RESSI 2023
RAPID SATAN - PROJET DE RECHERCHE COLLABORATIF

Projet	SATAN (plateforme de **S**imulation d'**AT**taques via **A**ppre**N**tissage automatique)
Consortium et TRL	Silicom (Porteur) et Amossys - TRL visé : 5
Convention projet	Régime d'Appui Pour l'Innovation Duale - Convention # 202906030
Période couverte	16 mars 2020 au 16 octobre 2022 (durée : 30 mois)

1. Contexte du projet, verrous scientifiques et techniques

La création du projet SATAN découle d'une analyse du marché, qui fait état des constats suivants :

- Les schémas actuels d'évaluation de la sécurité (Critères Communs et CSPN notamment) se concentrent principalement sur l'analyse de conformité et de robustesse des fonctions de sécurité. Néanmoins, on s'aperçoit d'un besoin de tester également l'efficacité des fonctions métier des produits de LID (Lutte Informatique Défensive), telles que les capacités de capture, de détection et de réaction.
- Les référentiels de qualification des prestataires PDIS/PRIS (Prestataires de Détection d'Incidents de Sécurité / Prestataires de Réponse aux Incidents de Sécurité) poussés par la LPM (Loi de Programmation Militaire) analysent essentiellement les processus des prestataires et la sécurité des systèmes d'information des prestataires. Il y a actuellement un manque de visibilité sur le niveau d'efficacité des prestations produites, et le marché, en particulier les OIV (Organismes d'Importance Vitale), commencent à remettre en cause le niveau de compétence et la capacité réelle de détection/réaction de leurs prestataires. Il y a donc ici un besoin de pouvoir estimer l'efficacité opérationnelle et la capacité de réaction des équipes SOC/CERT (Security Operation Center / Computer Emergency Response Team), de même qu'il est important de pouvoir former ces équipes face à des scénarios d'attaque réalistes.
- Les prestations pentest manquent encore d'automatisme et de déterminisme. Il est courant de voir des résultats très différents suivant les prestataires, qu'ils soient qualifiés PASSI ou non. Il est ici fait le constat d'un besoin de pouvoir tester la résistance d'un système d'information à la fois en profondeur et en couverture, avec une approche plus déterministe.

Ces trois constats convergent vers un besoin similaire, à savoir la mise au point d'une capacité avancée de simulation d'attaques, dont les objectifs visent à permettre de générer et d'exécuter des scénarios d'attaque réalistes, s'adaptant à l'environnement ciblé et prenant en compte des réactions potentielles de la défense (produits de LID et/ou équipes SOC/CERT). Le projet SATAN se positionne ainsi sur le domaine de la simulation d'attaques, également appelé RTA, pour Red Team Automation (automatisation de la Red Team) et a pour ambition de tirer parti de techniques du domaine de l'intelligence artificielle, et plus précisément du sous-domaine de l'apprentissage automatique, pour élaborer des scénarios d'attaque réalistes et efficaces.

Le paragraphe ci-dessous recense les verrous scientifiques et techniques marquant pour ce projet, ainsi que les actions que nous avons réalisées pour résoudre ces verrous :

- **VS. Temps de simulation du Cyber Range ralentissant l'apprentissage** - Modélisation et simulation du comportement du SI adaptée au contexte et à la vision de l'attaquant pentester. Développement de ce modèle dans l'outil "FAST SI" qui accélère drastiquement la vitesse d'apprentissage et permet à l'IA d'apprendre sur des SI variés, complexes et corrélés avec la réalité.
- **VT. Taille du SI attaqué et nombre "d'attaques unitaires et d'états" important (passage à l'échelle)** - Etude, conception et développement d'un algorithme IA, qui combine "Genetic Programming", "Reinforcement Learning" et "Mécanisme d'Attentions" pour ancrer des corrélations {Observations multi variants ; Actions ; Récompenses} permettant de délaisser les états et enchaînements d'actions sans intérêt. Grâce à plusieurs innovations, "Paramétrisation d'actions", "Nouveau mécanisme d'attention" et "Nouveau système de décision", l'IA est en mesure d'apprendre le chemin d'attaque optimal sur des SI dont les configurations sont diverses (topologies, caractéristiques des machines et vulnérabilités). La capacité de passage à l'échelle pour de grands SI n'a cependant pu être vérifiée lors de l'exploitation sur le Cyber Range faute de configuration suffisante.
- **VS. Problème de généralisation** - Plusieurs innovations apportées pour la généralisation : "FAST SI" diversifie autant que possible les SI simulés ; la "Paramétrisation d'actions" prend une décision adaptée malgré des paramètres à valeurs variables (ex : adresse IP) jamais observées pendant l'apprentissage. La généralisation a été démontrée grâce au Transfer Learning sur le cyber range. Et ce, sans aucune modification du modèle malgré des configurations différentes de celles utilisées pour l'apprentissage.

2. Approche méthodologique suivie

La boucle méthodologique a suivi une trajectoire classique :

- Etats de l'art continus : algorithmes RL et neuro évolutifs, mécanismes d'attention, mécanismes d'exploration, systèmes de récompenses extrinsèques et intrinsèques, paramétrisation d'actions, méthodes d'accélération des simulations SI, modélisation de simulations SI.
- Identification des solutions prometteuses et POC.
- Intégration des POCs pertinents et améliorations continues.

3. Résultats obtenus et démonstration

Les résultats obtenus sont les suivants :

- Une plateforme complète de simulation d'attaque a été mise au point. Cette plateforme permet de reproduire de manière réaliste des scénarios d'attaque représentatifs de modes opératoires d'attaquants (*kill chain*). Cette plateforme met à disposition une API permettant de construire des scénarios non déterministes (pilotage opportuniste), des scénarios déterministes (pilotage IA, rejeu d'une séquence d'attaques unitaires) et peut également être employée en mode interactif (pentest semi-automatisé). L'apprentissage s'appuie sur "Fast SI" et ses 44 attaques unitaires.
- L'outil d'apprentissage de pentest AD est opérationnel et est extensible pour supporter d'autres objectifs (ex : exfiltration de fichiers, exploits, etc) :
    - Le meilleur chemin d'attaque est appris, y compris lorsque le SI comporte plusieurs sous-réseaux. La prise de décision est explicable du fait du "Nouveau mécanisme d'attention" et du "Nouveau mécanisme de décision".
    - Grâce à 2 innovations notables ("Paramétrisation des actions" et "FAST SI"), l'outil développé permet d'exploiter un modèle appris, sur un Cyber Range, grâce au "Transfer Learning".
- La démonstration mettra l'accent sur l'apprentissage dans Fast SI (représentation visuelle de l'attaque, courbes d'apprentissage du nombre d'actions pour compromission) et sa validation dans le Cyber Range.

4. Bilan scientifique

Les points positifs sont les suivants :

- Des innovations majeures en IA (notamment en Reinforcement Learning) ont été introduites pour satisfaire les contraintes spécifiques au pentest :
    - Nette amélioration de l'algorithme "*CBWAR: Classification de Binaires Windows via Apprentissage par Renforcement. In Computer & Electronics Security Applications Rendez-vous. C&ESAR 2018*".
    - Nombre d'actions variables et paramétrisation d'action : cela a permis de démontrer qu'une IA à base de Reinforcement Learning était capable d'optimiser le chemin d'attaque AD quelles que soient la topologie du SI visé et la nature des machines et services qui le composent.
    - Apprentissage de corrélations {Observations multi variant ; Actions ; Récompenses} explicables participant aux prises de décisions également explicables.
- Il a été démontré qu'une IA combinant plusieurs techniques (Reinforcement Learning, Genetic Programing et Mécanismes d'Attention) est capable d'apprendre le chemin de compromission optimal.
- Le simulateur "Fast SI" offre une possibilité d'apprentissage pentest accéléré.

Les espoirs déçus sont les suivants :

- L'amélioration des performances des Cyber Range s'est révélée être un échec. Nous n'avons pas été en mesure de trouver une solution pour accélérer les retours de l'environnement SI simulé avec des conteneurs virtuels.
- Les POC "Systèmes de récompenses" ne se sont pas révélés concluants pour le cas d'usage d'apprentissage Pentest. Cependant, des idées d'emploi dans un contexte multi-objectifs commencent à émerger.

De nombreux pans de recherche sont encore ouverts :

- Adaptation à la réaction de la défense. A moyen terme, il est prévu de pouvoir confronter l'IA développée à des plateformes de défense automatisant les réactions de défense pilotées par des IA RL.
- IA multi-objectifs et multi-tâche : niveaux de furtivité combinés à des compromissions variées.
- Paramétrisation d'actions à plusieurs paramètres.
- Interaction humain/IA.