Stage détection des fraudes

Le thème du stage serait lui sur "outlier detection" avec l'idée de

1) faire un état de l'art

2) voir les points communs entre ""outlier detection" et les classifieurs "one class"

3) choisir deux, trois méthodes à tester

4) faire des expérimentations sur des données benchmark et des données télécoms

Le contexte général du stage est décrit ci-dessous. Le stage peut être suivi d'une thèse.

Lieu du stage : Lannion
Rémunération de l'ordre de 1000€ par mois

Profil : Data scientist avec goût pour info et mathématique appliquées
Contact : Vincent Lemaire (vincent.lemaire@orange.com)
http://vincentlemaire-labs.fr/

Le contexte général est la sécurité et la cybercriminalité (détection des fraudes). Face à ces préoccupations, les systèmes de détection de fraudes, de comportements anormaux ou de recherche de signatures d'attaques sont l'une des pièces maîtresses des dispositifs actuels de protection des systèmes. Les tentatives de fraudes, d'intrusions ou d'attaques sont protéiformes, évolutives et rendues furtives par la simple volumétrie Big Data du trafic et/ou des transactions. Dans la suite, nous qualifions la fraude, l'intrusion ou une attaque par acte frauduleux. La surveillance, d'un réseau ou de transactions nécessite la prise en compte, d'une part de la spécificité d'actes frauduleux et d'autre part des contraintes liées au traitement de données variées, volumineuses et à haute vélocité. L'approche de détection par comparaison à une base de données de comportement dits normaux, par exemple à l'aide de règles métiers, présente des limites. L'établissement des règles est un travail complexe, puisqu'elles doivent suivre constamment les évolutions des fraudes connues. Par ailleurs, la détection de d'actes frauduleux non répertoriées n'est par essence pas possible avec cette méthode.

La sophistication des fraudeurs est constamment croissante. On peut même trouver en ligne des « white paper » essayant de donner une analyse des systèmes antifraudes et suggérant des politiques pour ne pas se faire « pincer » comme par exemple en essayant d'imiter un comportement humain [0].

Face à ces constats, les méthodes d'IA, par apprentissage automatique, élargissent le potentiel de détection des fraudes, qu'elles soient connues ou inconnues, et peuvent être associées à des technologies ou des mécanismes relatifs au Big Data.

Les techniques d'IA jouent un rôle important dans la détection d'intrusion, de fraude, de comportements anormaux ou de recherche de signatures d'attaques. Les axes de recherche en cours portent sur des systèmes de sécurité en boucle fermée autour d'un expert (dans un domaine bien déterminé, expert de fraude, d'intrusion, etc) permettant un ajustement « perpétuel ». Ces systèmes comportent au moins les composantes suivantes qui font appel à l'IA :

- une analyse comportementale basée sur des données issues de différentes sources remontées par des sondes. Cette analyse a pour but de découvrir des patterns informatifs, par exemple des « attaques » (au sens large).
- des méthodes d'apprentissage qui vont produire des modèles permettant de prédire le caractère malveillant ou non d'un nouvel événement Ces modèles sont construits à partir d'un recueil d'annotations (ou de qualifications du caractère malicieux d'évènements) par l'expert dans un processus continu d'apprentissage.

Il y a un gros défi à combiner ces techniques pour éviter certains écueils qui peuvent conduire à des actions couteuses :

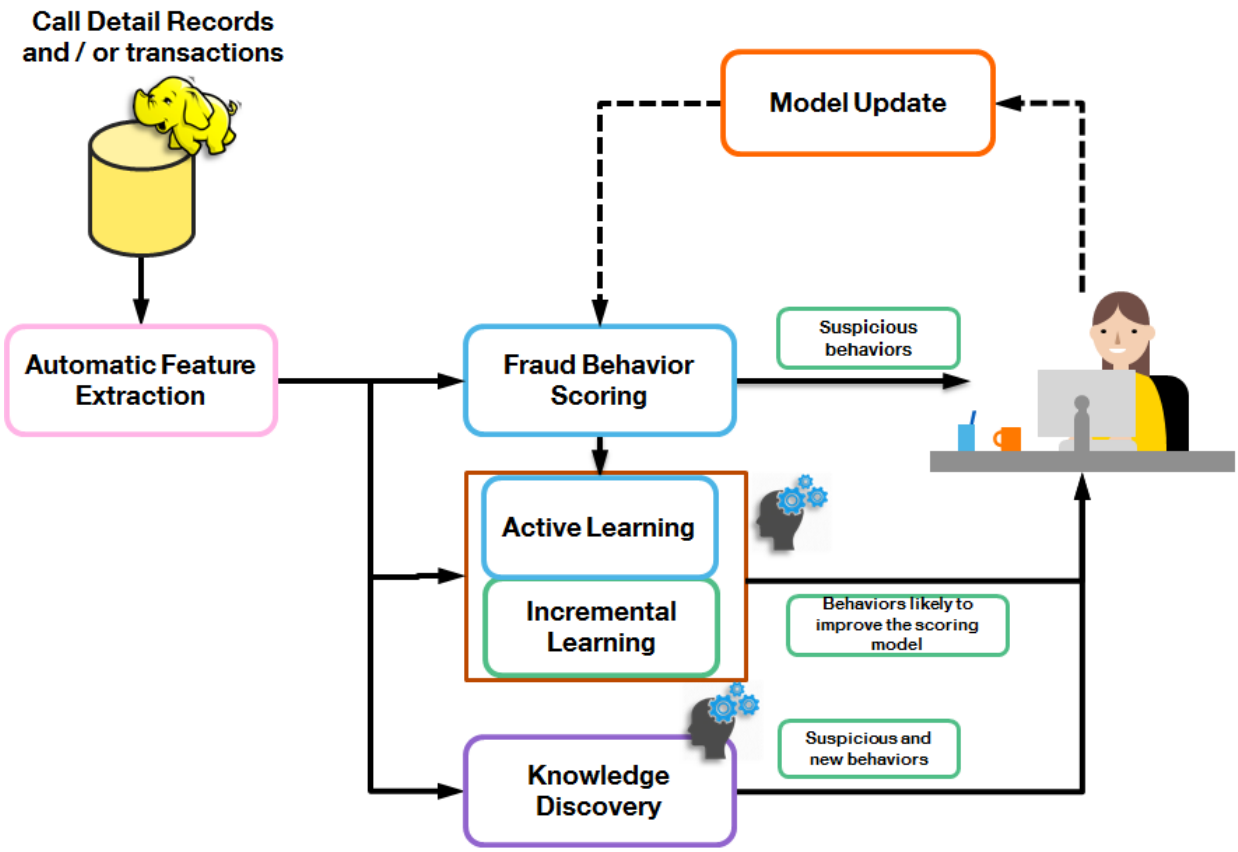
- Génération d'un grand nombre d'alertes
- Problèmes de faux positifs et donc de fausses alertes
- Problèmes de faux négatifs et donc non détection d'attaques
- Imprécisions qui nécessitent des interventions trop fréquentes de l'expert
- Incapacité à découvrir de nouveaux « types d'attaques »
- …

Différents types d'algorithmes d'apprentissage automatique existent (modes supervisé, non-supervisé, semi-supervisé, actif, architecture hybride et combinaison de classifieurs et/ou de méthodes). Lorsque ces algorithmes sont appliqués à la détection d'actes frauduleux on peut établir la un schéma fonctionnel sous forme d'une plateforme (voir figure ci-dessous) comportant :

- une infrastructure « big data » en boucle fermée [1]
- un module de construction d'une représentation adéquate au problème à traiter (Automatic Feature Extraction) [2, 2 bis]
- un module « supervisé » qui détermine le type de comportement (fraude /non fraude) : souvent un prédicteur de la fraude (fraud behavior scoring) entrainé soit sur des données étiquetées à au moins deux classes, soit dans certains sur des données ne comportant qu'une classe du problème [3]
- un module dédié à prendre en compte le feedback de l'expert des actes frauduleux et améliorer le prédicteur constamment au cours du temps :
    - un module d'apprentissage (incrémental) [4]
    - un module d'apprentissage actif pour apprendre de ses succès mais aussi de ses erreurs, ou pour prendre en compte les exemples les plus à même d'améliorer le prédicteur cité ci-dessus [5]
- un module de découverte d'information :
    - détection d'anormalité,
    - détection d'outliers, … [6]
- un utilisateur dans la boucle qui :
    - permet un étiquetage des données.

- possède aussi un ensemble de connaissances expertes et des règles métiers

Un bon système de prévention et détection de la fraude se doit de faire cohabiter intelligemment l'ensemble de ces modules.



Références:

[0] : " How to beat antifraud and start earning money?", Latvia, RIGA, White paper Antrax, (2015)

[1] "AI^2 : Training a big data machine to defund", Veeramachaneni et al., IEEE International conference on Big Data Security (2016)

[2] "Towards Automatic Feature Construction for Supervised Classification", Marc Boullé, European Conference on Machine Learning (ECML) (2014).

[2bis] "Representation Learning: A Review and New Perspectives". Y. Bengio et al. IEEE Trans. PAMI, special issue Learning Deep Architectures..

[3] "One-class classification: taxonomy of study and review of techniques," S. S. Khan et al. The Knowledge Engineering Review, vol. 29, pp. 345–374, 6 (2014).

[4] "A survey on supervised classification on data streams", Vincent Lemaire et al, Lecture Notes in Business Information Processing (2015)

[5] "Active Learning", Burr Settles, Synthesis Lectures on Artificial Intelligence and Machine Learning (2012)

[6] "Outlier Detection Techniques", Hans-Peter Kriegel, Tutorial Notes: SIAM SDM 2010