Analyse comportementale : la clé de la sécurité ?

Introduction

Détecter les signaux faibles d'une compromission dès qu'elle est survenue. C'est le rêve que l'analyse comportementale entend mettre à portée de la main des RSSI. Elle s'appuie sur des modèles d'analyse qui ne sont pas forcément nouveaux, mais profite du Big Data pour les appliquer à des volumes de données considérables, jusqu'ici largement inexploitées, à commencer par les logs des systèmes connectés à l'infrastructure informatique.

Surtout, cette analyse prétend aujourd'hui tirer profit du Machine Learning (ou apprentissage statistique) pour détecter spontanément des motifs comportementaux sans passer par leur modélisation préalable par des analystes. Une modélisation qui limite forcément les recherches et laisse ainsi un plus grand champ de manœuvre aux attaquants.

C'est donc naturellement que l'analyse comportementale apparaît comme un nouvel eldorado de la sécurité. Un eldorado vaste et porteur de multiples promesses. Mais qui reste encore largement à explorer.■

VALERY MARCHIVE

TechTarget / LeMagIT

L'analyse du comportement utilisateur, nouvel eldorado de la sécurité

Quel est le point commun entre Fortscale et SentinelOne, finalistes de l'édition 2015 de l'Innovation Sandbox de RSA Conference, et HP ? L'analyse du comportement des utilisateurs, ou *User Behavior Analytics* (UBA).

En l'occurrence, HP a profité de la dernière édition de RSA Conference pour présenter, la semaine dernière, ArcSight User Behavior Analytics. Cet outil doit permettre de disposer d'une visibilité accrue sur le comportement des utilisateurs et ainsi identifier les menaces liées à ces derniers en découvrant des motifs de comportement anormaux.

Pour produire ce travail d'analyse, la solution de HP s'appuie, comme d'autres, sur les rapports d'activité, les logs, consolidés par le SIEM. D'où le lien avec ArcSight. Mais là où d'autres s'appuient sur des technologies et des algorithmes développés en interne, HP préfère intégrer les outils de Securonix.

Un marché en pleine ébullition

Et le groupe a probablement eu l'embarras du choix, tant ce marché naissant de l'UBA est animé. Dessus, Gartner recense notamment des grands groupes comme BAE systems, IBM, Lockheed Martin, Oracle, Raytheon, ou encore SAS. Mais les plus prometteurs sont peut-être parmi les plus jeunes, comme Exabeam, Fortscale, GuruCul, iDetect, ou encore, donc, SentinelOne. Mais l'on peut également relever SpectorSoft, arrivé sur le marché mi-avril, ou encore Adallom, spécialisé sur la surveillance des accès Cloud.

Il faut reconnaître que les promesses ont de quoi allécher. Dans une étude, le cabinet explique que l'UBA s'appuie sur trois composants – l'analyse de données, l'intégration de données et leur présentation – pour réduire le nombre des alertes de sécurité, remonter les plus importantes, mais également accélérer les enquêtes et réduire les besoins en effectifs. Surtout, l'UBA promet de permettre de détecter les attaques – ou les comportements irrespectueux des règles de sécurité internes – plus vite, et sans interrompre l'activité.

Dans un billet de blog, Saryu Nayyar, Pdg de Gurucul, ne dit rien d'autre : pour elle, l'UBA doit permettre d'aller plus vite dans l'évaluation du risque associé aux comportements « en utilisant le Big Data et des algorithmes de *machine learning* ». Le tout en quasi-

temps réel, en s'appuyant sur les rôles et les titres des utilisateurs consolidés dans les applications des ressources humaines et les annuaires : ces données, ainsi que celles des accès et des activités, « sont corrélées et analysées en fonction des activités passées et actuelles ». Le tout devant permettre d'établir des modèles.

Utiliser l'UBA de manière opérationnelle

Dans une seconde étude, Gartner recommande de commencer modeste « ou avec quelques sources de données essentielles, comme une source d'activité telle que les logs Windows et une source de contexte utilisateur, telle qu'Active Directory », avant de pousser l'intégration plus avant. Cette sélectivité initiale permet de limiter « le bruit de fond » généré.

Mais rapidement, l'UBA mérite d'être exploitée de manière opérationnelle pour alimenter une console de gestion d'alertes. Et Gartner de citer l'exemple d'un important opérateur câble dont l'infrastructure produisait environ 500 000 alertes par jour ; en s'appuyant sur l'outil d'UBA de Bay Dynamics, après six mois, il ne restait plus que 800 alertes, dont 5 à 10 clairement identifiées comme les plus critiques.

Le cabinet mentionne également plusieurs clients d'Adallom ayant identifié les activités d'un logiciel malveillant ou encore celles de pirates, sur leurs applications SaaS, grâce à son offre d'UBA. Les services de NuData sont également mentionnés pour la découverte de tentatives d'attaques et de fraudes.

Des usages multiples

L'analyse comportementale a fait l'objet d'importants efforts dans les services marketing, notamment, au cours des dernières années. Mais elle a aussi trouvé toute son utilisé dans la lutte contre la fraude, pour de nombreux secteurs d'activité. Le rachat de Silver Tail Systems par RSA, en 2012, visait d'ailleurs explicitement ce domaine.

Mais c'est donc sans surprise que l'analyse comportementale gagne aujourd'hui du terrain dans la sécurité des systèmes d'information, pour détecter les menaces internes comme les attaques ciblées.

Fin 2011, Hugh Njemanze, co-fondateur d'ArcSight, relevait toute la pertinence du sujet : « l'analyse et l'identification de schémas comportementaux en sécurité, c'est comme dans le marketing ! On parle là

l'algorithmes personnalisés qui cherchent des événements ».

Et aujourd'hui, pour la sauce secrète des spécialistes de l'UBA tient au « machine learning », comme le soulignait récemment dans nos colonnes Idan Tendler, Pdg et co-fondateur de Fortscale, précisant que le Big Data n'apporte finalement que le socle technologique permettant de passer à grande échelle.

– *Valéry Marchive*

L'analytique au service de la sécurité : un domaine encore émergent

Ils sont à l'honneur les Fortscale, SentinelOne, Securonix, Exabeam, GuruCul et autres Adallom. Sans compter iTrust ou encore DarkTrace. Leur point commun ? Ils entendent appliquer l'analytique à la sécurité pour détecter les signaux faibles d'attaques ciblées, ou encore lutter contre la menace interne. Mais leurs approches sont différentes.

Eric Ahlm, directeur de recherche spécialiste de la sécurité chez Gartner, distingue ainsi trois camps distincts : ceux qui misent sur l'analyse des flux réseau, ceux qui se concentrent sur le comportement des utilisateurs, et enfin d'autres qui se positionnent à niveau plus élevé et promeuvent une approche plus globale.

Se protéger des utilisateurs et du détournement de leurs comptes

Mais l'approche centrée sur le comportement des utilisateurs, le *User Behavior Analytics* ou UBA, semble rencontrer actuellement un engouement particulièrement prononcé.

Eric Ahlm apporte un début d'explication : « l'UBA permet de découvrir des activités qui resteraient autrement non détectées. Et des rapports sur les menaces ajoutent à cet intérêt : elles montrent l'importance du rôle des détournements de comptes utilisateurs dans les attaques. Il s'agit d'un vecteur de menace important ».

Javvad Malik, ancien chercheur chez 451 Research devenu évangéliste sécurité chez AlienVault, ajoute à cela la question de la lutte contre les menaces internes : « L'UBA est une technique pour trouver des menaces internes. Ce défi n'est pas nouveau […] mais Edward Snowden est devenu la référence des départements marketing : que se passerait-il pour une entreprise avec un Edward Snowden caché en son sein ? »

Les apports et limites du Big Data

Mais que l'on parle d'UBA ou d'analytique au service de la sécurité, il s'agit dans tous les cas de collecter des données, d'importants volumes de données, et de les analyser. Javvad Malik relève ici le but : « corréler des données variées pour essayer d'identifier un sentiment ou une intention. »

Et la promesse est alléchante : « avec les bons outils et processus, traquer les menaces par anticipation peut

apporter d'importants bénéfices aux entreprises, en particulier lorsque des logiciels malveillants sont impliqués ».

Mais si le concept du Big Data au service de la sécurité n'est pas nouveau, comme le relève Eric Ahlm, tous les obstacles ne sont pas encore levés : « le problème, avec de grandes quantités de données de sécurité, c'est le temps. On peut collecter de grandes quantités de données, la question est de pouvoir les traiter dans un délai raisonnable ».

Qu'est-ce qu'un délai raisonnable, justement ? Le plus proche du temps réel… Car si traiter une requête prend trois jours, c'est trop pour se protéger efficacement… « alors que la brèche peut survenir en moins de 24h ».

Le Machine Learning à la rescousse

La question du temps n'est pas la seule. L'analyse de données repose massivement sur des algorithmes. Et pour Eric Ahlm, « il faut là des mathématiciens ; ce n'est pas un tâche simple » que de construire des algorithmes pertinents. « Et la plupart des équipes de sécurité ne disposent pas de mathématiciens de haut niveau ».

Dans la plupart des cas, donc, les entreprises doivent se contenter d'algorithmes pré-écrits. Et dont on imagine les performances plus ou moins limités, tout du moins dans les capacités d'évolution et d'adaptation.

C'est là qu'Eric Ahlm entrevoit un important apport du *Machine Learning*, « pour rendre ce travail sur les algorithmes plus accessible, plus intuitif ». En attendant, l'analytique de sécurité se trouve donc largement limitée aux grandes organisations capables d'attirer les profils requis. Mais tout cela est appelé à changer.

Une maturité à venir

Javvad Malik et Eric Ahlm soulignent tous les deux le très jeune âge de l'analytique appliquée à la sécurité, notamment « par rapport à la médecine ou l'astronomie, ou le décisionnel ». « Nous commençons juste à gratter la surface de ce qu'il est possible de faire », insiste Eric Ahlm.

Et de s'attendre donc à « observer beaucoup d'innovation » jusqu'à, « pourquoi pas, l'automatisation de certaines fonctions d'investigation ».

Pour l'heure, pour Javvad Malik, de nombreux industriels de la sécurité, mais aussi d'utilisateurs, « observent et attendant que ces technologies fassent leur preuves dans de vastes déploiements complexes ».

Alors il faudra s'attendre à des opérations de consolidation sur le marché : « une fois que la technologie aura été éprouvée, j'imagine [que certains acteurs] constitueront des cibles attractives pour des acquisitions ».

– Valéry Marchive

Comment l'analyse comportementale aide à lutter contre les attaques

En tant que professionnel de la sécurité de l'information, vous avez probablement déjà investi temps et argent pour chercher à comprendre ce qui se passe dans votre environnement.

Vous avez déployé des outils de gestion des logs, de gestion des informations et des événements de sécurité (SIEM), voire même des systèmes de renseignement opérationnel sur les menaces. Mais répondre à la question suivante reste difficile : comment savoir que quelque chose qui survient dans votre environnement ne devrait pas se produire ?

C'est tout l'objet de l'analyse comportementale (UBA, *User Behavior Analytics*). De nombreux éditeurs émergent tels que Bay Dynamics, ClickSecurity, GuruCul, Fortscale et Securonix déploient des techniques de Big Data pour établir rapidement un profil d'activité normal d'un environnement donné et détecter ensuite les anomalies indiquant des attaques. D'autres, comme Lancope, Solera et Splunk étendent leurs offres pour fournir également de telles capacités.

Ce guide se penche sur ce qu'il est raisonnable d'attendre de ces outils analytiques et sur les stratégies de déploiement afférentes.

Comprendre l'analyse comportementale

La dernière édition de la RSA Conference a accordé une place importante à l'UBA. L'utilisation des technologies analytiques est désormais au premier plan des architectures de sécurité. Et il y a une très bonne raison à cela : elles doivent aider à résoudre le problème de l'aiguille dans la botte de foin auquel sont confrontés tous les RSSI. Les outils analytiques les aident à trouver une signification aux vastes volumes de données collectés par les SIEM, les IDS/IPS, les logs, etc.

Les outils d'UBA utilisent des capacités analytiques spécialisées qui se concentrent sur le comportement des systèmes et de leurs utilisateurs. L'analyse comportementale est initialement apparue dans le domaine du marketing, pour aider les entreprises à comprendre et à prédire le comportement des consommateurs. Mais l'UBA peut s'avérer remarquablement utile dans le domaine de la sécurité.

Comment fonctionne l'UBA

Les outils d'UBA assurent deux fonctions principales. Tout d'abord, ils déterminent des comportements normaux pour des activités spécifiques à l'organisation et à ses utilisateurs. Ensuite, les outils d'UBA détectent rapidement des déviations qui requièrent l'attention des RSSI : ils identifient les cas où un comportement anormal est en cours. Ce comportement peut trahir ou non un problème ; ce sera aux équipes de sécurité de le déterminer par leur enquête.

La distinction entre UBA et les autres formes d'outils analytiques de sécurité et que l'UBA se concentre sur les utilisateurs plutôt que sur les événements ou les alertes. Autrement dit, l'UBA répond à la question suivante : l'utilisateur se comporte-t-il de manière anormale ? La distinction est subtile, mais importante : un événement peut être bénin dans un contexte donné, mais grave dans un autre.

Que chercher dans un outil d'UBA ?

De plus en plus de fournisseurs commencent à revendiquer l'intégration de capacités d'UBA dans leurs produits. Mais il n'y a qu'un petit nombre – quoiqu'en croissance – de véritables fournisseurs de solutions d'UBA. Les produits de ces éditeurs fonctionnent grosso-modo de la même manière : un moteur analytique exploitant des algorithmes propriétaires est alimenté par des sources de données existantes et les examinent. Les outils affichent alors leurs découvertes dans un tableau de bord destiné à l'utilisateur. Le but est de fournir aux professionnels de la sécurité des informations immédiatement exploitables.

Pour l'heure, ces outils n'engagent pas eux-mêmes d'actions défensives : ils fournissent à leurs utilisateurs les renseignements nécessaires pour déterminer s'il est nécessaire d'agir ou non. Mais il est raisonnable d'anticiper, d'ici 6 à 24 mois, l'apparition d'outils intégrés avec les systèmes de défense des entreprises, comme les pare-feu, pour automatiser la réponse aux menaces.

Les algorithmes analytiques sont la potion magique de ces outils. Lors de leur évaluation, les RSSI devraient des détails sur leur fonctionnement. Mais il existe d'autres points de différenciation.

Et cela commence par les sources de données supportées par l'outil : formats (CSV, Excel, etc.) et les types de fichiers logs. Il convient là d'interroger sur les capacités natives d'intégration et les possibilités de personnalisation. Le tout en fonction de sa propre infrastructure existante.

Le délai d'établissement des profils comportementaux de base, et son degré d'automatisation, sont également importants. Certains outils déterminent ces profils à partir de seulement quelques jours de données historiques, d'autres préfèrent en utiliser plusieurs semaines, voire mois. Plus l'historique est important, plus les profils comportementaux tendant à être prévis, parce qu'ils tiennent comptes de la variabilité saisonnière des activités.

Le délai de production de résultats se rapporte à la rapidité avec laquelle des résultats exploitables concrètement sont produits après l'intégration initiale. Mais attention : cet indicateur n'est pas aussi évident qu'il peut le paraître : une définition claire des « résultats » est nécessaire. Une bonne définition en est la fourniture d'informations inconnues au préalable.

La flexibilité du tableau de bord est également à prendre en compte. Celui-ci est-il par exemple conçu exclusivement pour des spécialistes de la sécurité ? Ou supporte-t-il des personnalisations permettant d'étendre son audience à des responsables métiers ?

Se pose enfin la question du déploiement, sur site ou en mode Cloud. Dans ce dernier cas, fortement appelé à se développer, se pose la question de la sécurité des données transmises à l'outil d'UBA.

Des données enfin mises à profit

Collecter les données n'est pas suffisant. Il est nécessaire d'investir dans des outils qui permette de trouver des informations et du sens dans ces données, capables de trouver ces indicateurs critiques d'une potentielle compromission : ces aiguilles dans la botte de foin. Les outils d'UBA peuvent fournir très tôt des indications sur des comportements suspects, d'utilisateurs, de systèmes et d'appareils. De quoi orienter les professionnels de la sécurité de manière précieuse.

– Johna Till Johnson, Nermetes Research

AUTEURS

Le document consulté provient du site www.lemagit.fr

Cyrille Chausson | *Rédacteur en Chef*

Valéry Marchive | *Rédacteur en Chef adjoint*

Linda Koury | *Directeur Artistique*

Neva Maniscalco | *Designer*

TechTarget
22 rue Léon Jouhaux, 75010 Paris
www.techtarget.com