LIVRE BLANC | MARS 2017

Sécurité des données d'entreprise : les bases de l'analyse du comportement de l'utilisateur

ca technologies

Table des matières

Résumé

Les cyberattaques font de plus en plus souvent la une des journaux. Et bien que la majorité des attaques de grande envergure (notamment les violations de sécurité majeures chez JP Morgan, Anthem et Slack) n'aient pas eu une origine interne à l'entreprise victime, le vol et l'utilisation malveillante des données par des utilisateurs à forts privilèges sont des incidents de plus en plus fréquents.

En réalité, 69 % des professionnels de la sécurité en entreprise affirment avoir déjà traité des cas de vol ou de corruption des données de l'entreprise par des membres internes « de confiance ».[1] D'autres violations de réseau sont aussi le fait de tiers travaillant avec l'entreprise, notamment des sous-traitants, des fournisseurs ou des partenaires, que ce soit par accident ou dans l'intention de nuire.

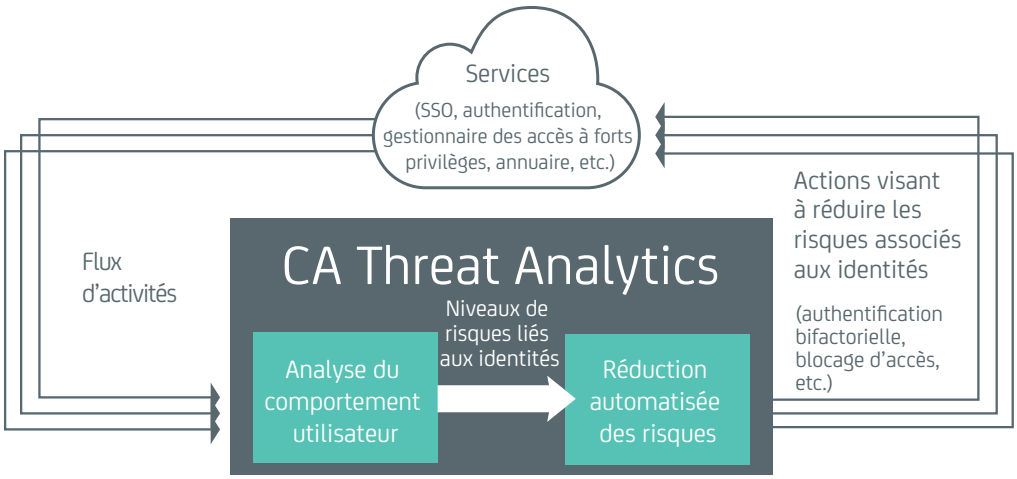
S'il y a une leçon à tirer de ces événements, c'est bien qu'il est urgent de protéger les accès à forts privilèges, et ce quelle que soit la taille de l'entreprise. Cependant, malgré des efforts de sensibilisation et une offre de produits de sécurité de plus en plus large, nombre de systèmes IT sont encore vulnérables.

Cela tient en partie au fait que les contrôles traditionnels utilisés en matière de gestion des accès et des identités (IAM, Identity end Access Management), bien que nombreux, sont tout à fait statiques. Et lorsqu'un utilisateur malveillant a réussi à accéder au système, il est libre d'exploiter celui-ci avec tous les privilèges du compte qu'il a usurpé.

Toutefois, en déployant une approche de sécurité centrée sur l'identité, qui allie une fonction d'analyse du comportement de l'utilisateur à une détection des anomalies dans un modèle à auto-apprentissage, les entreprises pourraient détecter rapidement les activités à risque et déclencher automatiquement des contrôles d'atténuation des risques, afin de limiter les dégâts subis par l'entreprise.

CA Threat Analytics

CA Threat Analytics protège les données de l'entreprise de la même manière que les cartes de crédit protègent votre argent. Bien que cette phrase évoque l'idée générale (une supervision permanente et l'utilisation d'outils d'analyse pour évaluer les risques et empêcher des personnes malintentionnées de dérober les actifs), elle n'offre que peu d'indications sur la façon dont cet objectif est atteint. Ce livre blanc explique comment CA Threat Analytics protège les données de l'entreprise en appliquant deux fonctionnalités liées : l'analyse du comportement de l'utilisateur et l'atténuation automatisée des risques.



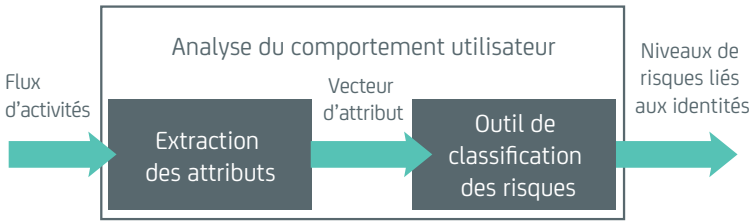
La fonction d'analyse du comportement de l'utilisateur permet à l'entreprise d'évaluer en continu le risque encouru et de détecter rapidement toute activité malveillante. Pour ce faire, elle se base sur un flux de données relatives à l'interaction entre une identité ou un groupe d'identités spécifique et les différents services ou applications, de manière à établir le niveau de risque associé à chaque identité d'entreprise.

La fonction d'atténuation automatisée des risques permet à l'organisation d'entreprendre automatiquement les actions de réduction des risques nécessaires et de contrer toute activité malveillante détectée. Elle change la façon dont les accès sont contrôlés pour chaque identité, en se basant sur le niveau de risque déterminé par les outils d'analyse du comportement utilisateur. Un exemple simple d'atténuation automatisée des risques consisterait à bloquer automatiquement l'accès d'une identité à haut risque pour une application ou un référentiel de données particulièrement sensible.

Bien que les fonctions d'analyse du comportement de l'utilisateur et d'atténuation automatisée des risques fassent toutes deux parties intégrantes du fonctionnement de CA Threat Analytics, ce livre blanc se concentrera volontairement sur l'analyse du comportement. Dans les sections suivantes, nous allons décomposer la fonction d'analyse du comportement de l'utilisateur évoquée ci-dessus afin d'en présenter les différents composants. Nous étudierons ensuite en détail chaque composant. Pour plus de simplicité, nous commencerons la discussion en étudiant la protection d'une seule identité sur un service unique. Après avoir expliqué les notions fondamentales des techniques utilisées, nous expliquerons la façon dont ces principes peuvent être améliorés en travaillant sur une population d'identités à l'échelle de plusieurs services.

Notions fondamentales

D'un point de vue conceptuel, la fonction d'analyse du comportement de l'utilisateur repose sur deux composants : l'extraction des attributs et l'outil de classification des risques.



Le composant d'extraction des attributs traite un flux d'activité et en extrait un ensemble d'attributs pertinents. Le terme « attributs pertinents » désigne les caractéristiques relatives à une identité spécifique ayant été observées dans le temps, notamment les aspects suivants :

- L'identité utilise un périphérique mobile inconnu.
- L'identité est exécutée sur un site distant.
- L'identité provient d'une adresse IP suspecte.
- L'identité fait partie d'un groupe à forts privilèges.
- L'identité a utilisé le service X en-dehors des heures de fonctionnement normales.

Le processus d'extraction des attributs est plus complexe qu'il n'y paraît, car il ne se résume pas à extraire les caractéristiques relatives à une transaction en cours. Bien que le flux d'activité arrive sous la forme d'une séquence d'événements isolés, la véritable source d'informations est le flux d'activité complet, depuis le début de la période. Cela vous permet de comprendre de façon globale (cumulée) l'utilisation et le comportement pour chaque identité. Sans étudier l'historique d'activité complet, vous seriez forcé d'évaluer les risques sur la seule base de chaque événement isolé.

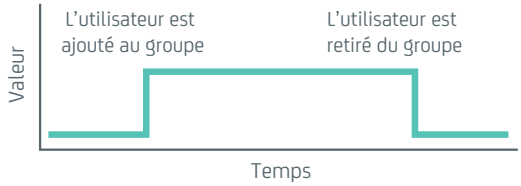
En prenant comme exemple l'un des attributs répertoriés ci-dessus, que signifie « heures de fonctionnement normales » dans le contexte d'un événement isolé ? Pour que CA Threat Analytics puisse se servir de ce type d'attribut important, la solution doit également calculer et utiliser les données d'historique.

En étudiant l'ensemble du flux d'activité, CA Threat Analytics fournit à l'entreprise considérablement plus d'informations que par le passé pour évaluer les risques et détecter les activités malveillantes. Ainsi, l'entreprise peut désormais évaluer les risques en se fondant sur les activités passées et les informations spécifiques à chaque identité. Cet avantage a toutefois un coût : la quantité importante de données à traiter, dont une bonne partie est redondante. Heureusement, grâce à la fonction d'extraction des attributs, la dimensionnalité des données est réduite. Les données redondantes peuvent être éliminées ou agrégées, tout en en mettant en lumière les informations nécessaires pour la seconde partie de la fonction d'analyse du comportement de l'utilisateur, à savoir la classification des risques.

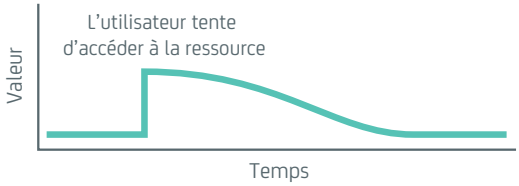
Déterminer la valeur dans un contexte temporel

Avant de continuer, intéressons-nous à un détail intéressant concernant les attributs observés sur la durée. Du fait qu'ils sont modifiés lorsqu'une activité se produit, ces attributs « vivent » théoriquement dans le domaine temporel. Autrement dit, leur valeur change au fil du temps. Lorsqu'un attribut est observé, CA Threat Analytics modélise l'observation en tant que fonction temporelle. En d'autres termes, si une activité entrante active un attribut, la « valeur » de cet attribut peut être à son maximum au moment où cette activité a lieu et changer à mesure que le temps passe.

La façon dont cette valeur change varie grandement suivant l'attribut extrait. Certains attributs sont complètement binaires, de sorte que quand ils sont observés, ils restent à leur valeur maximale jusqu'à ce qu'un événement viennent l'abaisser, comme illustré ci-dessous.



Il peut s'agir, par exemple, de l'attribut d'appartenance à un groupe sensible. Cet attribut est à pleine valeur sur la totalité de la période durant laquelle l'identité est associée au groupe. D'autres attributs sont modélisés sous forme d'impulsions déclinantes. Lorsqu'un attribut de ce type et observé, la valeur est au plus haut, puis descend progressivement avec le temps, comme illustré ci-dessous.



Il peut s'agir, par exemple, des tentatives d'un utilisateur d'accéder à une ressource pour laquelle il n'a pas d'autorisation. Bien que cet attribut soit pertinent aujourd'hui pour définir le niveau de risque d'une identité, il le sera moins dans une semaine et encore moins dans un mois. En réduisant progressivement la valeur des attributs dans le temps, CA Threat Analytics garantit que ces attributs contribuent à évaluer le risque de la façon la plus pertinente possible.

Outil de classification des risques

L'outil de classification des risques assure une fonction analytique qui convertit le vecteur d'attribut en trois niveaux de risque distincts :

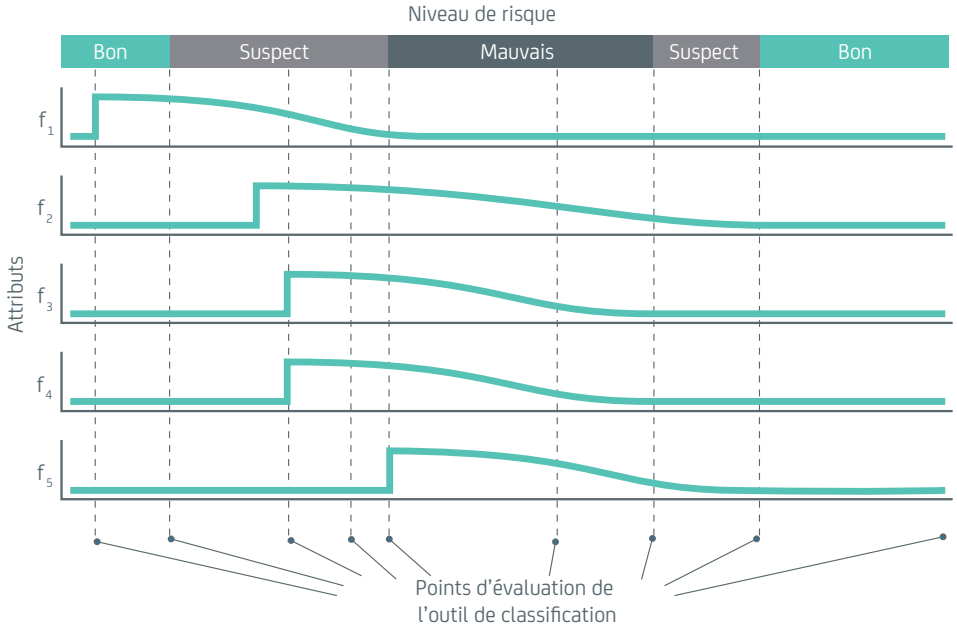
- **Bon** : l'identité présente un risque minime.
- **Suspect** : l'identité a été associée à des événements ou à des activités présentant des risques, mais ces risques n'exigent aucune action immédiate. Le système surveillera cette identité plus étroitement et exécutera éventuellement un ensemble initial d'atténuation automatisée des risques, suivant la politique de l'entreprise en la matière.
- **Mauvais** : l'identité est considérée à haut risque et exige une attention immédiate. Le système lance un processus d'alerte et d'atténuation automatisée des risques, conformément à la politique de l'entreprise.

L'outil de classification des risques utilise comme données d'entrée un vecteur des valeurs d'attribut et génère en sortie l'une des classes distinctes indiquées ci-dessus.

$$\text{Outil de classification } \overrightarrow{(\text{attributs}(t))} \longmapsto \{\text{bon, suspect, mauvais}\}$$

Comme nous l'avons expliqué précédemment, les attributs eux-mêmes sont fonction du temps ; l'outil de classification des risques agit donc également dans le domaine temporel. Il est appelé lors des points de décision critiques, généralement en réponse à des changements majeurs dans les valeurs du vecteur d'attribut. Lorsque l'outil de classification des risques calcule le niveau de risque pour un point donné dans le temps, toutes les fonctions de l'attribut sont évaluées pour l'identité ou l'entité concernée à ce moment précis. L'ensemble d'attributs actifs pour l'entité à cet instant T composent le vecteur d'attribut réel, que l'outil de classification des risques utilise pour évaluer le risque.

L'illustration ci-dessous présente les différents points auxquels l'outil de classification des risques effectuerait probablement son évaluation. Comme indiqué, l'évaluation a lieu lorsque la valeur d'un attribut augmente ou lorsqu'elle descend en dessous d'un seuil donné. Les valeurs prises en compte par l'outil de classification des risques correspondent à la valeur de chaque attribut au moment précis où l'évaluation se déclenche, illustré ci-dessus par les lignes verticales. Bien entendu, chaque cycle d'exécution de l'outil de classification des risques ne donne pas forcément un nouveau niveau de risque. En pratique, il existe bien plus de points d'évaluation que ceux illustrés ici. Il peut notamment s'agir de changements de valeur d'un attribut, d'une activité système ou de renseignements sur les menaces. En général, l'outil de classification des risques s'active dès qu'il pourrait y avoir un changement de niveau de risque.



Découvrons maintenant en quoi consiste exactement cet outil de classification des risques. Comment celui-ci traduit-il un vecteur d'attribut en une classe de risque ? Pour le savoir, il est intéressant de commencer par déterminer ce que cet outil n'est pas. Les outils de classification des risques de CA Threat Analytics ne sont pas de simples règles qui testent des attributs spécifiques, par exemple « si l'attribut X est actif, renvoyer Mauvais ». Il s'agit là d'une approche naïve appliquée par de nombreux produits de sécurité traditionnels. Cette approche échoue d'ailleurs de façon spectaculaire, car non seulement elle génère de très nombreux faux positifs, mais en plus, elle est fragile et facilement contournée. En outre, elle ne fait pas usage des informations critiques aussi bien pour détecter les activités malveillantes que pour rendre le système utilisable et convivial pour les utilisateurs légitimes.

Les fonctionnalités de CA Threat Analytics sont bien plus solides. L'outil de classification des risques de CA Threat Analytics examine les attributs non pas de façon isolée, mais dans le contexte de l'ensemble des attributs. Avec cette approche, il est possible de combiner plusieurs attributs, qui pris isolément, n'auraient aucun impact sur le niveau de risque, mais ensemble, permettent d'évaluer les risques de façon pertinente. De plus, CA Threat Analytics intègre le feed-back provenant des systèmes déployés, y compris différents aspects des utilisateurs et les changements dans la population d'identités, de manière à affiner ses décisions sur la durée. Il en résulte un système suffisamment flexible pour s'adapter aux nouveaux scénarios de déploiement et aux menaces émergentes.

Populations et services

Comme mentionné précédemment, plusieurs détails pratiques ont été simplifiés dans notre présentation ci-dessus. Tout d'abord, qu'en est-il des populations d'identités ? Dans un environnement d'entreprise tout particulièrement, certains aspects du groupe d'identités sont pertinents pour établir le niveau de risque d'une identité donnée. Voici quelques exemples :

- Accéder aux ressources avec davantage de périphériques qu'il n'est normal pour l'entreprise
- Engager des actions depuis un lieu autre que le site d'exploitation normal de l'entreprise
- Appartenir à un nombre anormalement élevé de groupes

La référence de base varie pour chaque entreprise en ce qui concerne les activités prévues et acceptables (ce qui inclut des facteurs tels que le nombre normal de périphériques associés à un utilisateur, les sites d'exploitation de l'entreprise ou le nombre approprié de groupes). En étudiant un groupe d'identités plutôt que chaque identité isolée, vous obtiendrez de nombreuses statistiques de population très utiles sur la base desquelles évaluer les identités individuelles. Bien entendu, tout cela a un coût. Au lieu de traiter le flux d'activité dans son ensemble pour une identité, cela impose d'effectuer une extraction des attributs sur tout l'historique d'activité et à l'échelle de l'entreprise.

De même, passer de l'analyse d'un seul service à l'analyse d'un groupe de services offre plusieurs avantages. En étudiant les actions qu'une identité engage sur différents services, nous sommes à même d'extraire des attributs pour créer des modèles d'accès type, que nous appliquons ensuite intelligemment pour assurer la sécurité de l'ensemble des services. Ces informations permettent à CA Threat Analytics de détecter les comportements anormaux et incohérents qui menacent cette identité en particulier ou l'entreprise dans son ensemble.

Conclusion

Ce livre blanc explique comment CA Threat Analytics protège les données de l'entreprise par le biais d'une analyse du comportement de l'utilisateur. Bien que les concepts de base d'une telle approche soient faciles à expliquer, les problématiques pratiques associées à l'extraction des attributs et à la classification des risques vont bien au-delà du champ d'application de ce document. Nous n'y avons d'ailleurs pas fait mention de nombre de situations réelles que gèrent nos équipes, notamment permettre une prise de décision en temps réel, garantir la précision du système dans le temps et fournir aux administrateurs système des informations pertinentes concernant les décisions en matière de risques.

Si vous souhaitez en savoir plus sur ces sujets et découvrir comment votre entreprise peut en tirer parti, consultez la page : https://www.ca.com/fr/products/ca-threat-analytics-for-privileged-access-manager.html

1 Accenture et HfS Research, « The State of Cyber Security and Digital Trust 2016 », juin 2016 : https://www.accenture.com/t20160704T014005__w__/us-en/_acnmedia/PDF-23/Accenture-State-Cybersecurity-and-Digital-Trust-2016-Report-June.pdf#zoom=50

CS200-259092