Demande n°3151 - ARED

Acronyme :DynHet

Identification

Identification

N°	3151
Année	2023
Statut	Soumis à l'établissement
Date de création	26/01/2023
Date de modification	30/01/2023
Date de soumission	30/01/2023
Acronyme	DynHet
Intitulé du projet	Détection d'intrusions au niveau système d'informations : détection d'anomalies par traitement IA dans des graphes dynamiques hétérogènes
Dispositif	ARED
Civilité du porteur	Monsieur
Nom du porteur	Hurfin
Prénom du porteur	Michel
Courriel	michel.hurfin@inria.fr

Priorité régionale

Priorité régionale

Priorité régionale	Cybersécurité
Regional priority	Cybersécurité

Stratégie régionale de recherche et d'innovation (S3)

Stratégie régionale de recherche et d'innovation (S3)

Domaine d'innovation stratégique	Économie numérique sécurisée et responsable
Domaine d'innovation stratégique	.
Levier thématique	Cybersécurité

Organisme de tutelle

Organisme de tutelle

Etablissement	INRIA
Commune	RENNES
Civilité	Madame
Nom	Morin
Prénom	Christine

Unité de recherche

Unité de recherche

Nom de l'unité	CIDRE
Civilité	Madame
Nom	Viet Triem Tong
Prénom	Valérie
Tél.	0299844573
Courriel	valerie.viettriemtong@centralesupelec.fr

Données complémentaires de l'unité de recherche

Libellé	Valeur
Statut et numéro de l'unité de recherche (UMR, USR, EA?)	Equipe Projet Inria Numéro RNSR: 201120989W
Le cas échéant, nom de l'équipe de recherche	CIDRE https://team.inria.fr/cidre/
Le cas échéant, nom du responsable de l'équipe de recherche	Madame Valérie Viet Triem Tong
Activités de recherche et points forts de l'équipe	Les travaux de l'équipe CIDRE visent à mettre en évidence les problèmes de sécurité et à proposer des contre-mesures préventives ou réactives pour les systèmes les plus largement utilisés. Trois axes principaux: 1) Compréhension de l'attaque 2) Détection d'attaque 3) Résistance à l'attaque
Si implication dans un/plusieurs programmes européens ou internationaux, préciser	2019-2022: Projet H2020 SPARTA (réseau de compétences en cybersécurité)

Co-directeur de thèse

Co-directeur de thèse

Civilité	Madame
Nom	Viet Triem Tong
Prénom	Valérie
Unité de recherche	Equipe Projet Inria CIDRE
Etablissement de tutelle	CENTRALE SUPELEC

Description

Description du projet

Libellé	Valeur
Résumé synthétique du projet (2000 caractères max)	La détection d'intrusions est un enjeu majeur de la défense en cybersécurité. De par la complexification des systèmes et l'utilisation de nouvelles briques technologiques, de nouvelles menaces apparaissent continuellement pouvant compromettre tous les niveaux d'un système d'information. L'activité d'un système d'informations ou de ses composants peut être modélisée sous la forme d'un graphe, qui peut être utilisé pour détecter des anomalies éventuellement symptomatiques d'une intrusion. Par le passé, la structure de graphes a ainsi été utilisée pour modéliser des communications réseau, des authentifications utilisateurs, l'activité au niveau d'un système d'exploitation ou celle d'un système d'informations plus globalement. Les graphes que nous envisageons de construire sont hétérogènes car les logs considérés et la définition des noeuds et arêtes des graphes sont multiples. Les graphes sont dynamiques car ils doivent permettre de modéliser le changement au cours du temps des interactions (arêtes du graphe) entre les éléments d'un système et l'évolution du système via l'ajout ou la suppression d'éléments (noeuds du graphe). Dans ce contexte, l'objectif de cette thèse est donc d'élaborer de nouveaux algorithmes de détection d'anomalies sur les graphes dynamiques hétérogènes.
Porteur du projet : date d'obtention de l'HDR ou dérogation	HDR obtenue en 2004
Hypothèses, questions posées, points de blocage, approche méthodologique, technique (4000 car. max)	Des premiers travaux, tels que l'approche Temporal Graph Network (TGN) [2020 - Rossi, Chamberlain, Frasca, Eynard, Monti, Bronstein], se sont intéressés à la prise en compte de la dynamique dans des graphes homogènes pour des tâches de classification de nœuds ou d'arêtes ou des tâches de reconstruction/prédiction d'arêtes. Cependant, bien que la dimension dynamique soit prise en compte au niveau de la temporalité des interactions entre éléments du graphe et permette de meilleures performances générales sur ces tâches, TGN ne permet pas de prédire des anomalies de types temporelles. D'autres méthodes directement développées pour la détection d'anomalies dans les graphes dynamiques [2021 - Liu, Pan, Wang, Xiong, Wang, Chen, Lee] nécessitent des ressources informatiques trop importantes pour être utilisées sur des graphes de grandes tailles et/ou manquent d'informations concernant l'évaluation de l'approche développée. Certaines méthodes nécessitent également une annotation du jeu de données d'entrainement [2022 - Anjum, Iqbal, Hamelin, Benoit] rendant difficile l'évaluation de leurs capacités d'adaptation à de nouvelles menaces. Le passage à l'échelle des algorithmes est également un problème, certaines méthodes n'étant entrainées et évaluées que sur une sous-partie d'un jeu de données temporelles conséquent. Nous proposons de développer de nouvelles méthodes robustes de détection d'anomalies dans les graphes dynamiques permettant d'être appliquées aux systèmes informatiques. Ces nouvelles approches ont pour objectifs de détecter des anomalies de différentes natures (anomalies structurelles, sémantiques et temporelles), correspondant à des interactions ou des dynamiques d'interactions nouvelles, c'est-à-dire non observées et/ou non attendues. Par ailleurs, les méthodes développées devront être applicables à de larges graphes hétérogènes pour offrir des solutions applicables à de réels cas d'usage cyber. D'un point de vue cyber, il faudra définir le ou les graphes permettant la représentation de l'activité du système d'information en reliant les aspects système, réseau et applicatifs. La définition du ou des graphes pourra prendre en compte les flux d'information entre les différentes entités représentées et cherchera à prendre en compte la temporalité à court terme et à long terme. Pour évaluer les modèles IA et maîtriser les capacités de détection, avant d'appliquer ces modèles à des données cyber, il sera certainement nécessaire de débuter les travaux de thèse par le développement d'un générateur de graphes temporels configurable. Les graphes temporels pourront être homogènes dans un premier temps pour faciliter la conception des modèles puis être hétérogènes. Les graphes simulés devront approximer ceux construits sur les données provenant des systèmes d'informations afin de contrôler au mieux les aspects

	systèmes d'informations afin de contrôler au mieux les aspects dynamiques de la détection d'anomalies. On développera d'abord la méthode de détection d'anomalies sur des graphes temporels homogènes puis sur des graphes hétérogènes. Enfin, la méthode sera évaluée sur des systèmes réels notamment en se basant sur des datasets académiques existants.
Environnement scientifique, positionnement dans contexte régional/national/international (2000 car.)	L'objectif est d'améliorer la détection des attaques avancées de type APT (Advanced Persistent Threat) en prenant en compte toutes les sources de logs disponibles au sein d'un système d'information (logs applicatifs, système et réseau). Ces informations peuvent être mises en relation dans une structure de graphe qui représente l'activité globale du système. La détection d'anomalies dans des graphes dynamiques hétérogènes est un sujet en plein essor dans le domaine de l'apprentissage machine. Le sujet de thèse se situe à l'intersection: il s'agit de définir des graphes intéressants pour la détection d'APT et de développer de nouvelles méthodes de détection d'anomalies sur ces graphes. Le candidat pourra soumettre ses travaux aux conférences internationales en sécurité (RAID, DIMVA, Usenix SEC, ACSAC, NDSS, etc.) mais aussi aux conférences en IA (NeurIPS, ICML, KDD, ICASSP, etc.). Il pourra présenter ses travaux à la communauté francophone au sein du GDR Sécurité ou des conférences telles que RESSI, C&ESAR et CAID. En plus des directeurs de thèse, Michel Hurfin (Inria) et Valérie Viet Triem Tong (CentraleSupélec), la thèse sera encadrée par Pierre-François Gimenez (CentraleSupélec) et Frédéric Majorczyk (DGA-MI). Les travaux seront aussi suivi par Barbara Pilastre (DGA-MI), Tristan Bitard-Feildel (DGA-MI).
Collaborations scientifiques (nature/partenariat/pays) et partenariat socio-économique envisagé	Les travaux de recherche permettront de renforcer les liens entre les chercheurs et enseignants-chercheurs de l'équipe Cidre (trois sont impliqués directement) et les membres de la DGA-MI (trois sont impliqués directement) .
Adéquation du projet avec le DIS de rattachement (3000 caractères maximum)	Le projet porte sur des problématique de Cyber-sécurité. Il a été pré-sélectionné par le conseil scientifique de CREACH LABS.
Si priorité régionale, préciser (200 caractères maximum)	Le projet porte sur des problématique de Cyber-sécurité. Il a été pré-sélectionné par le conseil scientifique de CREACH LABS.
Cotutelle internationale (indiquer "oui" ou "non"; si oui, préciser le pays et l'établissement)	non
Nom(s) du(des) cofinanceur(s) envisagé(s) ou confirmé(s)	
Si cofinancement industriel, justifier le choix du dispositif ARED au lieu du dispositif CIFRE	
État du cofinancement (indiquer "acquis" ou "demandé")	
Si demandé, date prévisionnelle de réponse du cofinanceur	
Si cofinancement refusé, autre(s) source(s) de cofinancement identifiée(s)	
Si reversement prévu, préciser (partenaire, montant)	

Candidat

Collaboration internationale

Collaboration internationale

Pays	
Etablissement partenaire à l'étranger	

Pièces jointes

Date de mi	Type de pièce	Commentaire	Document	Valide	Commentaire
30/01/2023	ARED - Dépôt - CV du porteur	Obligatoire - Le CV doit préciser si le porteur est titulaire de l'HDR et, le cas échéant, depuis quand, et mentionner la liste des doctorant·e·s déjà encadré·e·s et des publications associées.	CV-hurfin-02.pdf	Non traité	
30/01/2023	ARED - Dépôt - Avis directeur·rice d'unité	15 lignes maximum	lettre-soutien-dynhet-v2t2.pdf	Non traité	
30/01/2023	ARED - Dépôt - Justificatif cofinancement	Facultatif - Dépôt de demande, justificatif d'obtention, courrier du cofinanceur pressenti...		Non traité	
30/01/2023	ARED - Dépôt - Lettre partenaire étranger	Facultatif - Lettre d'engagement de l'établissement partenaire étranger en cas de cotutelle internationale		Non traité	
30/01/2023	ARED - Dépôt - Pièce complémentaire	Facultatif	CV-Co-Directrice.pdf	Non traité	
30/01/2023	ARED - Dépôt - Pièce complémentaire	Facultatif	CV-Encadrants.pdf	Non traité	
30/01/2023	ARED - Instruction - Lettre de saisine	Lettre de l'établissement à transmettre à la Région (utiliser le modèle disponible dans la rubrique "Documents à télécharger")		Non traité	
30/01/2023	ARED - Versement - Convention de partenariat	Le cas échéant (si autorisation de reversement)		Non traité	
30/01/2023	ARED - Versement 1/2 - CV doctorant.e	Pour versement tranche 1 (cohortes jusqu'en 2018) Pour versement tranche 2 (cohortes à partir de 2019)		Non traité	
30/01/2023	ARED - Versement 1/2 - Copie du contrat doctoral (ou CDD ou attestation))	Pour versement tranche 1 (cohortes jusqu'en 2018) Pour versement tranche 2 (cohortes à partir de 2019)		Non traité	
30/01/2023	ARED - Versement 1 - Lettre d'approbation école doctorale	Pour versement tranche 1 (cohortes jusqu'en 2018)		Non traité	
30/01/2023	ARED - Versement 2 - Attestation bon déroulement travaux année 1	Pour versement tranche 2		Non traité	
30/01/2023	ARED - Versement 2 - Bilan financier année 1	Pour versement tranche 2 ARED		Non traité	
30/01/2023	ARED - Versement 3 - Attestation bon déroulement travaux année 2	Pour versement 50% tranche 3 ARED		Non traité	

30/01/2023	ARED - Versement 3 - Bilan financier année 2	Pour versement 50% tranche 3 ARED		Non traité	
30/01/2023	ARED - Versement 4 - Bilan financier année 3	Pour versement solde année 3 ARED		Non traité	
30/01/2023	ARED - Versement 4 - Certificat administratif précisant statut thèse	Pour versement du solde de la tranche 3 (à accompagner des pièces justificatives correspondantes pour les cohortes jusqu?en 2018 : avis du jury, attestation de réinscription en ED, justificatif de prolongation ou d'abandon)		Non traité	
30/01/2023	ARED - Versement 4 - Avis du jury si thèse soutenue	Pour versement solde tranche 3 ARED (cohortes jusqu'en 2018)		Non traité	
30/01/2023	ARED - Versement 4 - Attestation de réinscription en ED	Pour versement solde tranche 3 ARED (cohortes jusqu'en 2018)		Non traité	
30/01/2023	ARED - Versement 4 - Justificatif de prolongation	Pour versement solde tranche 3 ARED (cohortes jusqu'en 2018)		Non traité	
	ARED - Versement 4 - Justificatif d'abandon de thèse	Pour versement solde tranche 3 ARED (cohortes jusqu'en 2018)		Non traité	