WAVESTONE

LES NOUVELLES MÉTHODES DE LUTTE CONTRE LA FRAUDE BANCAIRE EN LIGNE

L'ère du numérique s'accompagne d'une multiplication de délits et attaques à l'encontre des établissements manipulant des données bancaires.

En effet, selon la Banque de France, 4,7 millions de cas de fraude ont été observés en 2015 sur les moyens de paiement, engendrant près d'1 milliard d'euros de pertes. La fin de l'année 2016 a justement été marquée par une attaque informatique affectant 40 000 comptes clients de la banque britannique Tesco Bank. Le bilan de l'attaque montre des mouvements frauduleux sur 9 000 comptes, induisant le blocage du système de transactions en ligne pendant 48h et le remboursement de l'ensemble des clients touchés.

Force est de constater qu'une gestion inadaptée de la fraude peut avoir des conséquences extrêmement dommageables : impacts financiers d'une transaction frauduleuse non détectée, impacts d'image et de confiance client, impacts opérationnels (cellule de traitement de la fraude, gestion de crise…). Les techniques de fraude évoluent sans cesse, et il doit en être de même pour les dispositifs de lutte anti-fraude afin d'augmenter leur efficience sans pour autant dégrader l'expérience client.

TROIS AXES MAJEURS POUR LUTTER CONTRE LA FRAUDE

Afin de sécuriser le parcours client, 3 axes majeurs sont à intégrer :

- **La protection des parcours des clients,** afin de sécuriser la réalisation des opérations sensibles via l'implémentation de solutions de protection et la sensibilisation des clients ;
- **La détection de la fraude,** dont l'objectif est de détecter les fraudes réalisées ou en cours de réalisation ;
- **La réaction suite à la fraude,** afin d'alerter, investiguer et réagir rapidement en cas de fraude, suite aux alertes issues de la détection.

AUTEURS



MARTIN DESCAZEAUX
martin.descazeaux@wavestone.com

MATHIEU COUTURIER
mathieu.couturier@wavestone.com

YASMINE EL HIMDI
yasmine.el-himdi@wavestone.com

INTERVIEW

FRÉDÉRIC GERMAIN (SOCIÉTÉ GÉNÉRALE)

Pour les parcours client, l'approche traditionnelle repose sur des solutions de protection multiples telles que l'authentification (simple ou à double facteur) qui constitue une première couche de sécurité essentielle. L'authentification à double facteur offrant des gages de sécurité bien plus importants que les autres types d'authentification, son utilisation se démocratise largement, notamment pour les paiements et les opérations sensibles.

Pour autant, l'authentification étant bien souvent le premier niveau de sécurité rencontré par un client, il est aussi le premier à être attaqué. Des moyens de protection additionnels existent (limitation des options proposées, temporisation des opérations…), mais ils dégraderaient l'expérience client, ce qui va à l'encontre de la simplification et de la fluidification des parcours client recherchées par ces entreprises.

La protection pouvant atteindre ses limites dans de nombreux cas, les entreprises investissent actuellement dans des mécanismes **permettant de mieux détecter la fraude et réagir au plus vite**, tout en continuant à proposer une expérience client satisfaisante. Cette orientation suit également les directives de la règlementation, qui demande aujourd'hui aux établissements financiers d'aller plus loin que l'authentification en étant capable de détecter en temps réel des événements suspicieux ou frauduleux.

TENDANCES DE LA DÉTECTION DE FRAUDE EN LIGNE

De l'approche classique vers le machine learning

L'approche classique de détection de fraude, largement répandue aujourd'hui en France, consiste à détecter des schémas de fraude déjà rencontrés par le passé. Cette approche est principalement fondée sur l'application de règles préétablies, simples ou avancées, sur le flux de transactions :

/ **La détection unitaire**, qui consiste à définir une règle métier où le non-respect d'un critère peut générer une alerte (virement vers un compte/IBAN sous surveillance…) ;

/ **La corrélation d'événements**, qui consiste à mettre en œuvre des règles métiers plus avancées corrélant plusieurs types de données (réalisation d'une opération depuis un pays sous surveillance et dépassement d'un seuil cumulé sur 24h…).

Dans certains contextes où la menace n'est pas prépondérante et où le risque de subir une fraude sophistiquée est faible, **ces solutions peuvent être suffisantes** et ont prouvé leur **efficacité** dans le cas de fraudes usuelles.

Pour autant, face à la complexité et à la diversité des attaques, les stratégies de détection évoluent et s'orientent vers une connaissance client plus poussée. Ce recours à la donnée donne naissance à de **nouvelles approches plus innovantes** fondées sur des algorithmes et technologies d'analyse de larges volumes de données générées et traitées à grande vitesse. Ces techniques permettent de détecter des schémas de fraude connus, mais également d'être proactif face à des situations inconnues susceptibles d'être frauduleuses.

Détecter proactivement grâce au machine learning

Le machine learning est un domaine d'utilisation d'algorithmes capables d'**apprendre au travers d'exemples**. Ce modèle statistique est fondé sur des corrélations découvertes au sein d'échantillons représentatifs de données. Ces algorithmes se développent dans des versions de nature et de complexité diverses en utilisant les dernières avancées techniques, notamment les réseaux artificiels de neurones. La supervision humaine, bien qu'indispensable, s'en retrouve de moins en moins prépondérante.

La conception et l'utilisation d'un algorithme de machine learning passe par trois étapes :

/ **La collecte d'informations à analyser** : Ces données peuvent être soit **internes** (données techniques, données comportementales, données métiers…) et provenant d'un ou plusieurs canaux, soit **externes** (données issues de réseaux sociaux, de site d'informations, de partenaires…).

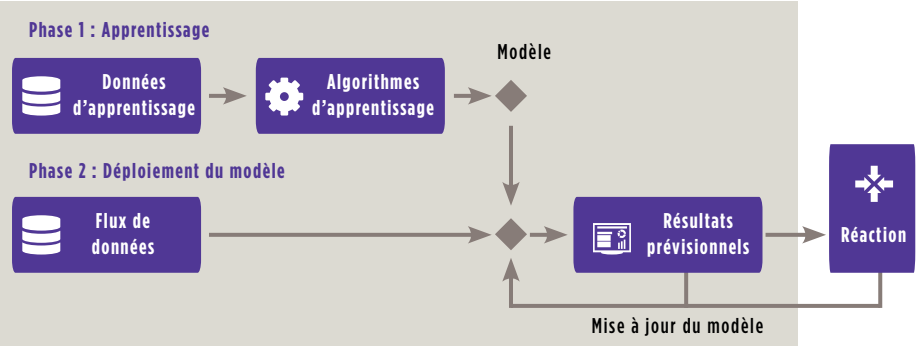
/ **L'apprentissage** : Pour apprendre, le modèle se nourrit des données pour établir, en plus ou moins grande autonomie, des **rapprochements statistiques**, des **règles de décision** et des **paramètres précis** pour améliorer la finesse de la détection. En plus d'une grande disponibilité des données, les techniques de machine learning ne nécessitent pas une fiabilité ou une complétude absolue au sein de leur base : fonctionnant sur des principes statistiques, une erreur ou un « manque » ponctuel n'aura pas d'impact significatif sur le résultat de l'apprentissage.

/ **La prédiction** : La prédiction correspond à la phase d'exploitation finale de l'intelligence auto-conçue. Les données entrantes sont exploitées en temps réel par le modèle, et la cible à prédire est souvent un score de risques : de 0 si la transaction est sûre, à 100 si c'est une fraude certaine.

> Mieux détecter la fraude et réagir au plus vite tout en continuant à proposer une expérience client satisfaisante

En illustration de ces concepts, le machine learning dans le cadre de la banque en ligne revient souvent à **créer des profils clients sur la base d'historiques et d'informations collectées** (terminaux utilisés, heures et lieux de connexions habituels, parcours de connexion et de réalisation des opérations…), puis de **prédire le caractère frauduleux de l'opération en cours en comparant le comportement actuel du client par rapport à son profil**. Il est à noter que ce profil est mis à jour continuellement en fonction des nouvelles opérations réalisées par le client.

Fonctionnement du machine learning



Le marché de la détection de fraude

Que ce soit en France ou à l'international, le **marché de la lutte contre la fraude sur le périmètre des paiements**, premier touché, est **mature**. De nombreuses solutions innovantes, notamment basées sur des techniques d'analyse comportementale et d'apprentissage automatique, sont déployées.

Le périmètre de la banque en ligne est quant à lui **toujours en quête de maturité**. De plus en plus d'initiatives sont observées et, en conséquence, de plus en plus d'acteurs se positionnent.

De manière générale, le marché est composé :

- D'acteurs intégrant une brique d'analyse comportementale à leur solution d'authentification ;
- D'outils à installer sur les terminaux clients ;
- De moteurs de règles à intégrer directement aux sites transactionnels ;
- De solutions de corrélation d'événements (SIEM) adaptées pour la détection de fraude ;
- De solutions de machine learning.

Certains acteurs proposent des solutions combinant plusieurs de ces segments. Pour autant, ces solutions sont souvent spécialisées par canal (monétique, banque en ligne, agence…) et **peu de solutions transverses sont aujourd'hui capables de collecter et d'analyser des données pertinentes en provenance de plusieurs canaux**, permettant d'avoir une stratégie globale de lutte anti-fraude.

Cadre règlementaire et éthique

Les algorithmes fondés sur le machine learning présentent des résultats nettement plus performants en matière de détection de fraude lorsqu'ils sont alimentés par un volume important de données et d'informations pertinentes qui sont, pour la plupart, des données personnelles. Dans ce cadre, certaines exigences de la loi informatique et libertés (LIL) et du règlement européen pour la protection des données personnelles (GDPR) **limitent les cas d'usage des données personnelles et de l'intelligence artificielle**. Il est à noter que les données biométriques (frappe au clavier, mouvement de souris…) sont considérées comme des données à caractère personnel, mais leur utilisation dans le cas spécifique de la lutte contre la fraude fait encore l'objet aujourd'hui de discussions.

Par ailleurs, d'un point de vue éthique, quelle pourrait-être la réaction d'un client s'il apprenait que sa banque bâtissait un profil détaillé de son comportement et de ses usages, sans qu'il n'en soit explicitement prévenu ? C'est pour cela qu'en raison de la nature des données que les algorithmes de machine learning sont amenés à traiter, une vigilance et une transparence forte doivent être apportées afin de garantir la confidentialité des données, respecter les règlementations et préserver l'image de marque et la relation de confiance client.

ALERTES ET CONTRE-MESURES

La mise en place d'outils de détection de fraude automatisés nécessite d'étudier la remontée d'alertes et les contre-mesures associées. L'enjeu pour les entreprises est double : premièrement, **réagir au plus vite** face à une suspicion de fraude, et deuxièmement, **arriver à bloquer in fine la transaction frauduleuse** avant qu'elle ne soit effective.

Automatisation des contre-mesures

L'automatisation des contre-mesures permet de réagir instantanément face à une suspicion de fraude, mais engendre un **risque de gêne client en cas d'erreur**. Les solutions actuelles, même celles fondées sur des règles « simples » ou sur des algorithmes de machine learning, peuvent générer un nombre non négligeable de faux positifs, et donc d'erreurs. Le **degré de confiance** dans les dispositifs de détection est donc le **premier critère de l'automatisation**.

Architecture cible

Certaines solutions de détection proposent des fonctionnalités de gestion des alertes, mais elles sont souvent limitées et l'ajout d'un module complémentaire est souvent nécessaire. Il peut se matérialiser par :

- Le recours à un outil du marché, notamment au travers des solutions de gestion des incidents ou d'investigation des alertes ;
- L'utilisation d'un outil interne, notamment à l'aide de solutions de gestion de notifications et de communication ;
- Le développement d'une solution interne qui pourra s'appuyer sur des composants existants.

Par ailleurs, l'intégration des contre-mesures automatiques peut entrainer **certaines modifications d'architecture majeures** du SI cœur de l'entreprise (blocage ou annulation d'opérations, authentification forte pour valider l'opération…). Ces impacts doivent être anticipés dès le début des projets.

Adaptation des processus métiers

Au-delà de l'identification des périmètres qui doivent être intégrés dans la gestion des alertes et la réaction, il est nécessaire de **travailler avec les métiers porteurs de ces processus** et de **les transposer dans un modèle plus automatisé**. Il s'agira également d'identifier qui est alerté : les équipes d'analystes ? les conseillers ? les clients directement ?

Les défis auxquels les établissements manipulant des données bancaires sont confrontés en matière de fraude ont une incidence considérable sur la confiance client et sur la lutte engagée pour limiter les pertes. Une analyse en temps réel des données clients par des outils basés sur le machine learning peut constituer la recette idéale pour assurer la détection proactive de la fraude, sans négliger l'importance de la réaction qui permettra in fine de bloquer les opérations frauduleuses. L'arrivée de nouveaux modes de paiement tels que l'instant payment renforce encore la nécessité de détecter en temps réel et d'automatiser la réaction, tout en respectant le cadre règlementaire qui encadre de plus en plus ces pratiques.

INTERVIEW : FRÉDÉRIC GERMAIN

DIRECTEUR DU PROGRAMME DE SÉCURITÉ DES SI DE LA BANQUE DE DÉTAIL SOCIÉTÉ GÉNÉRALE

Aujourd'hui, quelles sont les menaces auxquelles vous devez faire face dans la lutte contre la fraude ?

Depuis plusieurs années, les SI des banques s'ouvrent de plus en plus vers les clients et donc au monde extérieur, et s'exposent ainsi à des menaces qui se sont grandement diversifiées et accentuées allant du phishing à l'usurpation d'identité, du « simple » vol de mot de passe à l'ingénierie sociale et au vol massif d'informations sensibles.

Ce contexte de menaces nous encourage fortement à renforcer notamment nos capacités de protection et de détection au profit du Client au travers de la solution innovante du machine learning.

Pourquoi l'utilisation du machine learning pour lutter contre ces menaces ?

Pour répondre aux menaces toujours plus sophistiquées, nous souhaitons en effet passer d'une détection réactive de fraudes connues à une détection proactive de fraudes inconnues en utilisant le machine learning et envisager à terme la biométrie comportementale.

C'est pourquoi le Programme de sécurité des SI de la Banque de Détail Société Générale, lancé en 2015, outre le renforcement des dispositifs de lutte contre la fuite d'information, a comme principal objectif de mettre en œuvre des dispositifs innovants capables d'apprendre par eux-mêmes et de détecter les événements « frauduleux » avant tout impact négatif pour les clients et la banque.

Quelles ont été vos principales difficultés dans le cadre des projets de détection et d'alerting ?

Au-delà des challenges techniques de développement et de méthodologie projet posés par ces technologies encore récentes, une des principales conditions de réussite se trouve dans la mise à disposition d'une masse importante de données hébergées au sein d'un Big Data et de leur corrélation.

L'objectif, pour lutter contre la fraude, est de pouvoir constituer un profil d'habitudes de nos clients et d'estimer un niveau de risques en temps réel pour chaque action de navigation et de transaction, en toute conformité avec les dispositions de la CNIL.

Ces « difficultés » ont été levées grâce, notamment, à l'étroite collaboration avec les équipes en charge du Big data et des Directions « clients ».

Une autre condition de réussite est d'adopter une « nouvelle » politique de recrutement tournée vers la réactivité et la mise en valeur des profils encore rares de Datascientists.

Quels sont vos enjeux pour 2017/2018 ?

Nos principaux enjeux pour 2017/2018 seront d'étendre les périmètres de nos travaux à l'ensemble des marchés de clientèles afin de répondre aux besoins de nos différentes directions marketing & commerciales et de renforcer la performance de nos solutions de machine learning, tout en étant attentif à des solutions du marché dont la maturité pourrait satisfaire nos attentes.

WAVESTONE

www.wavestone.com

Wavestone est un cabinet de conseil, issu du rapprochement de Solucom et des activités européennes de Kurt Salmon (hors consulting dans les secteurs retail & consumer goods). Il figure parmi les leaders indépendants du conseil en Europe.

La mission de Wavestone est d'éclairer et guider ses clients dans leurs décisions les plus stratégiques en s'appuyant sur une triple expertise fonctionnelle, sectorielle et technologique.