Caractérisation tactique d'un attaquant évoluant dans un réseau compromis

Aimad Berady*, Valérie Viet Triem Tong*, Gilles Guette* et Mathieu Jaume†
*Equipe CIDRE : CentraleSupélec, Inria, Univ Rennes, CNRS, IRISA, Rennes
†Equipe MoVe : Sorbonne Université, CNRS, LIP6, Paris
e-mails : prenom.nom@(irisa | lip6).fr

***Résumé*—Les menaces dites « persistantes et avancées » laissent couler beaucoup d'encre et sont malheureusement souvent étudiées qu'à travers un prisme technique. Dans cet ensemble de travaux, nous avons cherché à considérer le facteur opérationnel qui guide les attaquants au cours de leurs campagnes.**

I. Introduction, menaces persistantes avancées

Les menaces dites « persistantes et avancées » (*Advanced Persistent Threats*) ont été théorisées par le NIST en 2011 [1]. L'institut les caractérise selon trois points : le groupe d'attaquants considéré comme une menace persistante avancée (1) poursuit ses objectifs de manière répétée sur une longue période de temps ; (2) s'adapte aux efforts des défenseurs pour y résister ; et (3) est déterminé à maintenir le niveau d'interaction nécessaire pour atteindre ses objectifs. Considérés de niveau « étatique », ces groupes sont constamment étudiés par les experts en analyse de la menace (*Threat Intelligence*), qui cherchent à comprendre leurs fonctionnements pour alimenter les chasseurs de menaces (*Threat Hunters*) en renseignements techniques directement utilisables pour leurs opérations défensives.

II. Phases opérationnelles

Dans la contribution titrée *Modeling the Operational Phases of APT Campaigns* [2], nous avons proposé un modèle du cycle de vie d'un attaquant. Il complète la vision décrite dans de précédentes publications telles que la populaire *Cyber Kill Chain* de Lockheed Martin [3]. L'apport de notre contribution est la considération de deux facteurs déterminants pour un attaquant : la notion de régression en cours d'attaque et celle de l'atteinte d'un objectif final de façon répétée. Le modèle a été confronté à deux attaques réelles dont les analyses techniques ont été publiées : la fuite de données d'Equifax (2017) et le sabotage de TV5Monde (2015). Le modèle s'articule en trois grandes phases opérationnelles dans lesquelles peut se trouver un attaquant après l'obtention d'un accès initial au réseau de la victime. Ces phases sont : la phase d'exploration ; la phase d'exploitation ; la phase décisionnelle.

Au cours de la phase d'exploration, l'attaquant passe successivement par deux états. Dans l'état de *Network propagation*, il va gagner de la connaissance et des privilèges sur le réseau de la victime. Dès qu'il aura découvert l'équipement qui accueille son objectif final, il va chercher à le dompter techniquement pour arriver être en mesure d'y atteindre ses objectifs finaux. C'est l'état *Asset dominance*.

Par nature, c'est dans l'état de propagation dans le réseau (*Network propagation*) que l'attaquant, évoluant dans un épais brouillard, risque de se dévoiler puisqu'il avance à tâtons et qu'il provoque de nombreuses erreurs sur le système d'information de sa victime ; et que ces événements peuvent être captés par des dispositifs de sécurité mis en place par le défenseur. C'est cette voie que nous avons choisie suivre pour nos travaux.

III. Propagation dans le réseau

En nous focalisant sur l'état de propagation dans le réseau, nous avons étudié les éléments manipulés par l'attaquant et perceptibles par le défenseur au cours d'une campagne. Dans l'article titré *From TTP to IoC : Advanced Persistent Graphs for Threat Hunting* [4], nous décrivons un modèle formel reliant entre eux tous les éléments de l'attaquant et du défenseur. Il en ressort que les paramètres qui instancient les procédures de l'attaquant traversent une chaîne d'exécution puis de détection et se retrouvent dans les mains du défenseur sous la forme d'objets présents dans des événements d'intérêt (EoI). Le cas échéant, ces événements alimentent les bases d'indicateurs de compromission (IoC) du défenseur.

Chacun de son côté, l'attaquant et le défenseur peuvent donc tracer ces objets à fins de capitalisation. Il s'agit d'objets que l'attaquant a conscience d'exposer ou que le défenseur a détectés. Nous représentons les ensembles de ces objets sous la forme de graphes persistants. Pour l'attaquant, chaque exécution de procédure est un graphe dont le nœud central est le composant du réseau de la victime sur lequel la procédure a été exécutée ; l'union de ces graphes représente une campagne. Pour le défenseur, chaque événement d'intérêt (EoI) est un graphe, dont le nœud central est le composant sur lequel a été détecté l'événement ; l'union de ces graphes représente la vision du défenseur d'une attaque (*i.e.*, espace de propagation).

Il est important de préciser qu'aucune des deux visions n'est parfaite : l'attaquant peut exposer inconsciemment des objets et le défenseur, puisqu'il est tributaire de ses règles de détection et de sa politique de journalisation, peut ne pas être alerté de certains événements malveillants ou encore être pollué par trop de faux positifs.

Au cours de cette étude, nous avons conduit une expérimentation sur un jeu de données issu du projet public *Mordor*. Les résultats observés nous ont permis d'identifier trois paramètres qui pouvaient être ajustés par le défenseur pour optimiser sa chaîne de détection : les configurations des sondes ; les règles de détection ; la base d'IoC.

Afin de pouvoir étudier des données issues d'un autre scénario d'attaque, nous avons initié un projet étudiant qui vise à reproduire une attaque documentée dans un environnement adapté. C'est l'ouvrage de Sparc Flow [5] qui a été utilisé pour cette première mouture. Une étude approfondie de la publication nous a permis de comprendre l'architecture du système d'information et de le reproduire dans un environnement virtualisé. Afin d'assurer la fonction de journalisation des événements de sécurité, l'architecture a été enrichie d'un SIEM (*Splunk*) et des sondes système (*Sysmon* et *auditd*). Nous avons ensuite extrait les étapes de procédures décrites par l'auteur avant de les rejouer sur l'infrastructure hébergée dans notre laboratoire. Nous allons poursuivre ces travaux afin de disposer de nouvelles données pour valider notre modèle.

IV. Comportement d'un attaquant

Lorsque l'attaquant se propage dans le réseau de sa victime, il compromet un grand nombre de systèmes jusqu'à découvrir celui qui présente un intérêt pour lui. Dans le cadre de l'expérimentation *Pwnjutsu*, nous cherchons à apporter des éléments de réponse à la question de l'attractivité d'un système du point de vue de l'attaquant. C'est-à-dire l'ensemble des éléments techniques ou organisationnels qui amèneront un attaquant à s'intéresser à un composant du système d'information plutôt qu'à un autre. Cette notion d'attractivité se concrétise sur deux axes : celui du positionnement au sein d'un réseau d'entreprise, qui est fortement dépendant de l'architecture du système d'information ; et, d'une manière plus générale, sur celui de son apparence extérieure, qui se traduira par la perception que l'attaquant aura du système auquel il fera face.

Pour cela, nous avons mis en place une expérimentation qui confronte une vingtaine d'attaquants, choisis au sein de la communauté d'experts de *YesWeHack* et recrutés spécifiquement pour les besoins de cette étude, à une architecture cible représentative d'un système d'information. Le but donné aux attaquants est de compromettre successivement trois machines et d'y collecter des *flags*, à la manière d'une compétition de type CTF. La particularité de notre architecture réside dans le fait que les attaquants se trouvent régulièrement dans des situations où ils devront choisir une alternative d'exploitation. De plus, notre scénario présente la particularité de n'être composé que de vulnérabilités connues et dont l'exploitation est évidente. En observant les premières intentions de plusieurs participants, nous en dégageons une tendance qui nous indiquera quelles caractéristiques techniques d'un système sont préférées par les attaquants.

Les données brutes produites au cours de cette expérimentation auront la forme de journaux d'événements issus des sondes des machines compromises par les participants et collectées par un SIEM. Les sondes surveillent également les accès en lecture aux fichiers contenant les *flags*. Cela nous permettra, en phase d'exploitation des résultats, de reconstruire les chemins empruntés par les attaquants au cours de leurs progressions.

Cette expérimentation permet également à notre partenaire, l'IRSN, d'enrichir leur vision sur les caractéristiques que devrait présenter un pot de miel déployé dans un réseau à défendre avec une approche dite « active ».

V. Conclusion

Tout au long de ces travaux, nous avons systématiquement cherché à observer le comportement des attaquants sous un angle opérationnel. Dans un premier temps en décrivant les phases et états par lesquels il passait. Puis, en nous concentrant sur l'état de propagation dans le réseau, nous avons cherché à identifier la relation logique qui existait entre les procédures des attaquants et les événements d'intérêt observés par le défenseur. Enfin, nous avons conçu une expérimentation inédite qui met une vingtaine d'attaquants face à une infrastructure laissée volontairement vulnérable afin d'observer leurs premières intentions techniques au cours de leurs progressions. S'inscrivant dans la dynamique d'une thèse, ces travaux ont été l'occasion de mieux conceptualiser les Tactiques, Techniques et Procédures qu'un attaquant met en œuvre dans le cadre de ses campagnes.

Remerciements

Nous tenons à remercier tout particulièrement Olivier Fichot et Vincent Clément de l'IRSN pour leur participation active au projet de *Pwnjutsu* et Benoît Fournier de l'Inria pour son précieux appui technique dans les phases d'implémentation. Nous saluons également le travail de Lucas Altenburg, Matthieu Capitant, Farba Harouna Diop et Louis Vayssette, étudiants du Mastère Spécialisé® en Cybersécurité de CentraleSupélec et IMT Atlantique, dans le cadre du projet industriel *Hack like*. Enfin, nous sommes reconnaissants envers la communauté *YesWeHack* qui nous a permis de travailler avec des experts de haut niveau.

Références

[1] R. Ross, "Managing information security risk : Organization, mission, and information system view," 2011.

[2] A. Berady, V. Viet Triem Tong, G. Guette, C. Bidan, and G. Carat, "Modeling the Operational Phases of APT Campaigns," in *CSCI 2019 - 6th Annual Conf. on Computational Science & Computational Intelligence*. IEEE, 2019.

[3] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains," 2011.

[4] A. Berady, M. Jaume, V. Viet Triem Tong, and G. Guette, "From TTP to IoC : Advanced Persistent Graphs for Threat Hunting," *IEEE Transactions on Network and Service Management*, 2021.

[5] S. Flow, *How to hack like a pornstar : Master the secrets of hacking through real-life hacking scenarios*, 2017.