VECTRA
SECURITY THAT THINKS

LIVRE BLANC

L'IA qui sous-tend les solutions Vectra

SCIENCE DES DONNÉES
RECHERCHE EN SÉCURITÉ

NATIVE AU CLOUD
AUTOMATISÉE

SOMMAIRE

Vectra® protège les entreprises en détectant et en bloquant les cyberattaques.

Vectra® est le leader de la détection et de l'aide à la résolution des menaces pour les entreprises hybrides et multiclouds. La plate-forme Vectra utilise l'IA pour détecter rapidement les menaces au niveau du cloud public, des identités, des applications SaaS et des centres de données. Vectra ne se contente pas de signaler les stratégies différentes de celles employées habituellement. C'est le seul fournisseur qui optimise l'IA pour détecter les méthodes d'attaque utilisées par les cyberpirates, c'est-à-dire leurs tactiques, techniques et procédures (TTP). Sa plate-forme offre des signalements efficaces et des données contextuelles claires qui permettent aux équipes de sécurité d'intervenir sur les incidents plus tôt et de bloquer les attaques en cours plus vite. Les organisations du monde entier s'appuient sur Vectra pour assurer leur résilience face aux cybermenaces les plus graves et pour empêcher les ransomwares, les compromissions de la chaîne logistique, les usurpations d'identité et autres cyberattaques de nuire à leur activité. Pour plus d'informations, rendez-vous sur le site vectra.ai.

Introduction

La science des données est la boussole de l'IA que Vectra utilise dans ses solutions. Nous avons depuis longtemps la conviction qu'utilisées correctement, la science des données et l'IA sont des armes redoutables dans notre lutte contre les cyberattaques et représentent un atout considérable pour les équipes de sécurité. Cependant, l'IA présente une multitude de facettes. Dans ce livre blanc, nous expliquons ce qu'est l'IA, nous en définissons la terminologie, nous décrivons les deux principales méthodologies permettant d'appliquer l'IA à la détection des menaces et, enfin, nous présentons en détail la façon dont Vectra détecte les menaces grâce à l'IA.

Que vous éprouviez de la fascination pour l'intelligence artificielle ou que vous n'y voyiez pas grand intérêt, ce livre blanc est pour vous.

Qu'est-ce que l'IA ?

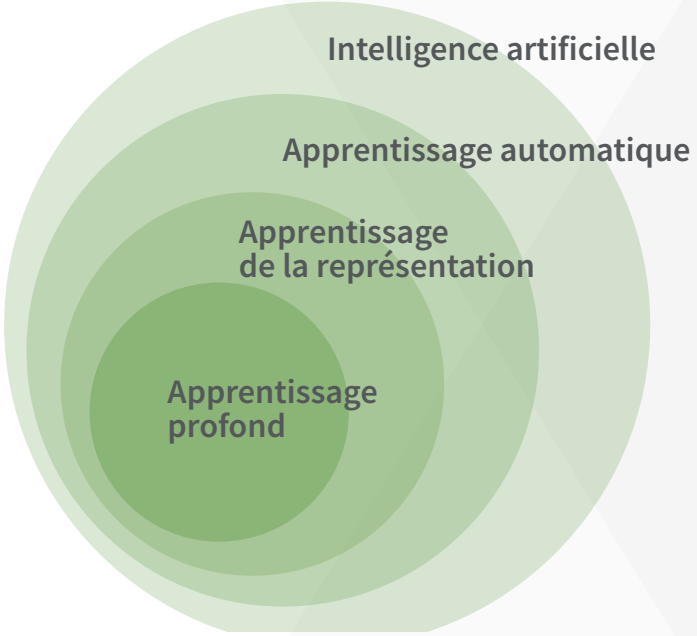
Définition de l'IA

Les termes « intelligence artificielle », « apprentissage automatique » et « apprentissage profond » sont souvent mal compris et assimilés à la même discipline ou associés à des niveaux de qualité différents. Ce n'est cependant pas le cas. Si ces termes sont liés les uns aux autres, chacun possède sa signification spécifique et distincte. Comprendre le sens de chaque peut aider à mieux cerner le fonctionnement des outils qui exploitent l'IA.

Intelligence artificielle (IA) : système capable d'automatiser le raisonnement et d'imiter le fonctionnement de l'esprit humain. Ce terme général englobe les sous-disciplines de l'apprentissage automatique, de l'apprentissage de la représentation et de l'apprentissage profond. Le terme d'IA recouvre de façon égale les systèmes qui s'appuient sur l'utilisation de règles explicitement programmées et ceux qui ont tiré de façon autonome des renseignements exploitables de grands volumes de données. La forme la plus récente de l'IA, qui entre dans la sous-discipline de l'apprentissage automatique et base son apprentissage sur les données, sous-tend des technologies telles que les voitures autonomes et les assistants virtuels.

Apprentissage automatique : sous-discipline de l'IA où les actions du système ne sont pas explicitement dictées par un humain, mais apprises à partir de données. Ces systèmes sont capables de traiter des dizaines de milliards de points de données afin d'apprendre à représenter de façon optimale les nouvelles instances de données et y réagir.

Apprentissage de la représentation : sous-discipline de l'IA peu connue, mais qui est centrale dans de nombreuses technologies d'IA utilisées aujourd'hui. Elle est axée sur l'apprentissage d'une nouvelle représentation abstraite à partir de données. Un exemple d'apprentissage de la représentation est la transformation d'images de différentes tailles en une liste de nombres de longueur cohérente, qui représentent une distillation des images d'origine. Cette abstraction permet avant tout aux systèmes en aval d'agir plus efficacement sur de nouveaux types de données.



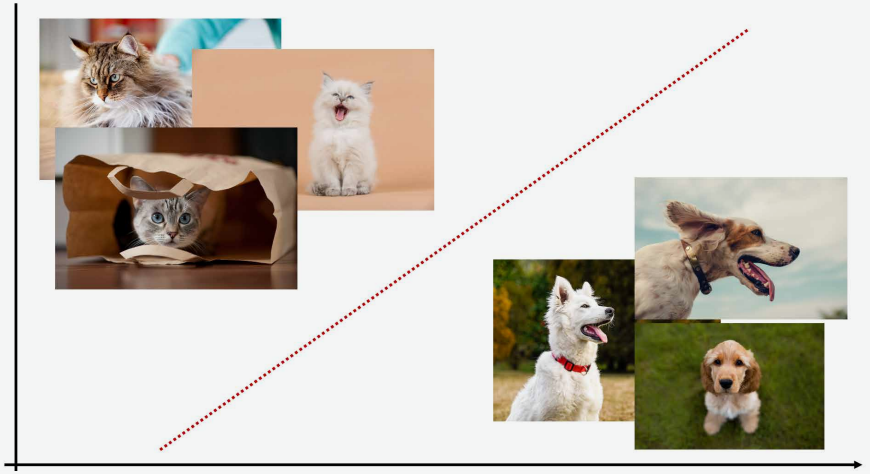
Relation entre les différentes sous-disciplines de l'IA.
Références : « Deep Learning », Goodfellow, Bengio et Courville (2016)

Apprentissage profond : souvent associé aux réseaux neuronaux, ce type d'apprentissage va plus loin que les sous-disciplines généralistes de l'apprentissage automatique et de l'apprentissage de la représentation. Il permet d'extrapoler à partir de données une hiérarchie d'abstractions qui représente les entrées sous une forme de plus en plus complexe. S'inspirant du cerveau humain, les modèles d'apprentissage profond exploitent les couches de neurones dont le coefficient synaptique s'adapte aux entrées. Les couches les plus profondes du réseau apprennent de nouvelles représentations abstraites qui simplifient des tâches telles que catégoriser une image ou traduire un texte. Si l'apprentissage profond peut être une technique efficace pour résoudre certains problèmes complexes, ce n'est pas la solution idéale pour automatiser l'intelligence.

Les différentes techniques d'apprentissage automatique

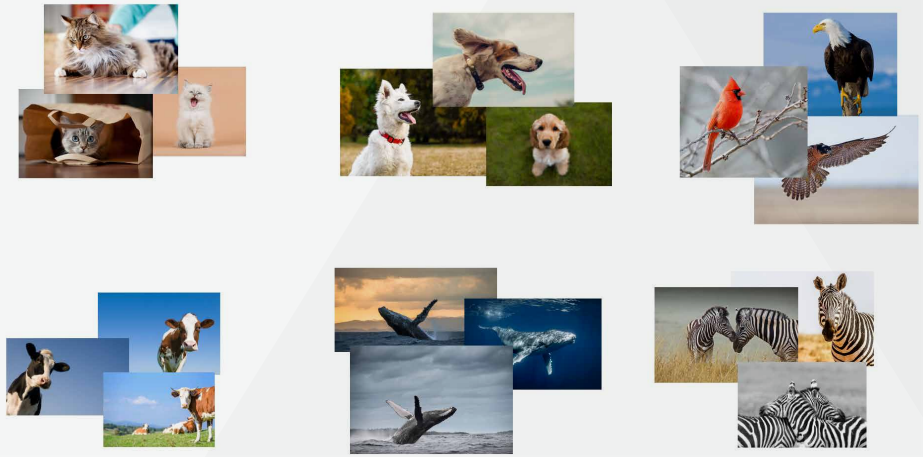
L'une des capacités clés des algorithmes d'apprentissage automatique est de classifier des instances de données dans différentes catégories. Plusieurs grandes catégories d'apprentissage sont compatibles avec cette capacité, les deux principales étant **l'apprentissage supervisé** et **l'apprentissage non supervisé**.

Dans l'apprentissage **supervisé**, le modèle apprend à partir d'un ensemble de données étiquetées. Une fois l'apprentissage effectué, le modèle peut prédire des étiquettes lorsque de nouvelles données sont ajoutées. Prenons un exemple. Si nous alimentons un modèle d'apprentissage supervisé avec un grand nombre d'images de chats et de chiens, il sera capable de prédire si une nouvelle image est celle d'un chat ou d'un chien. L'apprentissage supervisé exige un corpus volumineux de données d'entraînement étiquetées pour entraîner les modèles. Mais, une fois entraînés, ces modèles pourront catégoriser et étiqueter les nouvelles instances de données de façon très efficace.



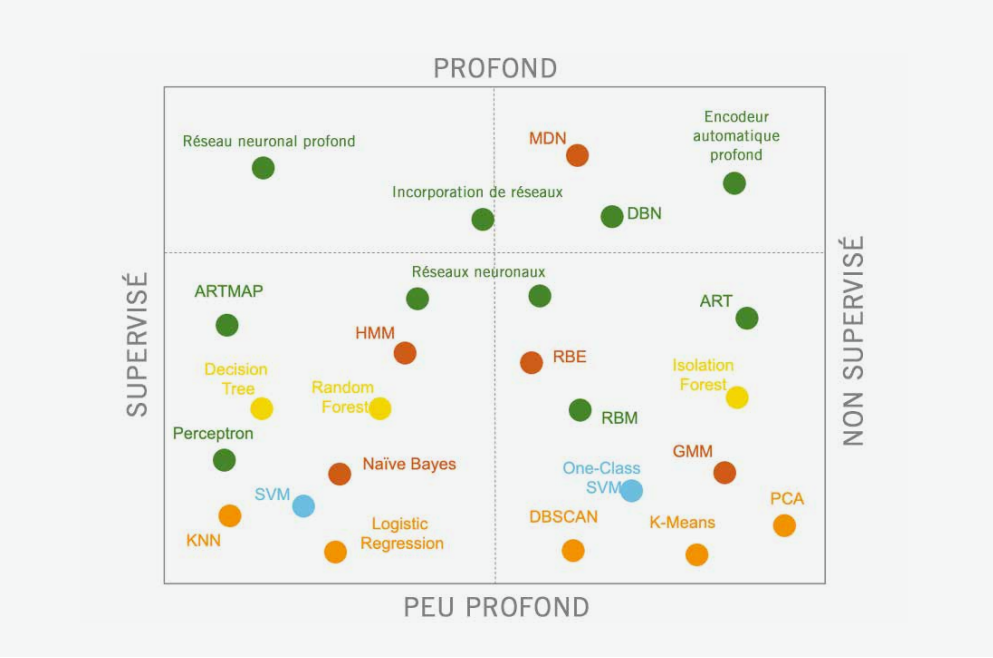
L'apprentissage supervisé utilise des données étiquetées pour identifier les facteurs qui distinguent les différentes étiquettes. Les modèles qui parviennent à exploiter ce type d'apprentissage sont capables d'étiqueter de nouvelles données.

Dans l'apprentissage **non supervisé**, le modèle apprend à partir d'un ensemble de données non étiquetées. Ces modèles apprennent la structure des données fournies, et sont ensuite capables de déterminer si et comment de nouvelles données s'intègrent à cette structure apprise. Les modèles d'apprentissage non supervisés présentent l'avantage de ne pas nécessiter d'entraînement préalable. Cette approche est très performante dans la détection des points de données anormaux, mais ne permet pas d'étiqueter facilement ces valeurs aberrantes.



L'apprentissage non supervisé porte sur la structure sous-jacente des données non étiquetées. Les modèles qui parviennent à exploiter cet apprentissage sont capables de mesurer le degré de correspondance des nouvelles données avec une structure apprise.

Ces approches générales incluent une variété d'algorithmes d'apprentissage, comme illustré ci-dessous, et les chercheurs continuent d'en créer de nouveaux. Pour compliquer davantage les choses, les algorithmes peuvent être combinés pour former des systèmes encore plus complexes. La question qui se pose alors est : comment fait un scientifique des données pour choisir l'algorithme ou les algorithmes appropriés pour résoudre un problème spécifique ? Est-ce qu'un algorithme peut être supérieur à tous les autres quel que soit le problème ?

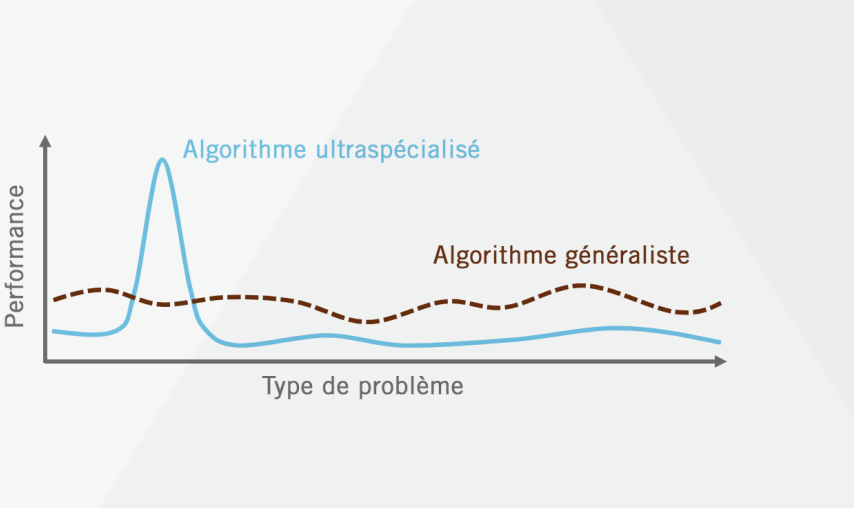


Il existe un grand nombre d'algorithmes d'apprentissage automatique, chacun présentant des points forts et des points faibles selon le type de problème à résoudre.

Théorème du « No Free Lunch »

Il n'existe aucun algorithme générique plus performant que les autres pour résoudre tous les problèmes possibles. Ce principe s'appelle « théorème du "no free lunch" ». En bref, pour tout problème donné, il y aura toujours un algorithme spécialisé qui fonctionnera mieux que les algorithmes génériques pour le résoudre. De ce besoin d'algorithmes spécifiques pour répondre à des problèmes spécifiques découle le besoin de produire toujours plus d'algorithmes, comme expliqué ci-dessus. Dans certains problèmes, un réseau neuronal supervisé offrira les meilleures performances ; dans d'autres, un partitionnement de données hiérarchique non supervisé fonctionnera mieux.

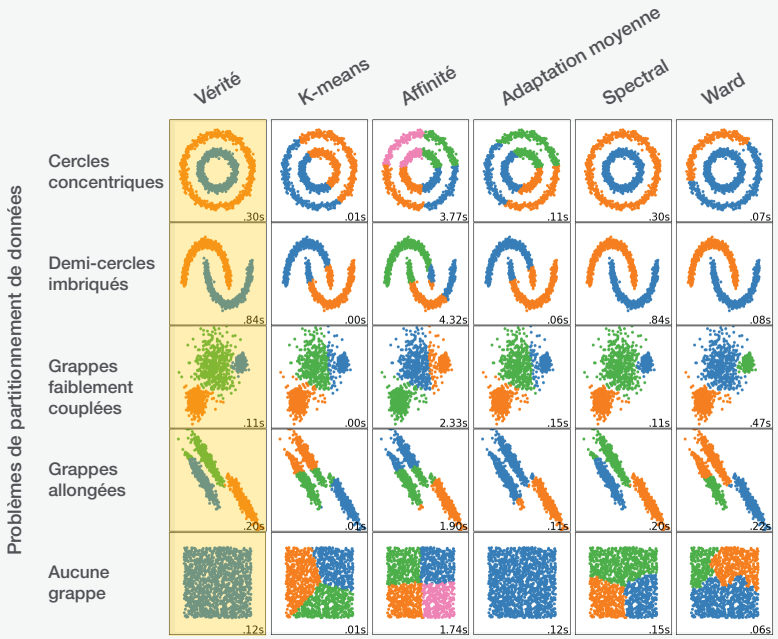
Par exemple, l'algorithme utilisé pour la reconnaissance d'images dans les voitures autonomes ne peut pas être appliqué pour traduire un texte. Chaque algorithme représente un choix spécifique, optimisé pour le problème à résoudre et les données qui servent à entraîner le modèle.



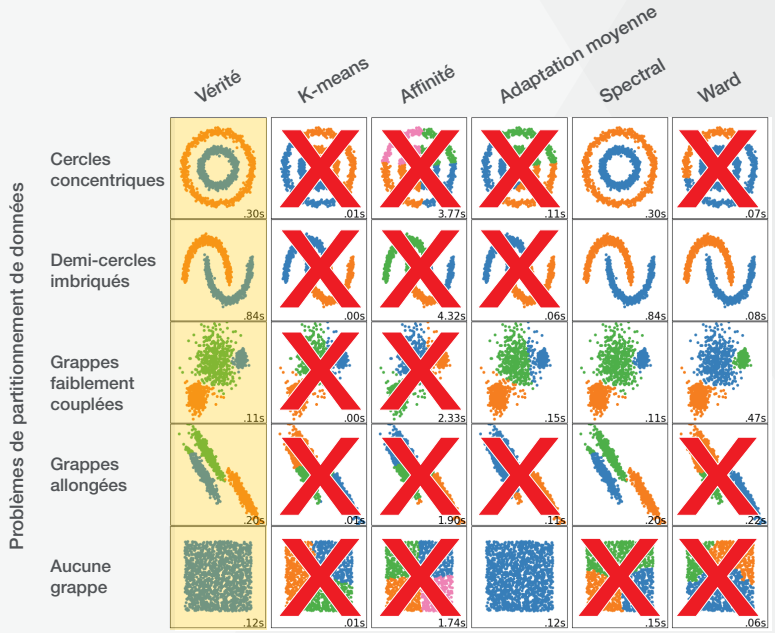
Théorème du « No Free Lunch » : aucun algorithme n'est plus performant que les autres pour résoudre un problème.

Trouver l'outil adapté à chaque tâche

Dans ce contexte, comment un scientifique des données choisit-il l'algorithme approprié ? En faisant appel aussi bien à la créativité qu'à la science. La paire « énoncé du problème/compréhension approfondie des données » peut le diriger vers la meilleure décision. Cependant, il est important de bien comprendre que tout mauvais choix peut donner des résultats pas seulement bancals, mais complètement erronés. Prenons l'exemple illustré ci-dessous. Pour chaque ensemble de données, le choix de l'algorithme donnera des résultats très différents. Il existe un algorithme idéal pour chaque problème, mais, et c'est très important, certains choix généreront des résultats vraiment mauvais. Il est donc fondamental d'adopter la bonne approche pour un problème donné.



Comparaison des résultats d'algorithmes d'apprentissage automatique (axe des x) avec différents ensembles de données (axe des y). Les bonnes étiquettes sont indiquées en jaune. Exemple adapté de scikit-learn.org.



Comparaison des résultats. Un X est placé sur les prédictions erronées qui généreraient des résultats non souhaités. Aucun algorithme n'est efficace pour tous les ensembles de données. Exemple adapté de scikit-learn.org.

Comment savoir si un algorithme est *adapté* ?

Pour choisir le bon modèle, les scientifiques des données disposent d'une méthode plutôt fiable qui est de décider comment en mesurer les performances. Lorsqu'on parle de performances du modèle, l'*exactitude* est un critère souvent mentionné.

$$Exactitude = \frac{(Vrais\ positifs + Vrais\ négatifs)}{(Vrais\ positifs + Vrais\ négatifs + Faux\ positifs + Faux\ négatifs)}$$

L'exactitude en tant qu'indicateur présente un certain intérêt, mais ne permet pas de révéler les véritables performances d'un modèle. Prenons l'exemple d'un problème de classification où l'objectif est d'apposer l'étiquette A ou l'étiquette B sur des données. Si l'étiquette A est utilisée 1 000 fois plus que l'étiquette B, vous obtiendrez facilement 99,9 % d'exactitude en ne faisant apparaître que les données portant l'étiquette A. Cette méthode offre donc une exactitude élevée, mais aucune donnée ne comportera jamais l'étiquette B. Si notre but est de trouver aussi des données portant l'étiquette B, il est clair que l'exactitude n'est pas l'indicateur approprié. Heureusement, les scientifiques des données disposent d'autres indicateurs qui leur permettent d'optimiser et de mesurer l'efficacité d'un modèle pour les cas qui les intéressent.

La précision en fait partie. Cet indicateur mesure si un modèle parvient à prédire une étiquette particulière par rapport au nombre total de prédictions qu'il formule.

$$Précision = \frac{Vrais\ positifs}{(Vrais\ positifs + Faux\ positifs)}$$

Les scientifiques des données qui visent un score de précision élevé doivent donc créer un modèle capable de prédire des étiquettes sans générer de nombreux faux positifs. Cependant, la précision ne nous dit pas si le modèle est réellement parvenu à étiqueter les cas qui nous intéressent. Un autre indicateur, le rappel, nous donne davantage de perspective à ce propos.

Le rappel mesure la fréquence à laquelle un modèle parvient à trouver une étiquette spécifique par rapport à toutes les instances de cette étiquette.

$$Rappel = \frac{Vrais\ positifs}{(Vrais\ positifs + Faux\ négatifs)}$$

Les scientifiques des données qui souhaitent obtenir des scores élevés de rappel doivent créer des modèles qui parviendront sans faille à détecter les instances qui les intéressent.

En effectuant le suivi de la précision et du rappel, et en trouvant l'équilibre entre les deux, les scientifiques des données peuvent mesurer et optimiser efficacement leurs modèles afin d'obtenir des performances élevées.

Pour choisir le bon modèle, les scientifiques des données disposent d'une méthode plutôt fiable qui est de décider comment en mesurer les performances.

Appliquer l'IA à la détection des menaces

L'IA et ses nombreuses disciplines ont un rôle majeur à jouer pour détecter et bloquer les cyberattaquants sur les réseaux d'entreprise modernes. Deux paradigmes permettent l'identification active des menaces de cybersécurité : l'un axé sur les maths, l'autre sur la sécurité. Dans cette section, nous allons détailler les différences entre les deux et expliquer pourquoi l'IA axée sur la sécurité offre de meilleurs résultats.

IA axée sur les maths : une approche insuffisante de la détection des menaces

Dans le paradigme axé sur les maths, les scientifiques des données génèrent des ensembles de statistiques simples à l'aide d'un nombre limité d'algorithmes génériques dédiés à la détection des valeurs aberrantes ou des nouveautés. Les chercheurs en sécurité combinent ensuite ces statistiques pour créer des centaines de règles statistiques. Si une nouvelle statistique est requise, la même approche générique est utilisée pour la créer. On ajoute souvent des filtres de suppression explicites en post-traitement à ces règles statistiques afin de gérer les volumes de détection supplémentaires que ce type d'approche occasionne (cf. le théorème du « free lunch » : les algorithmes génériques donnent des performances médiocres).

Paradigme axé sur les maths pour le développement de détections.

Prenons pour exemple la détection d'un canal C&C (Command & Control). L'équipe de scientifiques des données doit commencer par générer une statistique concernant la prévalence de tous les domaines externes. L'équipe de chercheurs en sécurité doit ensuite définir le seuil de prévalence permettant de détecter le canal C&C. Si un grand nombre de domaines utilisés par les équipements IoT est au-dessus du seuil de prévalence, un filtre de suppression doit être appliqué pour ignorer tous les équipements IoT. Des filtres de suppression supplémentaires doivent également être appliqués aux agents utilisateur, sous-réseaux et autres attributs jusqu'à l'obtention d'un volume d'alertes gérable. La nature générique de cette approche exige la présence de ces filtres de suppression, bien qu'ils présentent le risque de passer à côté d'une technique de contournement utilisée par un cyberattaquant.

IA axée sur la sécurité : couverture maximale, bruit minimum dans la détection des menaces

Le paradigme axé sur la sécurité est une approche qui tente de trouver le juste équilibre entre définir le problème (la méthode d'attaque) et trouver le modèle approprié. Les chercheurs en sécurité définissent le problème en identifiant une méthode d'attaque générale, pas juste un outil ou un exploit unique. De leur côté, les scientifiques des données déterminent l'algorithme approprié pour identifier cette méthode. Les deux équipes travaillent en étroite collaboration à la recherche de la solution. Cette approche permet de détecter directement la méthode d'attaque, et pas seulement les anomalies superficielles souvent constatées dans l'approche axée sur les maths.

Globalement, une approche axée sur la sécurité offre de meilleures performances, mesurées par les indicateurs de rappel et de précision. Elle offre en outre une capacité de résilience face aux changements apportés aux outils d'attaque et nécessite moins de types de détection, ce qui simplifie les opérations pour les équipes de sécurité. Lorsqu'une nouvelle méthode d'attaque montre une tendance à la hausse, le processus axé sur la sécurité démarre et une nouvelle détection est créée. Si la sophistication de cette approche peut nécessiter du temps de développement supplémentaire, les méthodes d'attaque sont très lentes à changer, et continuent d'apparaître en parallèle de méthodes plus anciennes et déjà bien connues.

Paradigme axé sur la sécurité pour le développement de détections.

Fonctionnement de Vectra

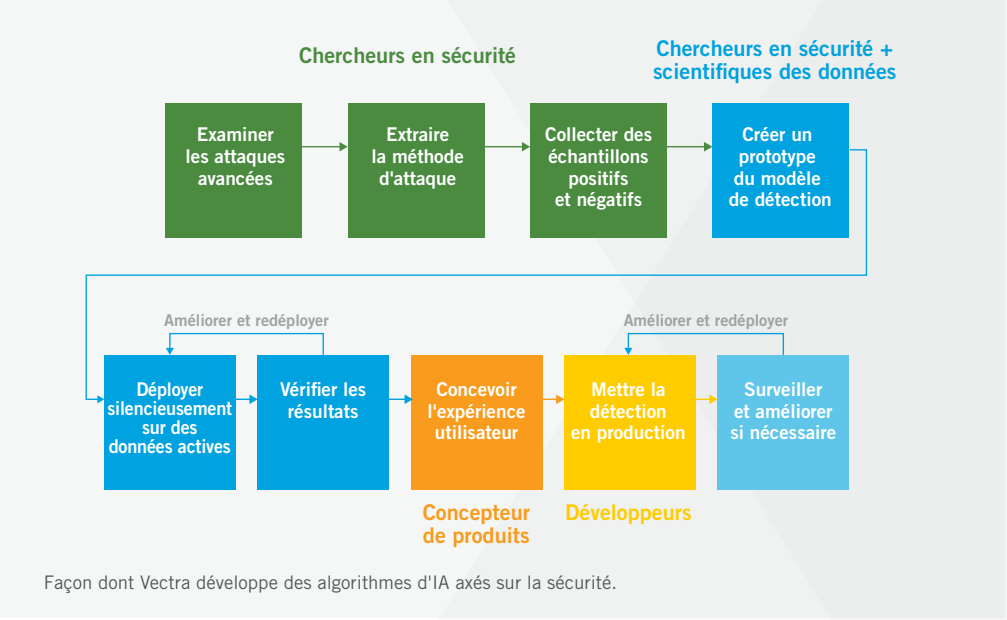
Vectra est un pionnier de l'approche axée sur la sécurité appliquée à la détection des méthodes d'attaque sur les réseaux ainsi que dans le cloud public, les applications SaaS et les identités. Dans les sections qui suivent, nous allons détailler l'étendue de la couverture et du processus de développement de Vectra, le moteur qui collecte et génère les détections, la corrélation entre des événements individuels dans des incidents de sécurité exploitables et le fonctionnement interne de deux détections Vectra.

Développement de la détection chez Vectra

Les détections Vectra sont explicitement dédiées à la découverte des cyberattaquants et à l'identification des méthodes d'attaque en cours, pas uniquement des anomalies. La couverture est élaborée par des chercheurs en sécurité de différents milieux et par des scientifiques des données experts dans l'extraction de valeur d'ensembles de données volumineux et complexes. Cela fait plus de 10 ans que ces deux groupes développent une approche très collaborative du développement de la détection des menaces, qui s'applique à plusieurs domaines de sécurité et types de données afin de détecter efficacement les comportements des cyberattaquants avec un minimum de bruit.

L'équipe de chercheurs en sécurité de Vectra est présente tout au long du processus de développement. C'est elle qui dirige le processus tout en surveillant et en examinant constamment les méthodes d'attaque utilisées dans la réalité. Les recherches ne sont pas axées sur des outils ou des groupes d'attaque spécifiques, mais sur les méthodes générales que les cyberattaquants exécutent. Par exemple, les chercheurs en sécurité peuvent observer l'utilisation de balises Cobalt Strike dans des incidents de ransomware.

Vectra est un pionnier de l'approche axée sur la sécurité appliquée à la détection des méthodes d'attaque sur les réseaux ainsi que dans le cloud public, les applications SaaS et les identités.



Façon dont Vectra développe des algorithmes d'IA axés sur la sécurité.

Au lieu de rechercher uniquement les balises Cobalt Strike, ils extraient les actions de cette technologie et étudient la méthode de *contrôle* du cyberattaquant. En se concentrant sur la méthode abstraite, Vectra peut protéger les outils actuels dont on sait qu'ils exécutent cette méthode et les outils futurs.

Une fois la méthode d'attaque identifiée, les chercheurs en sécurité s'efforcent de collecter un corpus d'échantillons malveillants et inoffensifs. Les échantillons malveillants sont recueillis à partir de plusieurs sources : clients qui partagent volontairement des métadonnées anonymisées, algorithmes de création de données synthétiques, cyberincidents documentés publiquement, attaques dans nos laboratoires internes. Les échantillons inoffensifs sont récupérés auprès de l'ensemble de données volumineux de Vectra, composé de métadonnées clients anonymisées.

Après avoir détecté la méthode d'attaque et collecté les données correspondantes, les chercheurs en sécurité travaillent avec l'équipe de scientifiques des données pour développer un prototype de modèle doté d'un seuil optimal de détection des méthodes d'attaque. Déployé en mode bêta silencieux, ce prototype s'exécute en arrière-plan et ramène des informations de synthèse à partir d'une large base de clients qui ont consenti à partager leurs données. Pour garantir l'efficacité maximale du modèle final, le prototype signale toutes les instances d'une méthode d'attaque observée et toutes les instances des incidents qui ressemblent à la méthode d'attaque – c'est-à-dire des événements qui se déroulent juste en dessous du seuil. Ce déclenchement en dessous d'un seuil permet aux scientifiques des données d'affiner leurs modèles et de s'assurer qu'aucun comportement n'est manqué. Ces modèles sont rapidement itérés jusqu'à ce que leurs performances de détection des méthodes d'attaque constatées dans la réalité répondent à des standards stricts de qualité.

Les étapes finales du développement de la détection impliquent la création d'une interface utilisateur dédiée qui affiche la totalité du contexte de la méthode d'attaque identifiée et, le cas échéant, d'autres informations sur les événements considérés comme normaux sur les systèmes en question. Ces modèles sont ensuite déployés en production où ils fonctionnent et signalent des incidents aux clients. Le pipeline de prototype utilisé pour collecter des données sert également à contrôler l'efficacité du modèle dans le monde réel et, si nécessaire, à améliorer davantage encore la détection.

L'avantage de ce travail est que les modèles ne nécessitent pas d'ajustements fréquents. En outre, ils sont efficaces sur les outils d'attaque existants et futurs. L'approche axée sur la sécurité de Vectra offre d'excellentes performances dans la détection des actions des cyberattaquants, pas seulement des événements « bizarres ».

Moteur de diffusion en temps réel pour des résultats exploitables

La vitesse de détection est fondamentale. Tout retard dans les alertes offre aux cyberattaquants la possibilité d'aller plus loin dans leur attaque. Les algorithmes de Vectra s'exécutent sur des données diffusées et non sur des lots périodiques. Ainsi, Vectra peut détecter les cyberattaquants sans délai, ce qui donne amplement le temps de bloquer leur progression.

L'échelle à laquelle l'opération est menée compte, car l'empreinte des réseaux d'entreprise, des déploiements cloud et des services SaaS augmente constamment, ce qui génère de plus en plus de données à traiter par la fonctionnalité de détection de Vectra. Avec son moteur de diffusion en temps réel, Vectra peut aider les plus grandes multinationales à extraire les données nécessaires à un apprentissage à long terme, sans problèmes liés à la taille des données.

L'efficacité des algorithmes, et en particulier de ceux qui utilisent l'apprentissage non supervisé, est très impactée par la quantité de données d'historique disponibles. L'exécution d'opérations de détection par lots limite la quantité de données pouvant être traitées dans un laps de temps raisonnable. Avec l'approche de diffusion de Vectra, les algorithmes extraient les éléments pertinents d'un événement et les factorisent sous forme de nouvelles références qui vont alimenter les modèles. Grâce à cet apprentissage basé sur les données de diffusion, les références sont créées à partir de données recueillies sur plusieurs mois et de millions d'événements, ce qui garantit des alertes de qualité élevée.

L'intelligence artificielle pour la mise en corrélation des menaces

L'IA de Vectra s'applique non seulement à l'identification des méthodes d'attaque individuelles, mais aussi à la mise en corrélation de ces actions pour identifier, catégoriser et prioriser activement les attaques en cours. Cette mise en corrélation est nécessaire, car les cyberattaquants peuvent exécuter plusieurs actions sur différents domaines afin d'atteindre leur objectif final. Un algorithme dédié de mise en corrélation analyse les comportements sur les comptes, les hôtes, le réseau et le cloud pour fournir un signal clair d'un incident de sécurité. Cet algorithme attribue ensuite les comportements à des ancrages stables sous la forme de comptes ou de systèmes.

Cet algorithme attribue ensuite les comportements à des ancrages stables sous la forme de comptes ou de systèmes.

Par exemple, dans les réseaux et les environnements cloud hybrides, des adresses IP transitoires sont attribuées à des systèmes stables en fonction des artefacts observés via un algorithme appelé host-id. Ces artefacts sont collectés à partir de métadonnées réseau qui incluent des informations telles que les hôtes principaux Kerberos, les adresses et cookies MAC DHCP, et à partir d'intégrations d'API comme les solutions EDR, vCenter, Azure et AWS. Une fois les artefacts attribués à un système donné, et à chaque fois qu'une adresse IP est observée avec un artefact donné, ce flux de métadonnées et le comportement d'attaque associé peuvent être attribués à ce système nommé, pas uniquement à l'adresse IP.

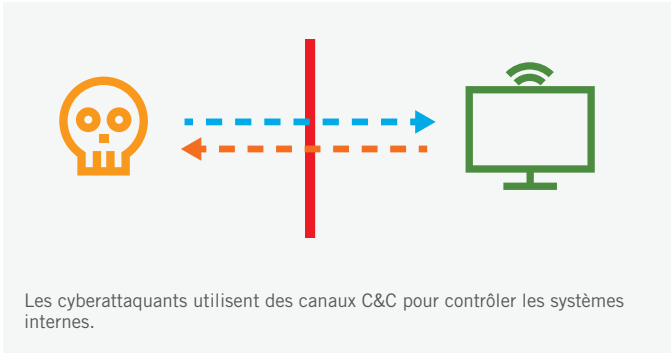
> L'IA de Vectra ne s'applique pas qu'à l'identification des méthodes d'attaque individuelles, mais aussi à la mise en corrélation de ces actions pour identifier, catégoriser et prioriser activement les attaques en cours.

AWS présente un défi d'attribution différent en raison de la façon dont les événements sont enregistrés dans le plan de contrôle AWS. En effet, les événements sont associés à des rôles présumés, pas aux comptes d'utilisateur sous-jacents. Tous les comptes peuvent assumer un rôle donné, mais, pour intervenir sur une attaque, il est primordial de connaître l'utilisateur IAM ou SAML réel qui assume ce rôle.

Les cyberattaquants avancés peuvent compliquer davantage encore l'action des systèmes de défense en enchaînant les rôles afin de dissimuler l'origine d'une attaque. Vectra utilise une technologie personnalisée appelée Kingpin qui peut revenir en arrière dans l'enchaînement des rôles et attribuer les attaques observées à un utilisateur sous-jacent, pas à un rôle ambigu.

Une fois les comportements d'attaque associés à un indicateur stable, ces comportements sont mis en corrélation entre eux afin de permettre l'identification du profil comportemental sous-jacent du système, qui ensuite étiquette et priorise les menaces en cours. L'algorithme de mise en corrélation a été conçu de façon à répliquer les actions entreprises par les analystes et chercheurs en sécurité de Vectra lorsqu'ils enquêtent sur les menaces. Cela leur permet de classifier des scénarios d'attaque avancés, comme les menaces externes ou les menaces internes de niveau administrateur, pour examen immédiat.

Étude de cas de la détection basée sur l'IA : canaux C&C chiffrés



Les cyberattaquants utilisent des canaux C&C pour contrôler les systèmes internes.

Méthode d'attaque

Au cœur de toute attaque basée sur le réseau se trouve l'utilisation d'un canal C&C (Command & Control). Un cyberattaquant disposant d'un accès à un système déploie un logiciel malveillant qui se connecte à un serveur externe. Bien que ce soit la machine interne qui initie la connexion, les réponses du serveur externe contiennent des instructions que le système infecté exécute, ce qui permet au cyberattaquant de poursuivre son attaque.

Les outils C&C sont utilisés dans des cadres d'attaques génériques, comme Cobalt Strike et Metasploit, ou développés en interne par des groupes d'attaque avancés. Ces cadres prennent en charge le chiffrement du canal ainsi que d'autres techniques comme l'utilisation de domaine-écran ou la gigue de session qui permettent aux cyberattaquants de contourner les techniques de détection.

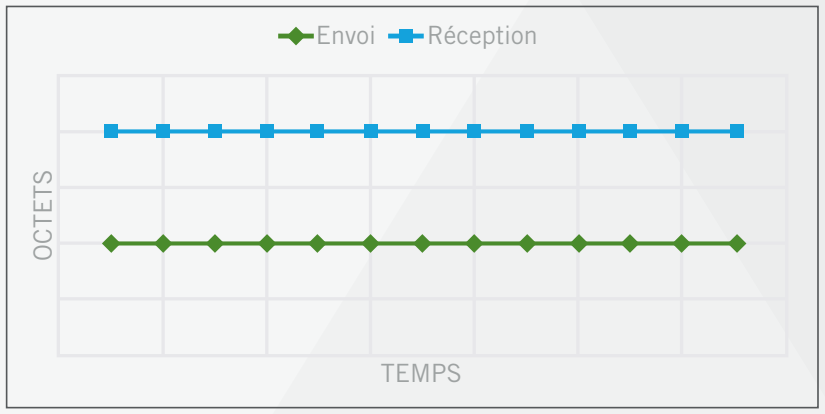
Vectra détecte les canaux C&C indépendamment du chiffrement ou des autres techniques de contournement.

Méthodologie de détection

Vectra détecte les canaux C&C indépendamment du chiffrement ou des autres techniques de contournement. Cette couverture découle de l'approche axée sur la sécurité décrite ci-dessus et permet de résoudre les nombreuses failles de l'utilisation d'une approche axée sur les maths pour résoudre les problèmes.

Lorsque l'équipe de chercheurs en sécurité de Vectra a extrait le comportement d'un canal C&C, elle a identifié que les indicateurs les plus évidents de cette méthode n'étaient pas des aspects circonstanciels du trafic, comme les domaines ou les agents utilisateur rares, mais plutôt la forme réelle du trafic réseau au fil du temps.

Prenons l'exemple ci-dessous, qui représente le trafic inoffensif d'un système externe.

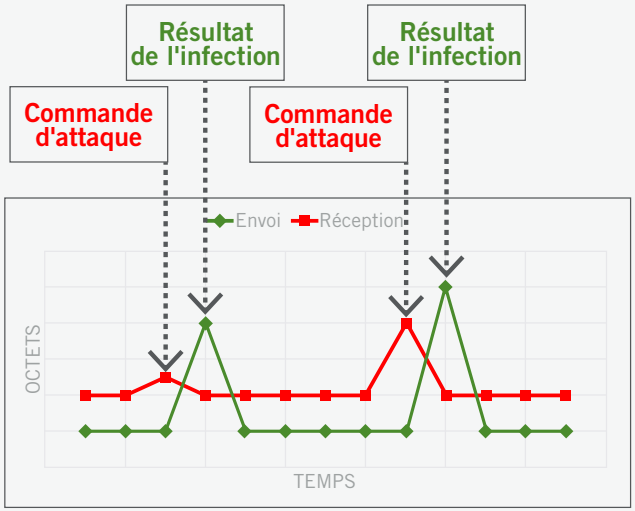


Trafic inoffensif

Exemple de trafic de transfert de données avec balise inoffensive.

Cet exemple de trafic caractérise un système qui envoie des balises extérieures à un serveur externe. Les balises sont une fonction réseau très courante utilisée par des services tels que les téléscripteurs boursiers, les applications de chat et les traceurs d'annonces. Elles permettent aux systèmes locaux et distants de rester synchronisés et de communiquer. Les canaux C&C malveillants utilisent eux aussi cette fonctionnalité.

Cependant, il existe une différence subtile dans la façon dont une balise apparaît lorsqu'elle est utilisée dans un téléscripteur boursier et lorsqu'elle est utilisée dans un canal malveillant. Voyons le schéma ci-dessous, qui représente un tunnel chiffré malveillant.

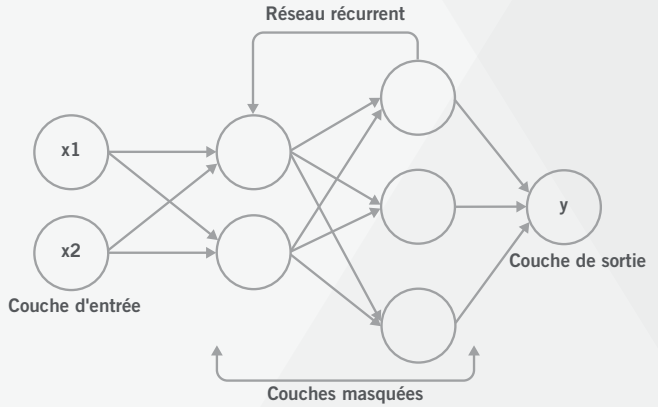


Exemple de trafic de transfert de données avec C&C malveillant.

Vous voyez les pics ? Ils se produisent lorsque le cyberattaquant envoie sa commande et que le système infecté envoie un résultat. Le premier pic de données survient de façon impromptue dans les octets de réception et est rapidement suivi par la réponse de la machine infectée.

Les scientifiques des données de Vectra ont étudié ces modèles et ont trouvé une approche optimale pour identifier ce comportement. Les données chronologiques qui caractérisent le comportement du canal C&C présentent de nombreuses similarités avec les données utilisées dans la reconnaissance vocale et le traitement du langage naturel, ce qui a incité l'équipe à opter pour l'utilisation d'un modèle d'apprentissage profond.

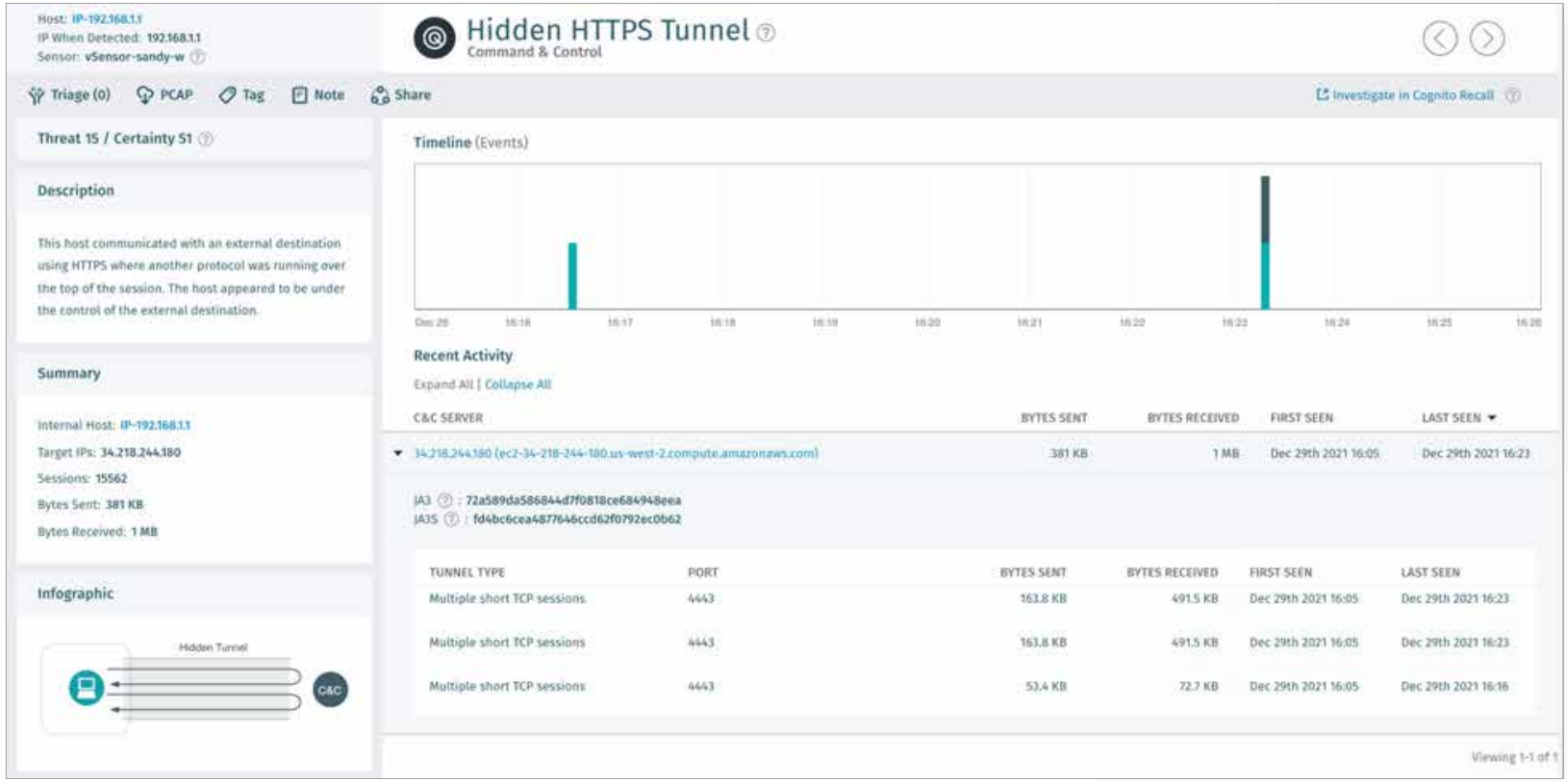
Vectra utilise une architecture spécifique de réseau neuronal récurrent connue sous le nom de LSTM (Long Short-Term Memory) pour identifier le comportement d'attaque. Ce type d'algorithme offre d'excellentes performances dans la compréhension des événements qui se déroulent à différentes échelles temporelles, ce qui est fondamental pour bien cerner la nature des données de conversation C&C. L'architecture LSTM est entraînée avec des échantillons réels et générés par algorithme. L'ensemble de données couvre une vaste plage de scénarios, d'outils, de configurations et d'environnements, ce qui permet au modèle d'identifier le signal généralisable d'un canal de contrôle indépendamment de l'outil utilisé.



Vectra utilise les réseaux neuronaux récurrents pour faire la différence entre communications C&C malveillantes et signaux inoffensifs.

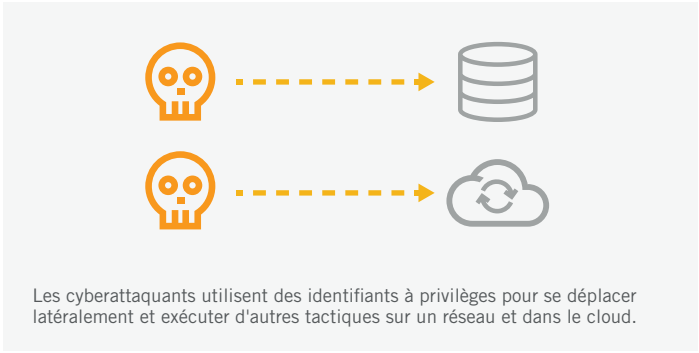
À noter que cette approche algorithmique a été rendue possible par la façon dont Vectra formate les données de session réseau. Si Vectra est capable de sortir des métadonnées de type Zeek, son analyseur personnalisé fournit une fidélité de métadonnées au-delà de celle assurée par un cadre Zeek standard avec une analyse des communications réseau à des intervalles de moins d'une seconde. Cette vue granulaire offre une visibilité élevée sur tous les types de communications, inoffensifs et malveillants, et permet aux équipes de scientifiques des données de Vectra d'utiliser les algorithmes qui assurent la meilleure couverture possible pour un large éventail de problèmes.

En combinant ces métadonnées uniques à une approche algorithmique sophistiquée, il est possible de détecter efficacement les cyberattaquants. La décision d'accorder la priorité aux données de communication elles-mêmes, et non aux signaux superficiels, autorise une grande résilience aux changements d'outils et au trafic chiffré. En outre, le signal de comportement clair élimine le besoin d'utiliser des filtres de suppression susceptibles de filtrer les canaux frontaux ou les actions de cyberattaquants furtifs.



Détection Vectra d'un canal C&C chiffré.

Étude de cas de la détection basée sur l'IA : utilisation abusive d'identifiants à privilèges sur des réseaux et dans le cloud



Les cyberattaquants utilisent des identifiants à privilèges pour se déplacer latéralement et exécuter d'autres tactiques sur un réseau et dans le cloud.

Méthode d'attaque

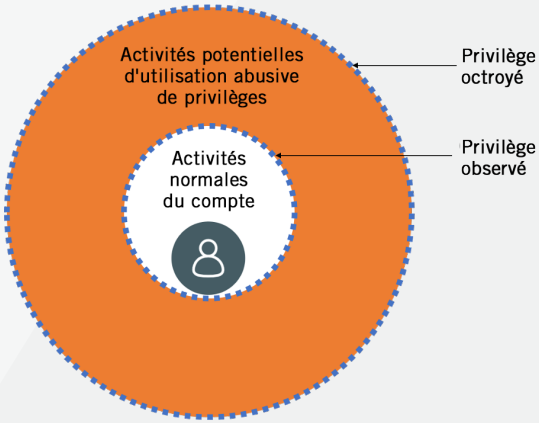
Un cyberattaquant qui parvient à voler des identifiants à privilèges bénéficie d'un vaste accès aux ressources réseau et cloud, sans avoir besoin de passer par un malware ou des charges actives qui peuvent laisser des traces ou déclencher des alarmes préventives. En accordant strictement aux utilisateurs un accès du moindre privilège, vous pouvez atténuer certaines attaques, mais des incidents récents démontrent que mettre en œuvre ce niveau de privilège représente toujours un défi.

Empêcher l'utilisation abusive d'identifiants volés implique de disposer de la capacité à détecter les menaces, les tentatives d'usurpation et l'exploitation d'un compte. Chaque action exécutée par un cyberattaquant est explicitement autorisée en fonction des autorisations qu'il a reçues. Les alertes basées sur des concepts comme les nouvelles interactions ne sont pas efficaces. En effet, les utilisateurs évoluent dans un environnement dynamique où l'accès aux nouvelles ressources est essentiel à leurs tâches quotidiennes. Un cyberattaquant qui a acquis une bonne compréhension d'un environnement essaiera de se dissimuler et, pour ne pas éveiller les soupçons, d'effectuer des actions non considérées comme nouvelles pour un compte particulier. Pour identifier efficacement l'utilisation abusive d'identifiants à privilèges, il faut utiliser une approche axée sur la sécurité afin de déterminer ce qu'un cyberattaquant essaie d'accomplir avec ces identifiants volés.

Méthodologie de détection

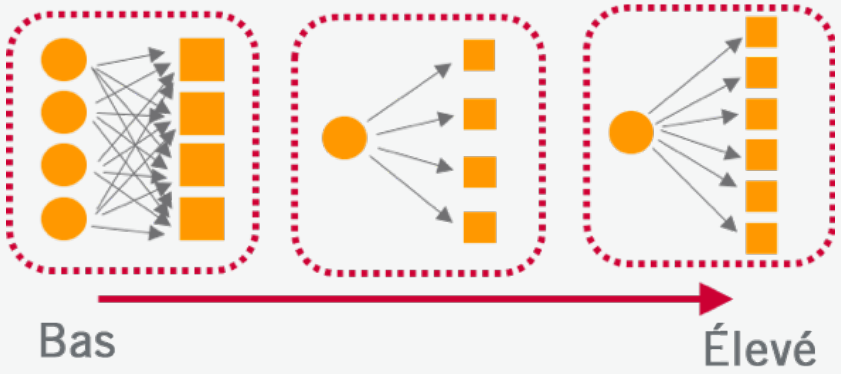
Vectra peut identifier l'utilisation abusive d'identifiants à privilèges volés sur des réseaux et des environnements cloud. Cette approche de la détection axée sur la sécurité implique de bien comprendre comment les cyberattaquants exploitent des identifiants volés. Pour un cyberattaquant, les identifiants à privilèges ont une grande valeur, car ils leur permettent d'accéder aux services et fonctionnalités stratégiques et privilégiés d'un environnement.

D'après les chercheurs en sécurité de Vectra, si vous connaissiez le privilège réel de chaque compte, système, service et opération cloud, vous disposeriez d'un plan de toutes vos ressources à haute valeur ajoutée. Si le concept de *privilège octroyé* est bien établi, cette représentation place une limite maximale à ce qu'est le véritable privilège de quelque chose comparé au privilège minimum nécessaire. Au lieu de cela, les équipes de chercheurs en sécurité et de scientifiques des données de Vectra ont identifié une nouvelle façon de représenter la valeur de systèmes dans un environnement, basée sur les événements observés au fil du temps. Cette vision dynamique et globale de la valeur est appelée *privilège observé*. Basée sur les données, elle offre une approche Zero Trust efficace de l'utilisation des identifiants sans configurations manuelles.



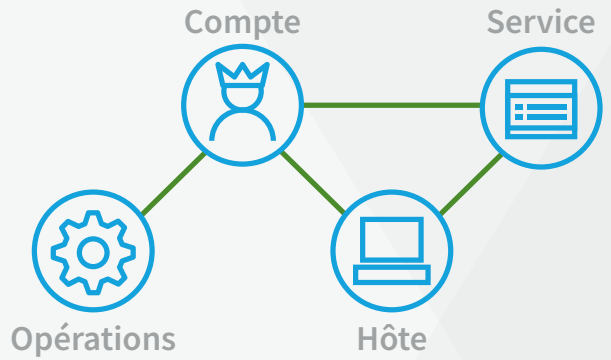
Le privilège observé est une vue Zero Trust du privilège normal dont un utilisateur a besoin pour effectuer son travail. L'utilisation d'un privilège supérieur à celui qui est normalement nécessaire exige un examen supplémentaire.

L'IA de Vectra calcule le *privilège observé* en tenant compte des interactions historiques entre les entités suivies, pas du privilège défini par un administrateur informatique. L'ampleur et la spécificité des accès et de l'utilisation contribuent énormément aux scores. Un système qui accède à plusieurs systèmes auxquels d'autres systèmes accèdent normalement disposera d'un privilège bas tandis qu'un système qui accède à de nombreux systèmes auxquels aucun autre système n'accède aura un score de privilège élevé. Grâce à cette approche, Vectra peut faire la différence entre comptes d'administration de domaines et comptes d'utilisateur normaux.



Vectra apprend les niveaux de privilège observés en fonction du comportement de l'utilisateur. Un compte qui accède à de nombreux services communs possède un privilège plus bas qu'un compte qui accède à des services peu utilisés par ailleurs.

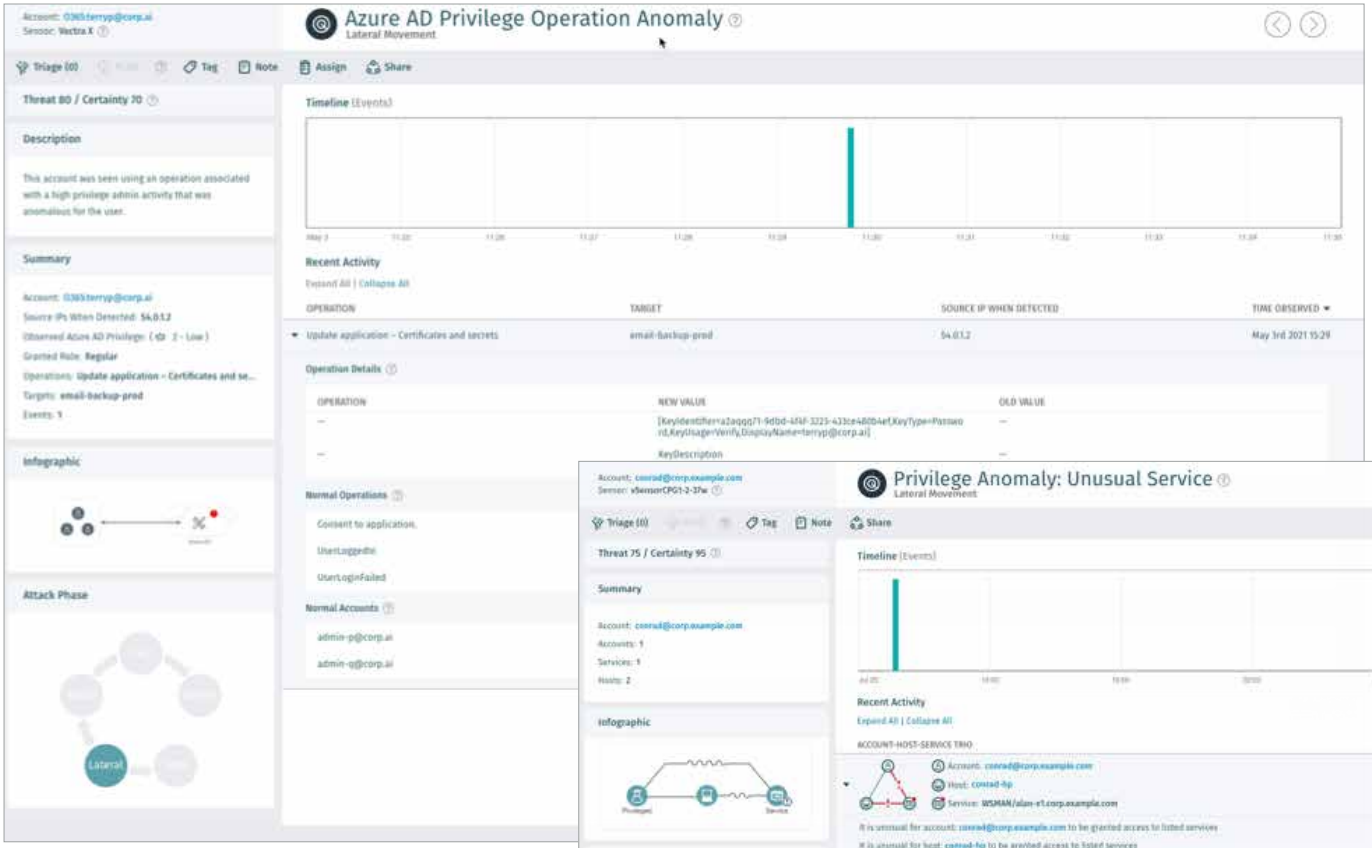
Une fois les scores de privilège observé calculés, toutes les interactions entre les comptes, les services, les systèmes et les opérations cloud sont mises en corrélation pour déterminer les interactions historiques normales entre les systèmes. Ensuite, une suite d'algorithmes d'apprentissage non supervisés qui tiennent compte des scores de privilège, et utilisent des algorithmes personnalisés de détection des anomalies ainsi que des implémentations de l'algorithme HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise), identifient les cas anormaux d'utilisation abusive des privilèges.

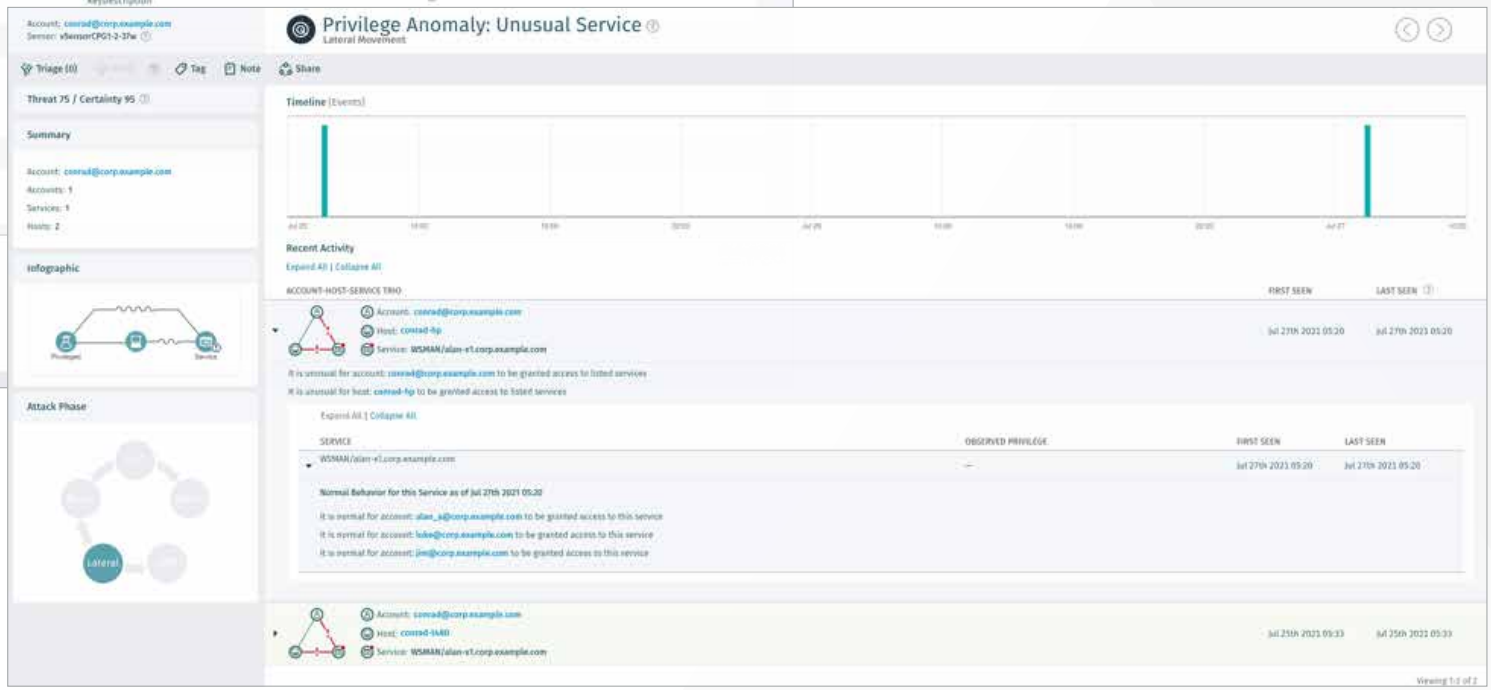


Pour détecter des utilisations abusives d'identifiants, Vectra applique un apprentissage non supervisé qui tient compte du privilège observé et des interactions entre les comptes, les hôtes, les services et les opérations cloud.

Cette approche sophistiquée axée sur la sécurité permet de repérer les identifiants volés utilisés de façon abusive à la fois sur des réseaux sur site et dans le cloud. L'indicateur du *privilège observé* repose sur la détection des actions anormales qui comptent, et autorise une précision plus élevée et un rappel plus précis qu'une approche qui ignore cette perspective critique.

L'IA de Vectra calcule le *privilège observé* en tenant compte des interactions historiques entre les entités suivies, pas du privilège défini par un administrateur informatique.





Détection Vectra de comptes qui utilisent des privilèges de façon abusive.

Nous innovons pour devancer les cyberattaques

Les cyberattaquants ne vont jamais s'arrêter d'innover, et c'est pour cette raison que les équipes de sécurité doivent faire de même. Au fil des ans, Vectra a constamment innové pour développer la plate-forme de détection et d'aide à la résolution des incidents la plus efficace possible pour les ressources réseau et cloud.

Vectra a développé plus de 100 détections optimisées par l'intelligence artificielle et axées sur la sécurité, et identifié un nombre infini de menaces sur les réseaux client et les environnements cloud, bloquant ainsi les cyberattaques avant qu'elles ne fassent des dégâts. Chacune de nos détections est basée sur notre compréhension approfondie de la façon dont les cyberattaquants agissent, et utilise les techniques d'apprentissage automatique les plus avancées disponibles. Au total, Vectra possède 33 brevets pour des technologies qui prennent en charge ces détections.

Au-delà de la couverture que les technologies brevetées de Vectra autorisent, nous pouvons nous enorgueillir d'être le fournisseur le plus référencé par la NSA et le cadre MITRE D3FEND, qui définit les contremesures que les équipes de sécurité doivent mettre en place pour protéger leur environnement. Le cadre D3FEND détaille les méthodes à disposition pour bloquer les attaques et contrer les techniques d'attaque définies dans le cadre MITRE ATT&CK. Au total, le cadre D3FEND cite 12 brevets Vectra comme des contremesures de défense de référence.

Chez Vectra, nous nous engageons à rendre le monde plus sûr et plus juste. C'est pour cela que nous allons continuer à exploiter l'IA axée sur la sécurité pour innover et développer des capacités de détection à même de bloquer les cyberattaques.

Pour plus d'informations, veuillez nous contacter à l'adresse info_france@vectra.ai.

E-mail : info_france@vectra.ai vectra.ai