Plateformes d'expérimentation cyber : création de nouveaux scénarios d'attaques, extension de l'activité normale, automatisation

— **Encadrants** : Frédéric Majorczyk, Maxime Lanvin
— **Mots-clés** : automatisation, TTPs (Techniques, Tactics and Procedures), VMs, IDS
— **Niveau** : M1/M2
— **Durée** : 4 à 6 mois

Sujet

Deux plateformes d'expérimentation cyber dédiées à la génération de logs ont été récemment mises à la disposition de la communauté sous la forme de projets open-source : la plateforme SOCBED [1] (Self-contained Open-source Cyberattack experimentation testBED) et la plateforme Kyoushi [2].

Ces deux plateformes fonctionnent à peu près de la même manière : à partir de descriptions du systèmes d'information, de l'activité normale et du scénario d'attaque, un ensemble de machines virtuelles est démarré et configuré puis l'activité normale et le scénario d'attaques sont joués. Ceci permet de tester des produits de sécurité (tels que des systèmes de détection d'intrusions) directement en ligne mais également de générer des logs qui pourront être utilisés hors ligne pour mettre au point et évaluer, a minima, des systèmes de détection d'intrusions.

Les articles [3, 4] décrivent les expérimentations menées par les auteurs avec ces deux plateformes.

L'objectif de ce stage est tout d'abord de faire une comparaison entre ces deux plateformes en les prenant en main, puis d'apporter un certain nombre d'améliorations à la plateforme la plus prometteuse. On cherchera tout d'abord à reproduire les résultats des articles. Pour SOCBED, il sera possible d'utiliser le git de l'évaluation de l'article : `https://github.com/fkie-cad/socbed-eval-acsac-2021`.

Dans l'objectif d'apporter des améliorations à cette plateforme, plusieurs axes seront explorés :

- spécification des points de capture réseau et automatisation de la collecte du trafic réseau ;
- amélioration de l'activité normale des utilisateurs ;
- amélioration du système d'informations (ajout de services) ;
- ajout de nouvelles étapes d'attaques et création de nouveaux scénarios d'attaques ;
- évaluation du passage à l'échelle de la plateforme ;
- génération de nouveaux datasets ;
- évaluation de sondes de détection d'intrusion réseau et/ou hôte.

Une fois la prise en main effectuée, la spécification des points de capture et l'automatisation de la collecte du trafic est une étape importante du projet. En effet, au moins pour SOCBED, la collecte des logs système est bien intégrée dans la plateforme actuelle mais ce n'est pas le cas pour le trafic réseau. Il sera important de permettre la définition de différents points de collecte dans le système d'information et d'automatiser le déploiement de la machine ou des machines virtuelles qui collecteront le trafic réseau.

Les 3 étapes suivantes du stage sont à la carte mais restent liées. Les attaques et l'activité normale des utilisateurs dépendent des services disponibles dans le système d'information et de leur configuration.

Concernant l'amélioration de l'activité normale des utilisateurs, il s'agira d'évaluer le réalisme de cette activité, puis de proposer des améliorations au niveau du réalisme des actions, de l'enchaînement des activités et de la diversité des actions disponibles [5, 6]. Il sera notamment possible d'ajouter des actions d'administration (éventuellement à partir d'une zone dédiée).

Concernant l'ajout de nouvelles étapes d'attaques et la création de nouveaux scénarios, il s'agira de s'inspirer d'APT réelles décrites dans des rapports d'analyse pour implémenter les différentes actions menées par les attaquants. Il sera possible de débuter par des scénarios déjà modélisés lors des évaluations MITRE [7].

Concernant l'amélioration du système d'information, il sera possible de mettre en place de nouveaux services (partage de fichiers, authentification, proxy, backup, git, intranet, etc.) et de restructurer le système d'informations (ajout d'une zone serveur, ajout d'une zone pour les développeurs, etc.). Il sera intéressant de pouvoir fournir plusieurs configurations possibles pour les différents services (plus ou moins sécurisés par exemple). Il faudra également configurer la collecte des logs pour ces différents services.

Une fois ces améliorations mises en place, l'étape suivante consistera en l'évaluation du passage à l'échelle de la plateforme : notamment l'impact de la collecte réseau, du nombre d'utilisateurs et de leur activité, des nouveaux services ajoutés, etc.

L'étape finale est la génération de nouveaux datasets qui pourront être fournis

à la communauté scientifique et l'évaluation de systèmes de détection d'intrusions hôtes et réseau sur ces nouveaux datasets (éventuellement dans plusieurs configurations différentes) [8, 9, 10, 11].

Références

[1] "SOCBED." `https://github.com/fkie-cad/socbed`. Accessed : 2022-11-03.

[2] "Kyoushi." `https://github.com/ait-aecid/kyoushi-environment`. Accessed : 2022-11-03.

[3] R. Uetz, C. Hemminghaus, L. Hackländer, P. Schlipper, and M. Henze, "Reproducible and adaptable log data generation for sound cybersecurity experiments," in *Annual Computer Security Applications Conference*, pp. 690–705, 2021.

[4] M. Landauer, F. Skopik, M. Wurzenberger, W. Hotwagner, and A. Rauber, "Have it your way : generating customized log datasets with a model-driven simulation testbed," *IEEE Transactions on Reliability*, vol. 70, no. 1, pp. 402–415, 2020.

[5] C. V. Wright, C. Connelly, T. Braje, J. C. Rabek, L. M. Rossey, and R. K. Cunningham, "Generating client workloads and high-fidelity network traffic for controllable, repeatable experiments in computer security," in *International workshop on recent advances in intrusion detection*, pp. 218–237, Springer, 2010.

[6] V. Kothari, J. Blythe, S. W. Smith, and R. Koppel, "Measuring the security impacts of password policies using cognitive behavioral agent-based modeling," in *Proceedings of the 2015 Symposium and Bootcamp on the Science of Security*, pp. 1–9, 2015.

[7] "Adversary emulation library." `https://github.com/center-for-threat-informed-defense/adversary_emulation_library`. Accessed : 2022-01-03.

[8] "Suricata." `https://suricata.io`. Accessed : 2022-10-03.

[9] "Zeek." `https://zeek.org/`. Accessed : 2022-10-03.

[10] "Wazuh." `https://wazuh.com/`. Accessed : 2022-10-03.

[11] "Sysmon." `https://learn.microsoft.com/en-us/sysinternals/downloads/sysmon`. Accessed : 2022-10-03.