Proposition de sujet de thèse :

Reconfiguration automatisée des scénarios de formation en cybersécurité, à partir de l'apprentissage automatique des données physio-psychologiques des opérateurs

Encadrants : David Espes, Christine Chauvin, Philippe Rauffet, Philippe Le Parc

Démarrage : Octobre 2021

Mots clés : Machine Learning et analyse des signaux physiologiques, Classification des tâches, Entraînement des opérateurs en cybersécurité, Reconfiguration automatisée de scénario de formation.

1 Sujet de Thèse

1.1 Contexte

Ces dernières années, le nombre de cyber-attaques n'a cessé d'augmenter. La complexité de ces dernières a également fortement évoluée [1]. Avec l'apparition des attaques persistantes et complexes, ces dernières sont suffisamment discrètes pour contourner les mécanismes de sécurité et compromettre leur capacité à les détecter. Les attaques sont donc bien souvent détectées lorsque la cible a déjà été atteinte ce qui oblige les équipes opérationnelles à fournir une réponse rapide et pertinente. Une telle réponse est délicate et repose très souvent sur l'expertise des opérateurs [2].

La cybersécurité repose de plus en plus sur des centres opérationnels de sécurité (SOC) qui se composent d'opérateurs dont le rôle est de répondre aux attaques lorsqu'elles sont détectées. Cependant l'environnement auquel sont confrontés ces opérateurs est en perpétuelle évolution. Les outils pour identifier l'attaque et pour prendre une décision ne cessent d'évoluer. De même, les procédures pour mener à bien les attaques varient fortement avec le temps et se complexifient grandement.

Un enjeu essentiel pour les acteurs de la cybersécurité est donc d'assurer la formation tout au long de la vie des opérateurs en place que ce soit durant leurs études ou durant leur vie active [3]. De nombreux travaux se sont focalisés sur les techniques à employer durant les études des opérateurs afin qu'ils soient formés à répondre aux attaques [4]. Ces techniques reposent majoritairement sur l'utilisation de jeux sérieux [5, 6, 7]. Cependant la formation des opérateurs ne peut se limiter à la formation prodiguée durant leurs études, car elle est bien souvent générique. Elle doit être adaptée aux instances auxquelles les opérateurs sont confrontés (c'est-à-dire que les mises en situation enseignée doivent correspondre au mieux à l'environnement dans lequel ils vont opérer) mais elle doit également être en capacité de renforcer les faiblesses de chaque opérateur. Les formations ou entraînements prodigués ne peuvent donc malheureusement pas être génériques et doivent être adaptés en fonction des opérateurs et de l'environnement dans lequel ils opèrent [8].

La formation doit poursuivre deux objectifs principaux : 1) Apprendre aux opérateurs à réagir aux nouvelles attaques pour en limiter l'impact et revenir à un état stable le plus rapidement possible, et

2) Former les opérateurs aux évolutions technologiques des outils d'analyse qu'ils utilisent au quotidien.

Aujourd'hui, la formation se fait très souvent par des mises en situation. Une partie de l'équipe d'opérateurs devient une RED team (c'est-à-dire qu'ils prennent le rôle de l'attaquant) et une autre partie joue le rôle d'une BLUE team (c'est-à-dire qu'ils prennent le rôle du défendeur) [9, 10, 11]. Une telle répartition des rôles n'est clairement pas adaptée à l'environnement de travail. En effet, le rôle principal dans lequel devrait se trouver l'opérateur du SOC est celui dans lequel il est au quotidien, c'est-à-dire qu'il devrait appartenir à la BLUE team. L'efficacité des formations actuelles est donc de 50% si on suppose que les temps passés en RED team et BLUE team sont partagés (ce qui est systématiquement le cas). Inversement, un opérateur qui réalise des tests de pénétration de réseaux devrait majoritairement appartenir à la RED team.

Les tâches quotidiennes de l'opérateur doivent donc être considérées tout au long de l'entrainement pour parfaire au mieux sa formation. Les scénarios d'entrainement doivent donc être conçus pour 1) renforcer la formation de l'opérateur au rôle qu'il a dans son travail quotidien (RED team / BLUE team), 2) améliorer l'expertise des opérateurs vis-à-vis des outils qu'ils manipulent au quotidien ou de leur évolution technologique, 3) identifier clairement les difficultés rencontrées dans un scénario d'entrainement et 4) fournir des explications précises en cas d'échec à un scénario.

1.2 Verrous

Comme nous l'avons expliqué plus haut, le principal problème des formations actuelles est qu'elles sont bien souvent très génériques. Les opérateurs remplissent uniquement un formulaire en fin de formation pour donner leur ressenti et les manques qu'ils ont pu rencontrer. En situation de stress et de difficultés, il leur est bien souvent difficile de faire un constat précis des problèmes qu'ils ont rencontrés [12, 13]. En effet, en réponse à un scénario d'attaques, les difficultés qu'un opérateur rencontre viennent soit : 1) de la complexité des outils utilisés, 2) de la difficulté à identifier correctement le type d'attaque et 3) de la difficulté à trouver la réponse appropriée à une attaque. Il est donc important de pouvoir identifier correctement quelle(s) phase(s) pose(nt) problème(s).

Les verrous d'un tel sujet de thèse sont donc de quatre ordres :

- **Identification des faiblesses des opérateurs dans le cadre des différentes situations de travail rencontrées**. Il faut pouvoir identifier les difficultés rencontrées par l'opérateur durant un scénario d'attaque, en distinguant les tâches ou les situations du scénario d'entraînement qui risquent de surcharger ou fatiguer l'opérateur dans son activité, de diminuer ses performances. Il est donc essentiel que le scénario d'entrainement repose sur une captation de l'environnement d'entrainement et des différentes sensations que rencontre l'opérateur durant l'entrainement. La difficulté est de trouver les bons paramètres qui vont permettre d'identifier la situation dans laquelle un opérateur se trouve en difficulté. La captation de ces paramètres doit être suffisamment précise et peu intrusive.
- **Prise en compte des facteurs psychologiques des opérateurs et des différences interindividuelles.** Le quotidien des opérateurs et les spécificités de chaque individu influent également sur les états mentaux des opérateurs et, par conséquent, sur les mesures des facteurs physiologiques. En effet, un individu ne va pas réagir de la même

manière à une situation de stress qu'un autre individu, chaque personne ayant une sensibilité différente à la contrainte (en fonction de son niveau d'expertise, de son expérience, etc.). De même, les problèmes du quotidien impactent chaque individu sur la qualité du travail qu'il est en capacité de fournir. L'analyse des mesures physiologiques doit donc tenir compte de ces facteurs extérieurs et ces différences interindividuelles afin de garantir la fiabilité des indicateurs mesurés.

- **Identification de l'environnement de travail et des instances dans lesquelles les opérateurs interviennent.** Les méthodes d'entrainement utilisées aujourd'hui sont principalement dédiées à un type de scénario donné qui peut largement différer des conditions opérationnelles des opérateurs. La grande majorité des scénarios d'entrainement sont destinés aux systèmes de contrôles industriels [14, 15, 16, 17]. Afin de garantir un entraînement individualisé, il est important de pouvoir identifier dans quel environnement travaille l'opérateur au quotidien. En effet, les entrainements prodigués doivent être le plus réalistes possibles et le plus proche des attaques auxquelles ils vont être confrontés. La difficulté de cette tâche est de pouvoir identifier de manière passive et automatique les instances dans lesquels les opérateurs travaillent.
- **Réalisation automatisée des scénarios dynamiques d'entrainement**. Le dernier point complexe à réaliser consiste en le déploiement automatisé des ressources nécessaires à la mise en œuvre des scénarios d'attaques. Un tel déploiement ne peut être générique et doit être adapté à chaque opérateur. Le déploiement doit également être dynamique, c'est-à-dire qu'il doit évoluer durant la durée de l'entrainement en fonction des paramètres identifiés sur les faiblesses de l'opérateur et sur l'environnement de travail de ce dernier.

1.3 Solutions

Les solutions proposées dans cette thèse vont répondre aux trois verrous énoncés précédemment. Une ébauche des solutions envisagées est proposée ci-dessous :

- **Identification des faiblesses de l'opérateur :** Quelques travaux de recherche [18, 19] ont proposé de mesurer l'efficience des opérateurs en cybersécurité pour pouvoir l'améliorer. Ces mesures reposent sur le traitement d'indicateurs de charge de travail, de stress, de performance… Ces travaux montrent un réel impact du niveau de stress sur l'efficacité et la qualité de la réponse réalisée. Il est donc important de détecter cet impact, et la survenue d'état cognitif délétère, qui peuvent dégrader la performance des opérateurs en cybersécurité lors de la survenue de menaces ou d'attaques. Pour répondre à ce premier verrou, il faudrait donc dans un premier temps pouvoir modéliser et évaluer les variations de la contrainte de la situation de travail : on pourra notamment distinguer différentes situations de cyberdéfense, en termes de densité d'évènements cybers, de nature de la menace/attaque, etc. qui génèrent un changement de l'état opérateur (au niveau comportement ou physiologique). On doit également pouvoir mesurer les conséquences de ces variations de contraintes sur l'opérateur, en termes de changements comportementaux (en lien notamment avec l'attention visuelle, l'activité motrice sur le clavier ou la souris ou le contrôle cognitif) ou physiologiques (variabilité du rythme cardiaque, oxygénation cérébrale, réponse pupillaire). Ce premier axe pourra notamment s'appuyer sur une première étude dans le domaine cyber, qui avait été menée au sein du simulateur C4 de l'UBS, et qui analysait l'effet de différents niveaux de criticité d'un

scénario de formation (phase de consolidation des défenses notamment en début d'exercice, de détection d'une menace imminente avec l'apparition de signaux faibles, ou de détection d'une attaque) sur le rythme cardiaque [20], mais aussi les communications et les actions de soutien au sein d'une BLUE team [21]. Dans un autre contexte, celui de la formation des pilotes de chasse sur avions Rafale, des travaux avaient également été réalisés pour identifier les tâches collectives qui généraient le plus de charge mentale et de dégradation du COFOR (Common Frame of Reference) en phases de vol ou en phases d'engagement d'une cible [22].

- **Prise en compte des différences interindividuelles :** Un autre point, qui n'est pas pris en compte dans les études actuelles, concerne les facteurs psychologiques qui sont susceptibles d'influer sur le comportement des opérateurs mais aussi sur les mesures. L'environnement quotidien de l'opérateur influe sur les facteurs physiologiques du stress. Il est donc important d'en tenir compte afin d'identifier les facteurs externes influant sur le stress et la sensibilité au stress de l'opérateur. Nous proposerons une méthodologie pour introduire ces facteurs psychologiques externes à la situation cyber avant de réaliser des mesures physiologiques qui permettront de pondérer les mesures réalisées. On pourra notamment s'appuyer sur des mesures de baselines pour la partie physiologique, ou utiliser des questionnaires pour évaluer la condition initiale de l'opérateur au début du scénario (qui peut varier en raison de facteur personnels, ou de la qualité de l'hygiène de vie en termes de sommeil, d'alimentation, etc.).
- **Identification de l'environnement de travail et des instances dans lesquelles les opérateurs interviennent :** Afin d'avoir un entrainement très proche des instances sur lesquelles opèrent l'opérateur, et toujours en vue d'un entrainement individualisé, il sera nécessaire d'avoir une bonne vision de l'environnement de travail. La mise en place de sondes au sein du SOC (et chez le client si le contrat entre le SOC et ce dernier le permet) aura pour vocation de réaliser une captation suffisamment fine pour identifier le type d'environnement. Pour cela, une solution d'inférence reposera sur l'utilisation de techniques utilisant le machine learning pour l'apprentissage de l'environnement de travail et la spécification pour l'identification des protocoles utilisés et spécifiques à l'environnement. L'utilisation de telles techniques permettra ainsi de profiler chaque opérateur, en identifiant les tâches ou les situations qui sont les plus critiques pour chacun. Par ailleurs, cela contribuera également à distinguer les situations très rarement rencontrées, qui peuvent amener une dégradation des performances à cause d'une faible familiarité à la situation rencontrée par l'opérateur. Il sera intéressant, d'identifier, dans de telles situations, les modes cognitifs adoptés par les opérateurs (plus ou moins anticipatifs ou réactifs) et de les analyser au regard de leur performance d'une part et de leur charge mentale d'autre part.
- **Réalisation automatisée des scénarios dynamiques d'entrainement :** A notre connaissance, il n'existe pas de modèle permettant de prendre en compte la dynamique de la formation d'opérateur de SOC. Un tel modèle sera dynamique, c'est-à-dire que le scénario d'entrainement évoluera en fonction des paramètres physiologiques et psychologiques captés durant l'entrainement. Ainsi, il permettra de viser principalement les difficultés rencontrées par les opérateurs et de s'assurer qu'à travers la répétition, l'opérateur sera en mesure d'y répondre convenablement lors d'une situation réelle. Pour la mise en œuvre d'un tel modèle, les réseaux programmables SDN et les réseaux virtualisés NFV seront utilisés durant la thèse. Ces technologies permettront de faire varier l'infrastructure et d'entrainer l'opérateur à prévenir ce type d'attaque dans des situations différentes. En fonction de l'adversaire, une attaque peut provenir de l'extérieur ou de l'intérieur du système. Bien que l'attaque utilise le

même vecteur de propagation, sa préparation sera bien différente (apprentissage de la topologie, des équipements…). Les scénarios doivent être composés d'un jeu suffisant d'attaque pour identifier les faiblesses de l'opérateur et suffisamment flexibles pour réaliser ces attaques sur différentes architectures. Cette richesse du jeu d'attaque pourra être proposée en combinant plusieurs outils utilisés par les Red Team [23].

Grâce aux solutions proposées, nous pensons pouvoir améliorer significativement l'efficacité des méthodes d'entrainement des opérateurs de SOC. Cette thèse a pour vocation de pouvoir modéliser et bâtir les fondations de l'opérateur cyber 2.0.

2 Cohérence de l'équipe encadrante

La difficulté d'un tel sujet repose sur sa pluridisciplinarité. Il y a trois champs disciplinaires qui interviennent : 1) la cybersécurité due au contexte spécifique des SOC, 2) le traitement du signal et le machine learning pour l'étude des mesures comportementales et physiologiques des opérateurs 3) l'ergonomie cognitive pour modéliser le travail cognitif et l'impact en termes d'état mental de l'opérateur avant l'exercice et après ce dernier.

Le montage de l'équipe encadrante a été réalisé pour prendre en compte cette pluridisciplinarité. L'encadrement du (de la) futur(e) doctorant(e) sera conjoint entre les équipes IRIS et FHOOX du Lab-STICC. Les travaux de l'équipe IRIS portent sur la protection, la défense et la résilience des infrastructures critiques qui sont bien souvent l'objet d'analyse des opérateurs de SOC. L'équipe FHOOX travaille à modéliser et optimiser la coopération humain-système. Ils étudient l'utilisation de mesures physiologiques ou psychologiques pour améliorer l'efficacité des opérateurs d'un système.La thèse sera ainsi dirigée par :

- Christine Chauvin (UBS – FHOOX)
- Philippe Le Parc (UBO – IRIS)

Elle sera également encadrée par :

- Philippe Rauffet (UBS – FHOOX)
- David Espes (UBO – IRIS)

Un tel sujet, de par sa thématique et son encadrement, est en totale cohésion avec les activités stratégiques de recherche de l'Alliance Universitaire de Bretagne (AUB). Cette thèse permet ainsi de créer une dynamique de recherche entre les deux établissements (UBO et UBS) qui la composent.

Références

[1] W. Tounsi and H. Rais, "A survey on technical threat intelligence in the age of sophisticated cyber attacks", in Elsevier Journal on Computers & Security, 2018

[2] M. Takano, "ICS cybersecurity incident response and the troubleshooting process", in IEEE Proceedings of the SICE Annual Conference (SICE), Japan, 2014

[3] G. M. Deckard, “Cybertropolis: breaking the paradigm of cyber-ranges and testbeds”, in IEEE International Symposium on Technologies for Homeland Security (HST), USA, 2018

[4] T. Pereira, H. Santos, I. Mendes, “Challenges and reflections in designing Cyber security curriculum”, in IEEE World Engineering Education Conference (EDUNINE), Brazil, 2017

[5] G. M. Deckard and L. J. Camp, “Measuring efficacy of a classroom training week for a cybersecurity training exercise”, in IEEE Symposium on Technologies for Homeland Security (HST), USA, 2016

[6] A. S. Andreatos, “Designing educational scenarios to teach network security”, in IEEE Global Engineering Education Conference (EDUCON), Greece, 2017

[7] A. Nagarajan, J. M. Allbeck and al., “Exploring Game Design for Cybersecurity Training”, IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Thailand, 2012

[8] NATO, “NATO Cooperative Cyber Defence Centre of Excellence”, in 11th International Conference on Cyber Conflict (CyCon), Estonia, 2019

[9] N. Veerasamy, “High-level Methodology for Carrying out Combined Red and Blue Teams”, in Proceedings of the Second International Conference on Computer and Electrical Engineering, Dubai, 2009

[10] A. Waksman, J. Rajendran et al., “A Red Team/Blue Team Assessment of Functional Analysis Methods for Malicious Circuit Identification”, 51st ACM/EDAC/IEEE Design Automation Conference (DAC), 2014

[11] J. Mirkovic, P. Reiher et al., “Testing a Collaborative DDoS Defense in a Red Team/Blue Team Exercise”, in IEEE Transactions on Computers, 2008

[12] A. Rege, S. Biswas et al., “Using Simulators to Assess Knowledge and Behavior of Novice Operators of Critical Infrastructure under Cyberattack Events”, IEEE Resilience Week (RWS), USA, 2017

[13] D. S. Henshel, M. Deckard et al., “Predicting Proficiency in Cyber Defense Team Exercises”, IEEE Military Communications Conference (MILCOM’16), USA, 2016

[14] A. Ferreira, F. A. C. R. Costa et al., “Use of Simulation to Achieve Better Results in Cyber Military Training”, USA, 2015

[15] J. Kim, K. Kim and M. Jang, “Cyber-Physical Battlefield Platform for Large-Scale Cybersecurity Exercises”, 11th International Conference on Cyber Conflict (CyCon), 2019

[16] K. Maennel, R. Ottis and O. Maennel, “Improving and Measuring Learning Effectiveness at Cyber Defense Exercises”, Nordic Conference on Secure IT Systems, 2017

[17] National Security Research Institute, “Cyber Conflict Exercise 2018”, 2018.

[18] J. Dykstra and C. L. Paul, “Cyber Operations Stress Survey (COSS): Studying fatigue, frustration, and cognitive workload in cybersecurity operations”, 11th USENIX Workshop on Cyber Security Experimentation and Test, USA, 2018.

[19] C. L. Paul and J. Dykstra, “Understanding Operator Fatigue, Frustration, and Cognitive Workload in Tactical Cybersecurity Operations”, Journal of Information Warfare, 2017

[20] Deline S., Guillet L., Rauffet P., Guerin C. (2016). Le stress dans un contexte de cyberdéfense: relations entre mesures subjectives et physiologiques, Congrès Annuel de Psychologie, Paris (France).

[21] Deline S., Guillet L., Rauffet P., Guérin C. (2019). Team cognition in a cyber defense context: Focus on social support behaviors. Cognition, Technology and Work.

[22] Lassalle J., Rauffet P., Leroy B., Guérin C., Chauvin C., Coppin G., Saïd F. (2017). COmmunication and WORKload analyses to study the COllective WORK of fighter pilots: the COWORK2 method. Cognition, Technology and Work. doi: https://doi.org/10.1007/s10111-017-0420-8

[23] J. Yuen, B. Turnbull and J. Hernandez, “Visual Analytics for Cyber Red Teaming”, IEEE Symposium on Visualization for Cyber Security (VizSec), USA, 2015