared as anomalous. This model has been demonstrated to possess a higher accuracy and efficiency of credit card fraud detection compared with other algorithms. SVM methods require large training dataset sizes in order to achieve their maximum prediction accuracy.



Each of these models provides different approaches unequally efficient regarding the nature of the problem. A combination of these approaches must be tuned for each business.

Local Adaptations

For each anti-fraud project, it is necessary to work with local experts not only to “train” or tune the algorithm, but also to understand the local specificities that may create leeway.
Major figures of the fraud landscape can be very different from one country to another and must be adapted or reevaluated for different institutions. This is also why self-learning and easily adaptable algorithms are privileged.

France:
5 Cite Rougemont
75009 Paris
France
(+33) 1 55 33 09 00

United States:
315 Montgomery St
#900
San Francisco CA 94104
(+1) 415 829 4577

UK:
Providian House
16 - 18 Monument St
London EC3R 8AJ
(+44) 207 868 5288

Luxembourg:
321 Route d'Arlon
L-8011 Strassen
Luxembourg
(+352) 31 35 02-1

Integrated Solution

Lusis can provide an integrated solution mixing TANGO risk management and other technologies that use various models or methods (such as SVM, FDS, GANN or others). The integration is done by TANGO. The models are using specific open source technologies. For instance our SVM implementation is done in Java although Tango is coded in C++.