Un mixte de machine learning et d'expertise humaine pour mieux prévenir les attaques informatiques

Combiner l'intelligence artificielle à l'expertise humaine pour améliorer la sécurité des systèmes informatiques. C'est la formule –gagnante– qu'ont élaboré pendant deux ans des chercheurs du CSAIL, le laboratoire de science informatique et d'intelligence artificielle (Computer Science and Artificial Intelligence Lab) du MIT (Massachussetts Institute of Technology), et la start-up PatternX spécialisée en machine learning (ML). Les deux équipes ont mis au point une plate-forme baptisée AI^2 capable de prévenir les cyber-attaques de manière beaucoup plus efficace qu'avec les systèmes actuels.

85% des attaques détectées

Testée sur un ensemble de 3,6 milliards de données de connexion (log lines), un volume généralement généré par des millions d'utilisateurs sur une période de trois mois, AI^2 a détecté 85% des attaques. Ce qui est grossièrement trois fois mieux que les tests sur des systèmes «traditionnels» aujourd'hui construits autour de règles rigides édictées par des responsables en sécurité, ou bien dans une approche d'auto apprentissage par la machine à partir de détections d'anomalies comportementales. Ce qui entraine immanquablement des faux-positifs qui finissent par perturber le système informatique nécessitant une intervention humaine. Selon PatternX, AI^2 génère cinq fois moins de faux-positifs.

Pour atteindre de telles performances, la plate-forme fait remonter des activités suspectes en regroupant les données dans des modèles utilisant du machine-learning non supervisé. Une analyse humaine confirme ensuite les événements effectivement associables à des attaques informatiques et intègre ce classement dans ses modèles pour la prochaine série de données à analyser. *« Vous pouvez considérez cela comme un analyste virtuel*, souligne le scientifique du CSAIL Kalyan Veeramachaneni dans un article publié par le laboratoire. *[Le système] génère en permanence de nouveaux modèles qu'il peut affiner en seulement quelques heures, ce qui permet d'améliorer ses taux de détection sensiblement et rapidement. »* Kalyan Veeramachaneni a co-développé AI^2 avec Ignacio Arnaldo, datascientist de PatternX et doctorant du CSAIL. Le scientifique a présenté ses travaux dans le cadre de la 2e conférence Big Data Security de l'IEEE (Institute of Electrical and Electronics Engineers) qui s'est tenue à New York (du 8 au 10 avril 2016).

Un système qui s'améliorer continuellement

Si cette approche bimodale montre sa pertinence, son efficacité dépendra néanmoins de la qualité des informations générées par les experts humains. Autrement dit, bien qualifier ce qui est une attaque de ce qui ne l'est pas. Une tâche résolument dévouée à un expert en sécurité. Lesquels ont

généralement d'autres tâches à accomplir qu'à cliquer sur des boutons pour confirmer, ou non, une tentative d'attaque.

Une problématique à laquelle AI^2 entend répondre en fusionnant trois méthodes différentes d'apprentissage non supervisé et en ne présentant que les événements majeurs aux experts humains. Ce qui génère un modèle supervisé constamment affiné dans un mouvement vertueux d'optimisation continue pour, au final, faciliter le travail du RSSI. Selon PatternX, la remontée de 200 alertes lors du premier jour d'utilisation AI^2 se bornera à 30 ou 40 au bout de quelques jours d'apprentissage. *« Plus le système détecte d'attaques et reçoit de retours de l'analyste, et plus la précision des prévisions futures est améliorée*, résume Kalyan Veeramachaneni. *Cette interaction homme-machine crée une cascade du plus bel effet. »* Les tentatives de fraudes, usurpations de comptes et autres abus de services ne manqueront pas de donner aux entreprises clientes l'occasion de pouvoir le vérifier.

Lire également
Cyberattaques : un cru 2015 très actif et plus criminalisé
Les entreprises démunies face aux cyberattaques sophistiquées
Cyberattaques : une entreprise européenne sur cinq a été touchée en 2015

Crédit photo : Lightspring-Shutterstock