PolyPublie
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title:	Proposition d'approche pour la détection de menace interne
Auteur: Author:	Amine Badaoui
Date:	2021
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Badaoui, A. (2021). Proposition d'approche pour la détection de menace interne [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/9995/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/9995/
Directeurs de recherche: Advisors:	Nora Boulahia Cuppens, & Michel Gagnon
Programme: Program:	Génie informatique

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Proposition d'approche pour la détection de menace interne

AMINE BADAOUI
Département de génie informatique et génie logiciel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie informatique

Décembre 2021

© Amine Badaoui, 2021.

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Ce mémoire intitulé :

Proposition d'approche pour la détection de menace interne

présenté par **Amine BADAOUI**
en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Soumaya CHERKAOUI, présidente
Nora BOULAHIA-CUPPENS, membre et directrice de recherche
Michel GAGNON, membre et codirecteur de recherche
Hanifa BOUCHENEB, membre

DÉDICACE

Une dédicace à mes Parents.

À mon père qui nous a quitté que Dieu lui fasse miséricorde.

REMERCIEMENTS

Tout d'abord je tiens à remercier Dieu tout puissant et ma famille.

À Polytechnique Montréal

Je tiens à exprimer ma profonde gratitude à ma directrice de recherche la **Professeure Nora Boulahia Cuppens** qui a supervisé ces travaux et qui m'a toujours encouragé. J'associe à ces remerciements, le **Professeur Frédéric Cuppens** qui s'est impliqué aussi dans le projet.

Je remercie le **Professeur Michel Gagnon** pour son engagement sur l'orientation du projet.

Ce projet a démarré avec le **Professeur José Fernandez** et je le remercie.

À Desjardins

Je remercierai toujours Monsieur **François Charest** pour son soutien durant mon passage à Desjardins, il a su être là en tout moment.

Je remercie Monsieur **Frédéric Michaud** pour son apport et surtout sa présentation sur la maturité du programme de menace interne.

Je remercie aussi Monsieur **Martin Rosa** avec qui j'ai eu des discussions de haut niveau sur l'investigation.

Je remercie toute personne avec qui j'ai eu la chance de collaborer au sein du réseau de Desjardins.

Je tiens à remercie toute personne dont je n'ai pas cité le nom. Un dernier remerciement pour le **Professeur Abdelkader Belkhir** mon ancien professeur.

Je remercie l'ensemble des mes amis qui m'ont supporté.

RÉSUMÉ

La menace interne est devenue récemment un sujet très abordé en cybersécurité suite aux différents incidents survenus dans de nombreuses entreprises. Les conséquences de cette menace pèsent aussi lourd que tout autre danger informatique. C'est pourquoi les industriels et les académiques s'y sont intéressés. Plusieurs solutions ont vu le jour sur le marché pour atténuer le risque et de multiples travaux de recherche ont été conduits pour répondre à cette problématique. Dans ce mémoire de maitrise, nous proposons une approche pour aider une grande entreprise à automatiser la détection de comportements malveillants dans un contexte de menace interne tout en réduisant le taux de faux positifs.

La première étape consistait à comprendre le programme de gestion des incidents de menace interne chez l'entreprise afin d'apporter des améliorations compatibles avec l'existant. Par la suite, nous avons fait appel aux graphes de convolution - GCN - afin de proposer un modèle qui évolue sur les attributs et le contexte des utilisateurs dans le réseau. Pour inclure le contexte de travail des usagers dans le graphe qu'on passe à notre GCN, nous avons eu à explorer plusieurs modélisations possibles basées sur les relations hiérarchiques et d'affaires qui relient les employés de l'entreprise. Chaque relation choisie impliquait un graphe différent. En plus, les relations étaient équivalentes et ne démontraient pas le niveau de voisinage entre les employés dans l'entreprise. Finalement, nous avons réussi à transformer l'ensemble des relations en une seule par le biais de Node2vec, qui à partir du graphe de l'ogranisation associe un vecteur à chaque noeud.

Nous avons obtenu des résultats divers que nous avons comparés aux modèles supervisés d'apprentissage machine. La dernière phase de nos travaux de recherche consistait en la création d'une ontologie ou une base de connaissances suite aux défis rencontrés lors de la collection des évènements produits par un usager en particulier sur les différents systèmes déployés dans le réseau de l'entreprise.

Dans l'objectif d'améliorer nos résultats, nous proposons de construire un nouveau jeu de données pour entrainer les différents modèles. Ce jeu de données se base sur l'activité des analystes lors de l'investigation d'un incident. L'activité sera extraite à travers un assistant virtuel qui permet de communiquer avec l'ontologie pour répondre aux différentes requêtes émises par les analystes.

ABSTRACT

The insider threat has recently become a very hot topic in cybersecurity following the various incidents that have occurred in many companies. The consequences of this threat are as heavy as any other computer danger. This is why industrialists and academics are interested in it. Several solutions have emerged on the market to mitigate the risk and multiple research studies have been carried out to address this problem. In this master's thesis, we propose an approach to help a large company automate the detection of malicious behavior in an insider threat context while reducing the rate of false positives.

The first step was to understand the internal threat incident management program at the company in order to make improvements compatible with the existing one. Subsequently, we used convolution graphs - GCN - in order to propose a model which evolves on the attributes and the context of users in the network. To include the work context of users in the graph that we pass to our GCN, we had to explore several possible models based on the hierarchical and business relationships that link the employees of the company. Each relation chosen involved a different graph. In addition, the relations were equivalent and did not demonstrate the level of proximity between the employees in the company. Finally, we succeeded in transforming the set of relations into a single one by means of Node2vec, which from the graph of the ogranization associates a vector with each node.

We obtained various results which we compared to supervised machine learning models. The last phase of our research consisted in the creation of an ontology or a knowledge base following the challenges encountered during the collection of events produced by a user in particular on the different systems deployed in the company's network.

In order to improve our results, we propose to build a new dataset to train the different models. This dataset is based on the activity of analysts during the investigation of an incident. The activity will be extracted through a virtual assistant which makes it possible to communicate with the ontology to respond to the various requests made by the analysts.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

CERT	Computer Emergency and Response Team
CMDB	Configuration and Management Database
CSIHO	Computer Security Incident Handling Ontology
CSV	Comma-Separated Values
DAML	DARPA Agent Markup Language
DARPA	Defense Advanced Research Projects Agency
DLP	Data Loss Prevention
EDR	Endpoint Detection and Response
FBI	Federal Bureau of Investigation
HIPAA	Health Insurance Portability and Accountability Act
IDMEF	Intrusion Detection Message Exchange Format
IDS	Intrusion Detection System
IDWG	Intrusion Detection Message Exchange Format Working Group
IETF	Internet Engineering Task Force
IOT	Internet Of Things
ITSDA	Insider Threat Security and Defense Architecture
KNN	K-Nearest Neighboors
LDAP	Lightweight Directory Access Protocol
OIL	Ontology Inference Layer
OSI	Open Systems Interconnection
OWL	Ontology Web Language
PCI	Payment Card Industry Data Security Standard
RDF	Ressource Description Framework
RDF	Ressource Description Framework Schema
ROC	Receiver Operating Characteristic
RUU	Are You You
ReLU	Rectified Linear activation Unit
SEM	Security Event Management
SIEM	Security Information and Event Management
SIM	Security Information Management
SIRPs	Security Incident Response Platforms
SOA	Security Orchestration and Automation
SOAR	Security Orchestration, Automation and Response

SOC	Security Operations Center
SWRL	Semantic Web Rule Language
TIPs	Threat Intelligence Platforms
TWOS	The Wolf Of SUTD
UEBA	User and Entity Behavior Analytics
XML	Extensible Markup Language
XSWRL	Extended Semantic Web Rule Language
DGL	Deep Graph Library
CC	Carbon Copy
BCC	Blind Carbon Copy
Adam	Adaptive Moment Estimation

LISTE DES ANNEXES

CHAPITRE 1 INTRODUCTION

Les technologies de l'information participent significativement à l'amélioration du travail au sein des entreprises. Présentement, toutes les activités d'une organisation sont liées à son système informatique, au point qu'il est impossible de les dissocier. Toutefois, des menaces pèsent sur ce système et peuvent avoir de graves conséquences et des impacts très coûteux pour l'entreprise, ses partenaires et ses clients. La protection d'un tel système fait déjà partie des actions considérées par les responsables de sécurité. De nombreuses solutions sont déployées pour détecter ou prévenir les différents types de menaces. On retrouve notamment les pare-feux sur les frontières du réseau ou entre certaines zones pour filtrer les communications sortantes ou entrantes, des solutions de détection d'intrusion au niveau du réseau pour détecter ou bloquer les attaques ciblant les actifs ou des agents installés sur les machines pour contenir les programmes malveillants tels les virus. L'ensemble des solutions de protection et de détection existantes réussissent dans la majorité des cas à identifier et bloquer les activités malveillantes et s'adaptent aux changements dans l'environnement qui peuvent apporter aussi leurs lots de risques comme le passage aux services infonuagiques directement ou à travers de nouvelles installations.

Ces moyens de mitigation perdent beaucoup en efficacité ou échouent lorsque les attaques sont effectuées par un usager qui possède un accès privilégié. Bien qu'ils offrent un niveau de sécurité même en interne en bloquant les tentatives d'accès non autorisées par exemple, ils restent inefficaces pour déterminer les actions malveillantes des utilisateurs.

En effet, les protections traditionnelles se focalisent plus sur des attaques externes faisant appel à des techniques de piratage pour réaliser l'objectif d'accès aux systèmes et aux données. Cet objectif est facile à atteindre pour un administrateur, il lui suffit de se connecter à la base de données et de récupérer les informations. Cette action malveillante passera inaperçue, car elle fait partie du travail habituel de l'administrateur (ce type d'accès ne déclenchera aucune alerte). En réalité, dans une entreprise, l'accès aux différents serveurs et aux données sauvegardées fait partie des activités quotidiennes des employés et les évènements enregistrés sont trop volumineux pour pouvoir tous les analyser un à un afin de confirmer la menace.

La menace interne a toujours figuré parmi les risques à considérer par les entreprises, mais aujourd'hui elle devient fréquente, difficile à intercepter et cause des dommages plus importants que ceux résultant d'attaques externes. Selon le rapport [4] publié par Securonix, la menace interne a touché en 2020 plusieurs secteurs de l'industrie. Les domaines pharmaceutique et financier figurent parmi les plus affectés. Le vol d'informations est la catégorie la plus dominante. Cela s'explique par le fait que ces deux secteurs possèdent des données précieuses

qui peuvent être utilisées par les compétiteurs ou à des fins de fraude.
La protection des organisations contre les cybermenaces devient un véritable défi lorsque les menaces proviennent de l'entreprise, en particulier des utilisateurs autorisés. Il est difficile de déterminer si ces utilisateurs font simplement leur travail ou font réellement quelque chose de malveillant. En 2020, de grandes entreprises ont été victimes des actes relatifs à la menace interne, ci-dessous deux exemples.

Vol de données chez Shopify Deux membres de l'équipe de support de Shopify ont abusé de leurs droits d'accès pour obtenir les enregistrements sur les transactions des clients pour certains marchands inscrits sur la plateforme [5].

Distribution d'information à caractère financière sur Amazon Le directeur principal du service fiscal du géant Amazon avait divulgué des données financières confidentielles aux membres de sa famille afin qu'ils puissent avoir un avantage sur les marchés financiers [6].

Dans *2020 Cost of insider threats Global report* [7], l'institut Ponemon présente une étude sur le coût de la menace interne. Au total, 204 organisations de partout dans le monde ont fait partie de cette étude et ont rapporté des cas de menace interne. Le nombre d'incidents associés à ce type de risque a augmenté de 47%, passant de 3200 en 2018 à 4716 en 2020. Respectivement, les coûts relatifs ont aussi augmenté de 8,76 millions de dollars US à 11,45 millions de dollars US.
Certains produits comme les solutions de prévention ou de détection de perte de données *Data Loss Prevention* (DLP) ou les suites d'analyse du comportement des utilisateurs et des entités *User and Entity Behavior Analytics* (UEBA) permettent de surveiller et détecter des comportements suspects dans le contexte de la menace interne.
Les DLP appartiennent à la catégorie des solutions basées sur des règles et comme les systèmes de détection d'intrusion *Intrusion Detection System* (IDS), ils génèrent beaucoup de faux positifs. Les DLP peuvent être déployés pour surveiller les impressions, les courriels sortants, le transfert de données effectué vers un support amovible, etc. Leur but est de protéger les données privées pour qu'elles ne soient pas exportées à l'extérieur du périmètre de l'entreprise sans autorisations sur l'ensemble des vecteurs de communications.
Les canaux de communication sont largement utilisés dans une grande organisation. Par exemple, le nombre de courriels envoyés à des entités externes comme des partenaires ou des clients peut parfois atteindre plusieurs milliers par jour. Avec une seule règle rigide qui stipule la détection de fichiers sur les courriels sortants, plusieurs alertes peuvent être déclenchées qui au final s'ajouteront à l'ensemble des faux positifs. La frontière entre l'action malveillante et une action légitime est encore plus fine que dans le cas des attaques connues et gérer un tel taux de faux positif a un coût significatif.

Les UEBA, quant à eux, tentent de produire un score de risque associé au comportement d'un usager en particulier, d'une machine, ou de toute autre entité du réseau. Les analystes consultent les scores, puis analysent les activités enregistrées pour un compte. Toutefois, le score représente un ensemble d'évènements qui se sont produits, mais parfois son calcul reste propriétaire à la solution et abstrait pour l'expliquer et confirmer la menace. Certaines UEBA utilisent l'apprentissage machine pour définir un modèle sur une entité, puis observent le changement du comportement pour cette entité par rapport à ce modèle de référence construit. Cette information est retournée à l'opérateur avec quelques détails sur les évènements impliqués dans la variation du comportement. Malheureusement deux complications existent. La première concerne l'algorithme utilisé pour l'apprentissage qui peut produire des faux positifs. La seconde est relative au changement de comportement habituel qui demeure normal pour l'usager, lorsque par exemple de nouvelles tâches lui sont attribuées. Dans un environnement de production incluant plusieurs milliers d'individus, la surveillance de l'activité devient un défi.

Souvent les UEBA ne prennent pas en considération l'organisation de l'entreprise en termes de déploiement de l'infrastructure informatique ou en termes de la position ou les rôles des employés. Par exemple, un usager peut avoir plusieurs comptes pour des fins particulières. C'est le cas lorsque la politique de sécurité oblige les administrateurs à avoir un second compte à haut privilège pour gérer les systèmes. L'usager avec les deux comptes sera représenté par deux entités sur la solution UEBA avec les évènements relatifs à chaque compte. Séparé, le score de chaque compte ne reflète pas l'activité de l'usager, en plus l'analyste doit établir manuellement la relation entre ces deux comptes pour avoir une visibilité correcte sur le comportement de cet usager.

Le retour des analystes ou les résultats d'investigation des anciens incidents ne sont pas pris en considération par les UEBA. Un même individu peut paraître plusieurs fois avec un score de risque élevé alors qu'il réalise des actions légitimes, pendant que d'autres activités malveillantes passent inaperçues dans tout ce bruit d'erreurs et d'incompréhension.

Les entreprises se sont orientées vers les techniques de l'intelligence artificielle pour automatiser et réduire le temps du traitement dans l'intention de répondre à la forte volumétrie des données à analyser qui est devenu un facteur significatif. Toutefois, l'adoption de ces techniques doit être contrôlée et approuvée selon le domaine. Contrairement à l'impact d'une erreur lors de la recommandation d'un film sur Netflix, l'impact d'une erreur de recommandation dans un domaine médical est bien plus sérieux. La cybersécurité s'oriente plus vers la même tolérance d'erreur qu'un domaine comme celui du médical ou tout autre domaine dont la criticité est élevée. La figure 1.1 reprend un peu cette idée, nous avons sur le côté droit de la figure un niveau élevé de tolérance aux erreurs, on y retrouve notamment Netflix ou toute

autre application semblable. Alors que dans le domaine médical la tolérance est presque nulle, car les erreurs ont des impacts graves. La menace interne est une branche de la cybersécurité

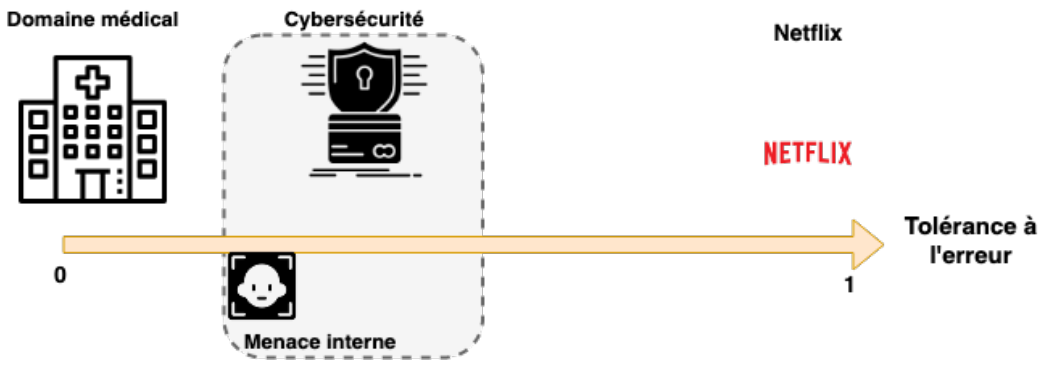

FIGURE 1.1 Tolérance à l'erreur dans l'utilisation de l'intelligence artificielle

qui implique des décisions importantes sur des individus au sein de l'organisation. Puisque ces décisions impactent des personnes, la tolérance à l'erreur est faible, voir même moins que pour la détection des programmes malveillants.

1.1 Problèmes liés à la menace interne

La détection d'individu comme intrus interne ou la protection en général contre la menace interne posent un véritable défi dans une grande entreprise. Les solutions actuelles telles les DLP et les UEBA répondent d'une manière partielle à ce problème. Les analystes observent beaucoup d'évènements et perdent énormément de temps à les analyser lorsque des alertes les concernant sont reçues, avant de se rendre compte que ce sont des faux positifs. Les anciennes investigations ou les anciens incidents ne sont pas considérés dans le processus de détection. Le tableau de bord des solutions de sécurité auquel se réfèrent les analystes ne permet pas d'introduire un retour pour pousser une correction sur le processus de détection et il assez difficile de déterminer la qualité d'une règle, car le contexte dans lequel l'activité d'origine a été effectuée inclut le comportement normal. Alors que les menaces réelles - vrai positifs continuent d'évoluer sur le réseau sans y être découvertes. La corrélation entre plusieurs sources d'évènement donne lieu à des erreurs qui sont introduites malheuresement dans les modèles d'apprentissage machine développés.

1.2 Objectifs de recherche

Les travaux de recherche consignés dans ce mémoire viennent répondre à la problématique de l'amélioration de la détection de la menace interne. nous proposons de mettre à jour le système en place pour qu'il intègre les investigations du passé et le contexte de travail des usagers. L'approche présentée doit toutefois s'adapter au processus en place et répondre aux besoins des équipes concernées. De cette manière, les résultats seront plus représentatifs des actions malveillantes.

1.2.1 Sous-objectifs

S.O 1 Comprendre le processus de gestion des incidents de menace interne d'une grande entreprise, et identifier les faiblesses rencontrées dans ce processus.

S.O 2 Introduire des modèles de l'intelligence artificielle et les évaluer pour la détection.

S.O 3 Proposer une ontologie pour éviter les incohérences lors du traitement des données, les impacts qui les accompagnent et améliorer le résultat des modèles que nous proposons dans le sous objectif 2.

1.3 Organisation du mémoire

Le deuxième chapitre discute sur l'évolution de la détection de la menace interne dans les différents travaux antérieurs, et les technologies disponibles actuellement pour la détection d'individus malveillants. Dans le troisième chapitre, on donne un aperçu de l'environnement en place dans lequel ce travail a été effectué et le processus de gestion des incidents en menace interne. Le quatrième chapitre décrit notre approche et comment nous avons proposé d'introduire les graphes de convolution pour répondre à notre objectif d'amélioration de la détection de la menace interne. L'entrainement et l'évaluation de notre approche ont été effectués sur des données réelles en production. Dans le cinquième chapitre, nous proposons de développer une ontologie et une base de connaissances sur un scénario de menace interne. L'ontologie permettra d'améliorer la qualité des données reçues des sources afin que les utilisateurs la consultent pour répondre à des requêtes pour l'investigation. Nous avons aussi opté pour un assistant virtuel afin de simplifier l'accès à la base de connaissance et pour avoir aussi une vision sur l'activité de l'investigation dans le but de construire un jeu de données pour les futurs modèles en apprentissage machine.
Finalement, une conclusion générale sur les résultats obtenus ainsi que des suggestions pour des perspectives.

CHAPITRE 2 REVUE DE LITTÉRATURE

2.1 Menace interne

La menace interne figure dans de multiples travaux antérieurs. Selon *Maybury et al.* [8] un acteur malveillant interne est une personne ayant des droits d'accès et une connaissance des systèmes, motivée à avoir un impact négatif sur la mission d'une entreprise par des actions qui compromettent la confidentialité, l'intégrité ou la disponibilité des informations. Alors que pour *Aleman-Meza et al.* [9] la menace interne fait référence aux actions malveillantes potentielles des employés au sein d'une organisation. Une autre version proposée par *Althebyan et Panda* [10] stipule que la personne suspecte connaît la structure du système d'information de l'organisation à laquelle elle est autorisée à accéder. En anglais les termes *Insider* et *Insider threat* sont définis séparément [11]. *Pfleeger et al. Sinclair et Smith Bishop et al.* [12–14] considèrent un *Insider* comme une personne avec des accès légitimes sur les ordinateurs, données ou les programmes que les personnes en dehors de l'entreprise ne possèdent pas. Cette personne peut aussi être représentée par une entité externe comme le cas des partenaires, des consultants ou les anciens employés. *Pfleeger et al. Thepharidon et al.* [12, 15] trouvent que *Insider threat* est l'ensemble des actions ou actes qui mettent à risque les données, processus ou les ressources d'une entreprise dont l'origine provient des usagers qui ont obtenu des droits d'accès à un système informatique (SI) et qui abusent de leurs privilèges, violant ainsi la politique de sécurité du SI.

La menace interne peut être classifiée selon aussi l'impact résultant des activités malveillantes effectuées par les individus. Quatre classes sont développées par *Al-Mhiqani et al.* [16] .

Vol de propriété intellectuelle Dans ce cas, l'employé malveillant vole les données sur la propriété intellectuelle puis les transfère à l'extérieur de l'organisation pour réaliser des profits. Les données relevant de la propriété intellectuelle diffèrent d'une entreprise à une autre et d'un domaine à un autre. Elles peuvent inclure du code source d'un logiciel, les informations sur les produits comme les formules ou la conception, les plans stratégiques, les soumissions d'appels d'offres ou encore les informations sur les clients. Le montant s'estime à 13,5 millions de dollars pour cet impact sur le jeu de données *Computer Emergency and Response Team* (CERT).

Fraude Elle survient beaucoup plus dans le domaine financier où l'individu tente de faire des profits au sein de l'entreprise.

Espionnage Un point souvent soulevé dans le domaine militaire, il existe aussi dans l'industrie. L'individu espionne l'activité dans l'entreprise et puis la relate à l'externe vers

des concurrents pour avoir des bénéfices.

Sabotage Un des impacts majeurs et dévastateurs pour une entreprise est le sabotage. Dans ce cas, l'individu tente de détruire les actifs physiques ou virtuels par son activité.

2.2 Intrus interne

Un moyen de faciliter la mitigation ou la détection d'un intrus interne est de le caractériser. *Al-Mhiqani et al.* [16] ont réalisé une organisation pour introduire de l'ordre dans les différentes appellations et distinctions sur les caractéristiques de la personne considérée ou l'activité effectuée. Comme pour un attaquant externe un individu interne a des motivations qui incitent son comportement, trois motivations ou facteurs ont été définis par *Cole et Ring* [17].

Facteur financier Il représente le principal facteur pour les incidents en menace interne, car la personne ayant accès aux différents secrets de l'entreprise peut les vendre à des prix importants aux concurrents par exemple.

Facteur idéologique Ce facteur existe lorsque l'individu n'approuve pas la politique de la compagnie dans laquelle il travaille, ou ses intérêts sont affectés par cette politique.

Facteur Personnel Des conflits existent dans une entreprise et les conflits peuvent évoluer lorsque des les informations personnelles sont utilisées. Cette situation nuit aux employés et à l'entreprise.

Trois types d'intrus internes sont cités dans la littérature [17–19]. Le plus discuté est celui de l'individu malveillant, tel que défini à la section 2.1, qui abuse des privilèges pour créer des préjudices à l'entreprise. Un autre type est celui de l'usurpateur qui vole l'identité d'un employé puis l'utilise afin d'accéder au réseau pour mener son attaque et atteindre les données ou les plateformes ciblées. Puis, les personnes qui effectuent des actions douteuses de manière non intentionnelle par méconnaissance de la politique en vigueur ou des mesures en place.

2.3 Jeux de données pour la menace interne

Un obstacle important pour réaliser des recherches en menace interne est d'avoir un accès à un jeu de données à analyser. Cependant, il existe certains jeux de données qui ont servi à proposer des techniques ou des approches de détection et de mitigation. L'ensemble des jeux de données présentés nous oriente vers les données à collecter afin de cerner l'activité d'un usager dans l'entreprise.

2.3.1 Jeux de données menace interne - CERT

Connu sous le nom CERT [20], ce jeu de données est une collection de plusieurs bases de données synthétiques générées par la division CERT de l'Université Carnegie Mellon et d'autres partenaires [21].
Elle inclut les journaux de l'authentification, l'historique de navigation Web, les courriels échangés, les journaux d'accès aux fichiers, l'utilisation d'équipements comme les dispositifs de stockage amovibles, les données *Lightweight Directory Access Protocol* (LDAP) [1] ainsi que des informations psychométriques sur les individus. Ce jeu de données est organisé comme suit sur le tableau 2.1

Tableau 2.1 Détails sur la base de données CERT

Web	ID, utilisateur, date, Ordinateur, URL, contenu de la page
courrier	ID, utilisateur, date, Ordinateur, from, to , CC, BCC, pièce jointe, taille, contenu du courriel
authentification	ID, utilisateur, date, Ordinateur, activité authentification ou fermeture de session
journaux fichier	ID, utilisateur, Ordinateur, nom de fichier, contenu du fichier
journaux équipement	ID, utilisateur, date, Ordinateur, attacher le périphérique/ détacher le périphérique

Ce jeu de données est assez diversifié et reprend des actifs réels qu'on retrouve le plus souvent dans les entreprises, mais il reste incomplet comparé aux actifs qu'un usager utilise présentement, comme les programmes lancés sur la machine. En plus, les actions qu'on observe pour un usager restent synthétiques et ne reflètent pas réellement l'activité d'un employé. Cependant, l'organisation et les sources utilisées ici nous donnent une base sur les sources de données qu'on doit ajouter afin de faire une analyse globale.

2.3.2 Jeu de données Schonlau

Cet ensemble de données a été créé par Mathias Schonlau, professeur à l'Université de Waterloo, pour son article sur la détection d'usurpateurs basée sur les commandes exécutées dans le cadre de la détection d'intrusion [23]. Ce jeu de données contient les commandes exécutées par 50 utilisateurs avec des rôles différents dans l'entreprise. Les données sont constituées de 50 fichiers correspondant chacun à un utilisateur. Chaque fichier contient 15 000 commandes.

1. LDAP (Lightweight Directory Access Protocol) est un protocole conçu pour communiquer avec un serveur d'annuaires afin de localiser des données sur des organisations, des individus et d'autres ressources telles que des fichiers et des appareils dans un réseau [22].

Les 5 000 premières commandes de chaque utilisateur contiennent seulement les commandes effectuées par lui et sont utilisées comme des données d'apprentissage. Les 10 000 commandes suivantes sont un mélange entre les commandes lancées par l'utilisateur lui-même et des commandes lancées par d'autres utilisateurs. Le but de la recherche effectuée sur le jeu de données est de réussir à identifier les commandes exécutées par les autres utilisateurs comme étant des anomalies.
Ce jeu de données regroupe seulement les commandes réalisées par les utilisateurs et ne donne pas une visibilité sur toutes les actions qu'un utilisateur a pu effectuer. Aussi, on suppose l'absence de commandes malveillantes sur les 5 000 commandes qui servent à créer un modèle sur le profil d'un utilisateur. Les utilisateurs sont susceptibles de recevoir de nouvelles tâches à réaliser et produisent ainsi de nouvelles commandes non-observées durant la phase d'apprentissage. Ce nouveau comportement va générer beaucoup d'alertes.

2.3.3 Jeu de données Are You You - RUU

Ce jeu de données a été créé à partir de l'activité récupérée sur 34 ordinateurs d'étudiants en informatique avec leurs consentements à l'Université de Columbia. La collection se faisait à travers un agent installé sur les machines et a permis d'avoir plus de 10Gb, en moyenne 7 jours de données par étudiant. Afin d'obtenir des artefacts qui représentent un usager interne malveillant, un exercice était conduit pendant 15 minutes sur des temps différents dans lequel 14 étudiants ont participé.
Cet exercice implique que les étudiants tentent de trouver des informations qui leur permettront d'avoir un gain financier sur un système de fichier rendu disponible seulement pendant la durée de l'exercice. Ce jeu de données inclut les évènements sur l'activité des registres pour les machines sous Windows, la création et la suspension de processus avec les détails sur les processus comme le nom, le chemin ..., les accès sur l'interface et le chargement des librairies dynamiques [24].
Bien que ce jeu reprend une activité réelle, il n'inclut que la phase d'accès aux fichiers disponibles pendant l'excercice qu'un acteur malveillant réalise. La suite des actions que les étudiants concernés accomplit avec les informations receuillies n'est prise en compte. On ignore aussi le fait qu'un utilisateur manipule les fichiers pour des fins de travail et on suppose qu'ils sont toujours utilisés dans une activité malveillante.

2.3.4 Jeu de données *The Wolf Of SUTD* - TWOS

L'ensemble de données TWOS *The Wolf Of SUTD*[2] a été collecté lors de la compétition de l'université de technologie et de design de Singapour en mars 2017 conçue afin d'obtenir des

2. STUD, Singapore University of Technology and Design

instances réalistes de menace interne malveillante. Le challenge impliquait la participation de six équipes composées de quatre étudiants qui se sont affrontées pendant cinq jours. Leurs activités ont été surveillées par plusieurs agents de collecte. Les données enregistrées correspondent aux activités relatives aux mouvements de la souris de l'ordinateur, les touches du clavier, les processus, les évènements du système de fichiers, le trafic réseau, les courriels et les connexions et déconnexions. Au total, 320 heures de données qui comprenant 18 heures de données d'usurpateur avec deux instances de données d'acteur malveillant interne. Ce jeu de données inclut aussi des réponses à des questionnaires psychologiques pour cerner la personnalité des participants. [25]
Lors de la compétition, les équipes tentent de gagner, ce qui genère beaucoup d'évènements liés à une activité qui est loin des actions prises par un employé pour réaliser son travail.
L'ensemble des jeux de données présenté inclue des comportements qui ne reflètent pas l'activité d'un employé dans une entreprise. De plus, la majorité des attributs concerne des opérations sur la machine locale et n'inclut pas nécessairement des actions sur des serveurs distants ou des services accessbiles telles les bases de données où des données importantes résident. Aucun contexte n'est disponible soit du côté de l'actif auquel l'utilisateur a accédé soit du côté des privilèges de l'utilisateur. Dans ce travail, en plus d'utiliser des données d'un environnement réel, nous explorerons le contexte de l'employé dans l'organisation.

2.4 Techniques et méthodes en menace interne

Liu et al. [18] se sont basés sur la chaine d'attaque du framework MITRE ATTACK [3] pour modéliser et catégoriser les techniques et méthodes utilisées par les intrus internes durant leurs activités malveillantes. Mais selon l'ancien directeur du *Federal Bureau of Investigation* (FBI) dans sa présentation à l'édition 2013 de la conférence *BlackHat* [4], les techniques évoquées dans le framework MITRE ne s'appliquent pas correctement dans un contexte de menace interne. En menace interne, des tactiques différentes sont suivies et un nouveau framework doit être établi pour rassembler les pratiques utilisées. En effet, les administrateurs internes n'ont pas besoin d'utiliser des exploits avec des charges utiles afin de s'introduire sur les machines auxquelles ils ont déjà accès. La figure 2.1 reprend les étapes d'un individu lors de ses actions malveillantes dans le contexte de menace interne tel que proposé lors de la conférence.

3. https ://attack.mitre.org, Le framework *MITRE ATTACK* est une base de connaissances mondialement accessible sur les tactiques et techniques de l'adversaire, basée sur des observations du monde réel. Elle est utilisée comme base pour le développement de modèles et de méthodologies de menaces spécifiques dans le secteur privé, au sein du gouvernement et dans la communauté des produits et services de cybersécurité [26].

4. Les Conférences *Black Hat* sont un événement unique qui rassemble officiellement des experts fournissant des points de vue nouveaux et exclusifs sur la sécurité de l'information.

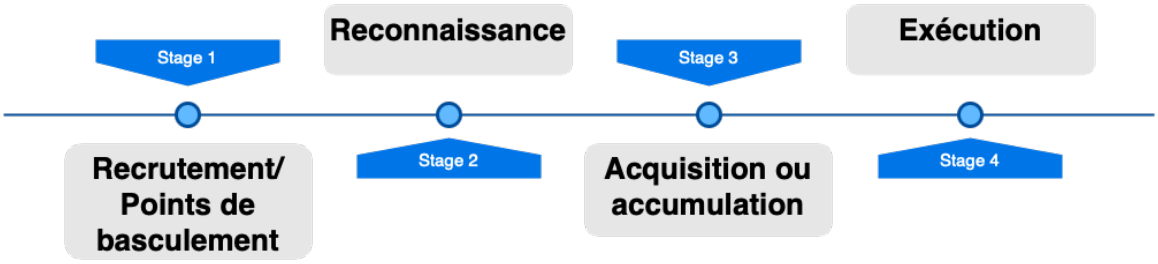

FIGURE 2.1 Chaine d'attaque pour la menace interne

2.4.1 Recrutement ou points de basculement

Les attaquants peuvent être malveillants dès le début de leur contrat ou suite à un évènement majeur. Cette étape est d'un domaine d'intérêt souvent négligé ou difficile d'accès, car il inclut la vie privée de l'individu qui est protégée par des lois. Mais l'identification des évènements, tels qu'un divorce ou des problèmes de dépendance aux jeux, peut permettre d'identifier les membres du personnel qui représentent un haut risque. Ces évènements fournissent également une plus grande assurance quant à l'authenticité d'autres alertes possibles produites plus loin dans la chaine d'attaque. Par exemple, un employé qui passe par des moments difficiles financièrement peut faire des manipulations pour tenter d'avoir des bénéfices avec les actifs auxquels il a accès au niveau de l'entreprise.

2.4.2 Reconnaissance

Cette étape de recherche et de reconnaissance peut fournir des avertissements précoces uniques, ainsi qu'une assurance supplémentaire lorsque ces alertes sont liées à d'autres alertes. Dans cette étape l'acteur malveillant recherche l'emplacement où sont stockées les données voulues et tente d'y accéder.

2.4.3 Acquisition ou accumulation

Cette étape regroupe l'ensemble des techniques manuelles et automatiques auxquelles l'individu a eu recours pour récupérer les données identifiées durant la phase précédente et les sauvegarder dans un seul endroit pour les préparer à l'action suivante.

2.4.4 Exécution

Dans cette dernière étape, l'individu met son plan en action. Par exemple, transmettre les données à des endroits auxquels l'entreprise n'a pas un contrôle.

2.5 Évolution de la recherche sur la menace interne

Un algorithme de prédiction pour prévenir la menace interne a été proposé par *Althebyan et Panda* [10]. Ce modèle se base sur un graphe de connaissance qui représente l'individu et les d'accès qu'il fait sur des documents ayant des priorités. Le graphe évolue lorsqu'un utilisateur accède à un objet, cette évolution se matérialise par une unité de connaissance qui s'attache au graphe de l'individu. Chaque unité est liée à un objet - un document dans ce cas -. Un autre aspect mentionné est la dépendance entre les documents qui vont servir par la suite à créer un groupe de documents que l'algorithme utilise pour évaluer la tentative d'accès et le risque. Un seuil est défini pour chaque risque. S'il est atteint ou dépassé suite à une tentative de lecture par exemple, tous les documents du même groupe seront interdits au demandeur. Le scénario sur lequel les auteurs ont déroulé leur modèle assume l'existence d'une dépendance entre les documents qui n'est pas toujours vraie dans le monde réel. L'exemple de scénario que les auteurs ont déroulé est une soumission pour un appel d'offre. Le projet à soumettre a été divisé en plusieurs documents qui incluent les coûts des matériaux, l'estimation en temps, et d'autres informations relatives aux besoins en termes d'employés et de ressources budgétaires. Dans cet exemple, chaque document dépend d'un autre. En réalité, les documents n'ont pas tous une relation La technologie est un point important dans la détection des activités d'un acteur malveillant. *Wang et al.* [27] suggèrent d'avoir une approche multidisciplinaire et ont développé Insider Threat Security and Defense Architecture *(ITSDA)* une architecture plus globale pour se défendre des risques engendrés par la menace interne. Sept aspects sont discutés.

1. La haut direction de l'entreprise pour accepter et intégrer l'architecture de sécurité dans les plans stratégiques et d'affaires.
2. Les politiques à renforcer pour se défendre contre la menace interne.
3. La sensibilisation et de la formation du personnel.
4. L'entreprise doit aussi inclure la menace interne à sa stratégie de management de risque et s'assurer de faire des audits pour confirmer l'implémentation des procédures et de leur maintiens.
5. La surveillance de la conformité pour s'assurer du succès de l'implémentation des stratégies.
6. La compagnie doit s'acquérir des technologies qui permettent de se défendre et mettre en place un plan de réponse à ce type d'incident.
7. Un plan de réponse à incident approprié pour traiter les incidents de menace interne.

Afin de collecter et de surveiller les activités des usagers, *Ali et al.* [28] présentent un modèle développé autour d'un agent qui est déployé sur les machines des clients. Cet agent envoie

l'activité observée sur le profil de l'utilisateur vers un serveur où sont sauvegardés tous les profils. Les analystes consultent et analysent manuellement la présence d'une violation de la politique par un usager selon des règles d'accès qui sont implémentées et définies pour des niveaux différents de profils d'utilisateurs. L'activité locale sur la machine comprend la navigation sur le Web, les opérations sur les partages, les processus et les applications utilisés par l'usager, les impressions et l'utilisation de lecteurs externes comme les clés USB. Cet agent ressemble beaucoup aux produits *Endpoint Detection and Response* (EDR) qui sont utilisés actuellement pour la surveillance en continu des évènements enregistrés sur une machine.
La surveillance de nos jours ne peut pas se limiter à l'ordinateur utilisé quotidiennement, car l'usager manipule d'autres services disponibles dans l'entreprise, tels que les services infonuagiques dont les évènements ne sont pas enregistrés sur la machine locale, mais sur le service distant. Pour s'adapter à l'évolution et la nouvelle vision du périmètre de l'entreprise, les mesures de protection doivent être adaptées. La solution apportée en [29] par *Yassen et Panda* révèle trois modèles pour la surveillance des bases de données relationnelles sur le nuage. Les données dans un environnement infonuagique sont dupliquées sur plusieurs zones pour des fins de haute disponibilité. Surveiller l'activité d'accès nécessite de prendre en considération cette duplication distribuée, car l'utilisateur peut extraire des données à partir de plusieurs endroits. Tout comme un modèle poste-à-poste où les zones sont liées une à une, les accès sont enregistrés sur une zone puis diffuser sur l'ensemble des zones de disponibilité. Cette approche possède un point négatif, car la base de connaissances sur l'ensemble des zones doit toujours être synchronisée, ce qui peut induire une latence dans l'exécution des transactions.
Le deuxième modèle utilise un coordinateur central qui reçoit la requête de l'usager, puis la transfère à l'instance. Une mise à jour est envoyée au coordinateur pour lui permettre de mettre à jour sa base de connaissances afin de vérifier les accès suivants. Le problème avec cette approche est qu'on aura un point de défaillance, car si le coordinateur tombe en panne, tout le système ne sera pas accessible. Une amélioration serait de mettre un coordinateur secondaire qui prend en charge l'activité à la place du premier avec un basculement automatique. Le dernier modèle se base sur la position géographique de l'individu, car ses accès se font généralement depuis la même zone. Cette solution vient consolider les limites de l'approche de poste-à-poste avec moins d'information à synchroniser entre toutes les zones.
Claycomb et al. ont fait une étude [30] sur les évènements qui précèdent le sabotage dans les systèmes informatiques dans le cadre de la menace interne. Le but de cette recherche est d'identifier les signatures ou l'ensemble des actions faites par l'opérateur malveillant avant son acte de sabotage, afin de permettre aux experts en sécurité de le détecter à l'avance et prévenir tout dommage. Avec la complexité des interactions dans une entreprise, les auteurs

ont opté pour une approche sociotechnique qui combine les actions techniques qui découlent de l'activité sur les différents systèmes et le comportement de l'individu dans l'organisation. À partir de 49 cas, ils ont relevé 449 évènements en suivant la méthode suivante. La première étape est de collecter les actions pour chaque cas, puis de les mettre sous format chronologique afin d'identifier les indicateurs clés de l'attaque. Ceci permettait aussi de comparer les résultats avec les comportements de la population non malveillante. Ensuite les suites temporelles résultantes sont analysées pour vérifier et confirmer les évènements et éviter toute incohérence. La troisième étape est de définir les évènements qui pouvaient être détectés avec la technologie courante, puis de les catégoriser. La dernière étape est de mettre des hypothèses sur l'occurence de l'action malveillante. Les résultats montrent qu'avant l'attaque, pour la majorité des cas étudiés, au minimum un évènement était enregistré et que les actions de comportements précédaient les actions techniques.

2.5.1 Jointures des données de sources hétérogènes

Un usager réalise des actions sur de multiples actifs de l'entreprise. Réunir tous les artefacts nécessite de recourir à l'ensemble des journaux enregistrés pour cet utilisateur sur des systèmes souvent hétérogènes. Une difficulté apparait lorsqu'on veut manipuler les données issues de ces systèmes pour construire un seul espace auquel les experts du domaine accèdent, soit pour faire des requêtes directes ou construire un modèle d'apprentissage machine.
Lors de la conférence des graphes de connaissance *Knowledge Graph Conference 2019* [31], les conférenciers ont discuté des problèmes rencontrés dans la vie réelle lors de la création d'un graphe de connaissance à partir des différentes bases de données et suggèrent d'implémenter une autre couche entre la collecte des données et le traitement. Cette étape inclut des ingénieurs en ontologie munis de méthodes pour construire l'espace de données global et s'assurer de la cohérence des données qui s'y trouvent. *Squeda* [32] propose de faire une projection des données depuis les bases de données vers les ontologies pour fixer l'écart de conceptualisation tel que défini dans la figure 2.2 et éviter d'avoir un impact négatif sur les résultats attendus par les experts du domaine. Une ontologie a été proposée en [33] par le CERT pour introduire et modéliser les concepts qu'on retrouve dans le domaine de la menace interne. Elle détaille les indicateurs d'activité malveillante.

La figure 2.2 reprend le schéma dans lequel les données transitent depuis les sources de données qui peuvent être des serveurs d'applications, des serveurs de base de données ou des solutions de sécurité vers la dernière étape des rapports.
Au moment de faire la liaison entre les données pour rassembler les activités enregistrées d'un usager, des erreurs peuvent se produire et sur une énorme quantité d'information. Ces

FIGURE 2.2 Écart de conceptualisation

erreurs peuvent passer inaperçues, car ce sont des exceptions rares dont le ratio est très petit comparé à l'ensemble des données. De plus, il sera plus difficile de les identifier plus tard dans le traitement ou même impossible, car les données d'origine peuvent être perdues si elles ne sont pas sauvegardées. En effet, les serveurs sont configurés pour garder les données pour une durée de rétention définie ou une taille de données particulière. Au-delà de ces deux paramètres, la rotation des journaux se déclenche et les nouveaux enregistrements écrasent les plus anciens.

2.5.2 Détection de menaces basée sur les règles

Les administrateurs peuvent configurer des règles sur les solutions de sécurité pour alerter les analystes sur la possibilité d'une mauvaise pratique qui indique une menace interne. Ces alertes sont mises en oeuvre par exemple sur les DLP au niveau des courriels sortants pour signaler la présence d'une chaîne de caractère spécifique telle un numéro de carte de crédit. Un tel scénario est devenu moins apprécié de nos jours, car il sera difficile de voir toutes les alertes en une période limitée du côté des opérations. Une grande organisation se voit communiquer avec l'externe très souvent ce qui implique un nombre important de communications à surveiller dans le contexte de prévention contre la perte ou le vol d'information. En plus, de la majorité de ces communications correspondent à un comportement légitime et seront considérées comme des faux positifs.

2.5.3 Détection de menaces basée sur les réseaux bayésiens

Ambre et Shekokar [34] proposent une architecture complète pour l'analyse de journaux et la corrélation entre les évènements. Cette architecture possède quatre modules. Le premier est utilisé pour collecter les fichiers depuis les sources d'évènements à travers des clients

installés sur les machines vers un serveur central pour que le deuxième module les analyse pour identifier les champs de chaque entrée. Ensuite le troisième module définit des relations de corrélation entre des actions qui peuvent se présenter comme un seul évènement sur une seule machine, plusieurs évènements sur une seule machine, un seul évènement sur plusieurs machines ou le cas échéant plusieurs évènements sur plusieurs machines. La corrélation permet d'avoir un niveau de connaissance supérieur et ainsi définir de nouveaux types de comportements et améliorer ainsi la détection.
Le quatrième module introduit la notion d'algorithmes bayésiens, dans lequel une probabilité est associée à un évènement puis une probabilité conditionnelle est calculée pour voir le ratio de détection ou le ratio de faux positifs avec la condition de présence de l'évènement ou sa présence sur un ensemble d'enregistrements. Un modèle en réseau bayésien a été proposé [35] pour prédire la menace interne dans lequel *Axelrad et al.* se sont basés sur des variables psychologiques pour produire des prédictions sur le comportement. L'utilisation de la probabilité conditionnelle a des avantages de simplicité et de probabilité. Elle est assez extensible pour inclure aussi le retour des analystes sur les évènements.
Le produit UEBA de la société Exabeam utilise des ajustements bayésiens pour affiner la valeur d'une anomalie observée depuis une source d'évènement comme l'atteinte d'un seuil avant de calculer le score final pour une session d'utilisateur (voir figure 2.3). Des valeurs sont

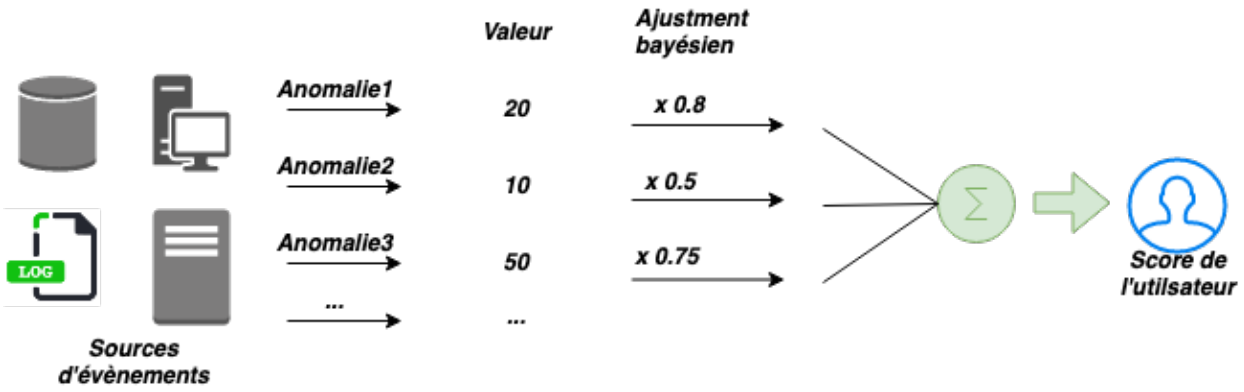

FIGURE 2.3 Ajustement des scores pour le produit UEBA de Exabeam [1]

définies pour chaque anomalie, si l'anomalie est déclenchée pour un usager la valeur associée sera multipliée par l'ajustement avant de faire l'agrégation globale.
Dans la réalité, les experts peuvent être en désaccord [36] sur la probabilité d'une action. En effet, le même évènement enregistré pour deux individus peut être différent selon le contexte. Dans une entreprise, un employé dans le département de marketing a tendance à communiquer souvent avec l'extérieur en envoyant des courriels alors qu'un ingénieur de données

ne le fait pas. Dans certains évènements, il faut tenir compte du contexte temporel, par exemple durant la période de déclaration des revenus, la majorité des usagers transmettent leurs documents fiscaux à leurs adresses ou espaces personnelles.

2.5.4 Détection de menaces basée sur K plus proches voisins

Le problème de menace interne a été considéré comme un problème de classification dans lequel l'algorithme du k plus proche voisin *K-Nearest Neighboors* (KNN) est utilisé pour déterminer la classe d'un élément. Dans leur approche, *Santos et al.* [37] sur le jeu de données APEX '07, les chercheurs ont considéré des séquences de cinq actions pour chaque analyste, si une des actions est identifiée comme malveillante la séquence le sera aussi. Le jeu de données TWOS a été utilisé dans [38] pour construire un modèle en KNN, des étapes de traitement de langage naturel ont été adoptées pour transformer le texte à l'intérieur des courriels en des vecteurs d'attributs.

2.5.5 Détection de menaces basée sur Machine à vecteurs de support - SVM

Les machines à vecteur de support SVM sont une méthode d'apprentissage machine supervisée basée sur la théorie d'apprentissage statistique [39]. Cette méthode a été utilisée pour la détection d'intrusion [40] en raison de ses performances. Elle minimise l'erreur de classification et maximise la marge géométrique entre deux classes afin d'avoir la meilleure séparation possible pour les données linéairement séparables ou pour celles non linéairement séparables [41]. L'une des difficultés à pouvoir appliquer cette méthode est le déséquilibre entre les classes qui résulte en de mauvaises performances. En effet, d'un côté nous avons la majorité des données qui représente l'activité normale et de l'autre une minorité pour l'activité de menace interne [42]. Afin de répondre à la situation de déséquilibre, SVM à une classe a été proposé dans [43] pour la détection des données aberrantes qui sont très différentes des données sur lequel le modèle a été entraîné. Cette technique tente de maximiser la distance entre un point et la majorité des données, tout élément qui se situe en dessous de cette distance sera considéré comme intrus. À la différence du SVM ordinaire le SVM à une classe s'entraîne seulement avec les éléments considérés normaux, considéré comme formant une seule classe.

Cette technique est applicable pour des éléments fixes, cependant les données en menace interne ne se basent pas sur un seul évènement ou une suite d'évènements fixes. Il se peut que pour un individu, trois actions soient suffisantes pour le considérer dans une investigation tandis que pour un autre usager, plusieurs sont nécessaires sur une longe période de temps. L'entraînement avec une seule classe suppose que toutes les données d'entraînement sont normales et n'incluent pas d'actions malveillantes.

2.5.6 Détection de menaces basée sur les chaînes de Markov cachées

Munis du jeu de données du CERT, *Tabish et al.* [44] organisent les données en séquence d'actions effectuées par l'usager durant une période d'une semaine afin de construire des suites temporelles.
Les séquences concernant les activités sont divisées comme suit.

- Connexion durant les jours de la semaine dans la période de travail habituel de 8am à 5pm
- Connexion en dehors des heures de travail pendant la semaine
- Connexion durant les week-ends
- Copie d'un fichier exécutable vers un lecteur amovible
- Copie d'un fichier zip, jpg, doc, pdf, txt
- Envoi d'un courriel en interne vers des adresses courriel de la société
- Envoi d'un courriel à l'externe de l'entreprise
- Visite de site Web
- Insertion de lecteur amovible durant les horaires de travail de la semaine
- Insertion de lecteur amovible en dehors des heures de travail de la semaine
- Insertion de lecteur amovible pendant le week-end

Avec cette organisation des activités, les chercheurs ont construit des séquences qui représentent les états et les transitions d'un état vers un autre de l'usager sur une période de temps. Le modèle utilise les chaînes cachées de Markov pour apprendre les séquences sur une période d'entraînement de quatre semaines puis prédire la probabilité des séquences sur les semaines suivantes. Leur modèle a réussi à avoir un bon ratio de détection comparé au ratio de faux positifs.
Un résultat très important de ces travaux de recherche est le fait que lors de la détection de l'intrus interne, le modèle permet de retrouver aisément la transition d'état et ainsi retrouver la séquence de l'activité. Ceci est très important pour les analystes car cela leur permet d'appréhender la différence entre l'activité normale et l'anomalie. Un des soucis avec cette approche est que l'on considère que pendant la période d'entraînement l'activité est normale, mais rien ne peut le confirmer. Un autre inconvénient est le fait que l'attaque se manifeste comme anomalie alors qu'elle peut très bien être superposée sur une activité ordinaire. Selon *Al-Mhiqani et al.* [16] le nombre d'états considérés a un coût sur le modèle.

2.5.7 Détection de menaces basée sur les graphes de convolution

Souvent les usagers qui font partie d'une même équipe ont tendance à faire les mêmes opérations. Dans un contexte de détection en menace interne, ceci peut servir à éviter de voir certaines opérations d'une même équipe comme activité douteuse sachant qu'elles ont été déjà explorées dans le passé pour un usager et ont été considérées comme légitimes. Par exemple, dans une équipe composée d'agents hypothécaires on retrouve souvent des courriels sortants dans lesquels ces agents communiquent avec leurs clients. Dans un cas où cette relation d'appartenance à l'équipe d'hypothécaire n'est pas prise en compte, les analystes auront à faire des investigations qui amènent à de faux positifs et la conclusion que ces actions font partie des tâches quotidiennes pour les personnes de cette équipe. Notamment, on peut profiter de cette notion de pairs pour éviter la charge d'investigation et réduire le taux de faux positif. D'un autre côté, l'appartenance à une équipe permet d'améliorer la confiance dans la détection et nous ouvre une voie vers la prévention. Par exemple, si on considère les privilèges d'accès des administrateurs, un évènement enregistré pour un administrateur nous permet de savoir que cette action est très douteuse sachant qu'il possède un privilège particulier.

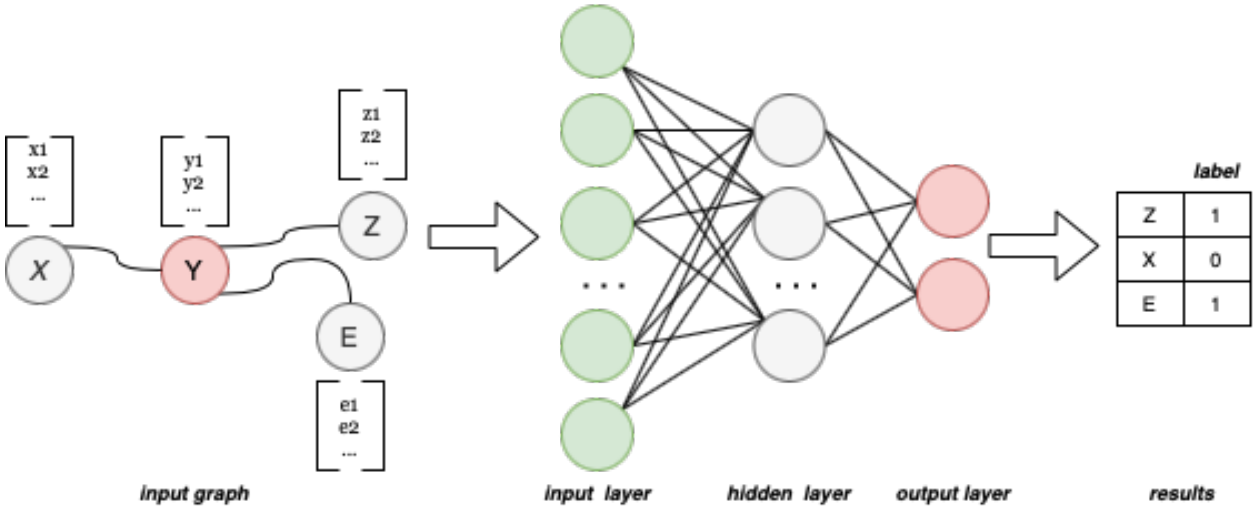

FIGURE 2.4 Réseau de graphe de convolution dans un contexte d'apprentissage semi-supervisé

L'utilisation des graphes de convolution était proposée en [45] pour tirer avantage de la notion de relation. Dans [45], le jeu de données CERT a été transformé en un graphe dont les noeuds sont les profils des usagers et les évènements enregistrés pour chaque profil sont mis sur un vecteur qui accompagne le noeud. Le jeu de données CERT n'inclut pas le concept de l'organisation qu'on retrouve dans les entreprises.

Afin d'introduire les relations entre les noeuds, les chercheurs ont d'abord proposé de lier les usagers qui s'envoient des courriels, mais cette idée ne permettait pas d'avoir des relations entre l'ensemble des usagers. En effet, plusieurs noeuds se retrouvaient isolés du reste du graphe et cela a un impact sur l'apprentissage, car ce dernier se base sur l'échange entre les noeuds liés. Afin de remédier à cette situation, ils ont proposé une fonction qui calcule les similarités entre les utilisateurs et ont construit une matrice d'adjacence qui représente les relations entre les noeuds. Ce module se base sur la relation directe et la pertinence dans les comportements, mais des détails manquent au niveau de leur fonction de calcul de similarité. La figure 2.4 reprend l'idée de cet article : en entrée nous avons le graphe dont les noeuds sont les utilisateurs, les vecteurs associés à chaque noeud sont les évènements enregistrés pour l'usager, les relations entre les usagers.
Avant de passer vers le réseau de neurones, une convolution est réalisée. Cette convolution s'exprime par une fonction qui prend le vecteur de chaque noeud puis réalise une agrégation avec les vecteurs des noeuds voisins. La convolution opère avant chaque couche de traitement sur le réseau de neurones. Par exemple si on soumet seulement deux couches au total, le voisin du voisin de chaque noeud est exploré. Dans une relation de gestionnaire ou supérieur direct, on peut atteindre les usagers qui ont le même gestionnaire.

2.6 Ontologie

En philosophie, l'ontologie est la théorie de la nature de l'être, elle a été introduite dans le but de faire une description détaillée du monde réel qui entoure la civilisation et pour résoudre le problème de sémantique entre les entités hétérogènes. L'ontologie a fait et fait l'objet d'utilisation dans la science de l'informatique, mais aussi dans d'autres domaines comme la biologie. Sa définition évolue depuis des années pour inclure de nouvelles perspectives. *"An ontology is an explicit, formal specification of a shared conceptualization of a domain of interest"*. La conceptualisation fait référence à un modèle abstrait d'un phénomène dans le monde en identifiant le concept pertinent de ce phénomène. Le mot explicite signifie que les types de concepts utilisés et les contraintes sur leur utilisation sont explicitement définies. Formelle signifie que les machines sont en mesure d'interpréter les concepts indiqués. Partagé indique que l'ontologie n'est pas spécifique à un individu, mais partage et accepte par un groupe d'individus. Le domaine d'intérêt spécifie qu'on s'intéresse à la modélisation d'une partie pertinente du monde [46]. Une ontologie définit un vocabulaire commun pour partager l'information d'un domaine d'intérêt [47] entre les machines et les usagers.
L'ontologie est composée de quatre composants principaux, **Concept**, **instance**, **relation** et **Axiome**

Concept ou Classe est un élément fondamental du domaine d'intérêt, en orienté objet cet

Tableau 2.2 Synthèse comparative des approches

Technique	Jeu de données	Forces	Faiblesses
Réseaux bayésiens [35]	Privé	L'utilisation de la probabilité conditionnelle a des avantages de simplicité et de probabilité. Elle est assez extensible pour inclure aussi le retour des analystes sur les évènements.	L'action dépend fortement du travail de l'employé, et si cette information n'est pas incluse la probablité s'applique à tous les employés.
K plus proches voisins [38]	TWOS	Cette méthode est facile à comprendre et à appliquer, elle peut aussi inclure le groupe pour voir la distance d'un individu de son équipe.	Assez difficile de définir dès le début le meilleur nombre de cluster à choisir. Les activité malveillante peuvent se cacher facilement dans un comportement normal.
SVM [42]	Privé	Offre un bon équilibre entre la qualité et l'efficacité.	L'activité de l'intrus interne peut se superposer sur une activité normale. De plus, les activité malveillante ne sont pas nombreuses.
Chaînes de Markov cachées [44]	CERT	Elle permet de trouver la séquence de l'activité qui a été identifiée comme malveillante.	Le nombre d'états considérés a un coût sur le modèle.
GCN [45]	CERT	Inclut le contexte de l'utiliateur dans le modèle.	Les relations choisies dans cet article ne dépendent pas de relation dans une entreprise, dans ce travail nous présentrons une limite de l'utilisation du GCN.

élément peut être considéré comme une classe d'objet (une personne ou une voiture par exemple).

Instance est un individu spécifique pour un concept, par exemple Canada est une instance du concept pays.

Relation utilisée pour démontrer la liaison qui existe entre deux concepts. Le premier concept pour la relation est considéré comme domaine et le second concept comme la portée.

Axiome permet d'imposer certaines contraintes sur les concepts et les relations pour assurer la cohérence de l'ontologie.

Une ontologie O est six-uplet $< C, R, H^C, H^R, I, A >$ [48] [49] où :

— C : désigne l'ensemble des concepts

— R : désigne l'ensemble des relations. $R_i \in R, et R_i \rightarrow C \times C$.

— H^C :désigne une relation d'ordre partiel sur C appelée hiérarchique ou taxonomie des concepts. Elle associe à chaque concept ses sous-concepts. $H^C \subseteq C \times C$, ou H^C(C1,C2) signifie que C1 est un sous-concept de C2

— H^R : désigne une relation d'ordre partiel sur R appelée hiérarchique ou taxonomie des relations. Elle associe à chaque relation ses sous-relations. $H^R \subseteq R \times R$, ou H^R(R1,R2) signifie que R1 est une sous relation de R2

— I : désigne l'ensemble des instances de C ou de R.

— A : désigne un ensemble d'axiomes.

2.6.1 Web sémantique

Les ressources sur le Web sont disponibles sous un format lisible pour l'homme. L'objectif du Web sémantique ou Web de données est de les rendre interprétables par les machines. Il permet la construction et la concrétisation des ontologies à travers les standards et les technologies. Formulée par *Berners et al.* [50], *"The Semeantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation"*. La vision est de permettre à l'utilisateur de rechercher des informations en fonction du sens plutôt que de la syntaxe. Pour les machines, il s'agit de traiter les ressources de données de façon automatique et de les relier ensemble.

2.6.2 RDF

Ressource Description Framework *(RDF)* est une structure de modèle de données pour décrire la sémantique d'une ressource Web. Il a été initié par Ramanathan V.Ghua [5] lorsqu'il travaillait chez Apple. En 1999, la W3C [6] a adopté la première version qui a été mise à jour en 2004. Le modèle RDF est composé d'un ensemble de triplets constituant un graphe orienté dont les noeuds représentent les différents **sujets** considérés comme ressource, des **objets** qui peuvent se présenter comme ressource ou valeur et les arcs représentent les **prédicats** ou les relations qui lient les sujets et les objets [46].
Ressource Description Framework Schema (RDFS) est une extension sémantique du vocabulaire RDF. *"Il fournit des mécanismes pour décrire les groupes de ressources associées et les relations entre ces ressources."* [52]. Cependant RDFS reste simple d'un point de vue logique, car on ne peut ajouter d'autres informations sur les classes ou les instances [53].

5. **Ramanathan V. Guha** est le créateur de RSS, RDF et Schema.org [51]

6. World Wide Web Constorium (W3C) est une communauté internationale qui développe des standards libres pour assurer l'évolution du Web, https ://www.w3.org

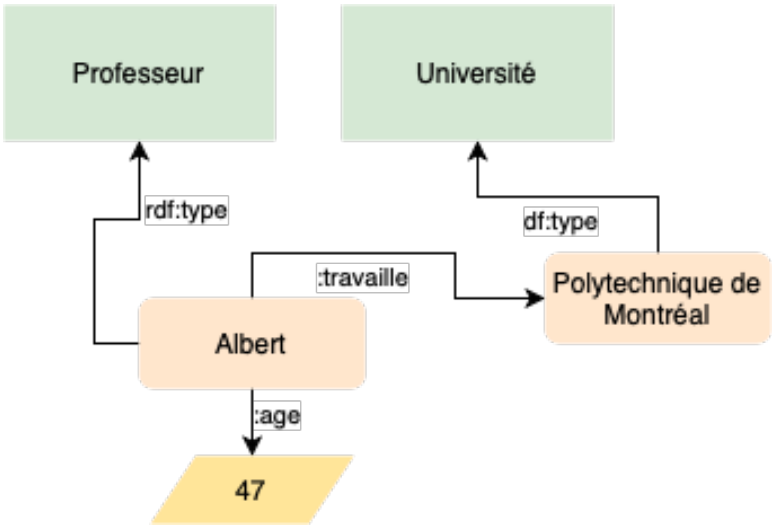

FIGURE 2.5 Exemple RDF

2.6.3 OWL

Ontology Web Language *(OWL)* a été développé pour ajouter une couche sémantique formelle aux données décrites en RDF/RDFS. Initialement créé grâce à la fusion de deux projets DARPA Agent Markup Language *(DAML)* proposés par American Defense Advanced Research projects Agency *(DARPA)* et le projet européen Ontology Inference Layer *(OIL)*, W3C l'a adopté en 2004. OWL fournit un vocabulaire additionnel, pour définir les relations d'intersection ou d'union sur les classes. OWL ajoute aussi la cardinalité. Du côté des propriétés ou des relations, OWL ajoute les caractéristiques de symétrie, transitivité et d'inverse. Tous ces ajouts permettent d'accroître le niveau de sémantique autour des données.

2.6.4 SWRL

Le langage OWL n'est pas capable d'exprimer toutes les relations, et son expressivité peut être étendue en ajoutant des règles à l'ontologie. *Semantic Web Rule Language* (SWRL) permet de combiner l'ontologie et des règles de logique. SWRL permet d'écrire des règles qui peuvent être exprimées par les concepts de OWL pour fournir des capacités de raisonnement. Les règles de SWRL sont des règles d'implication, elles sont sous le format :

$$Antécédents \implies Conséquences$$

Cette syntaxe implique que les conséquences existent si les antécédents sont satisfaits. Les expressions OWL peuvent apparaître à la fois dans les antécédents et dans les conséquences [54].

2.7 Ontologie en cyber sécurité

L'environnement d'une entreprise représente le monde réel dans lequel interagissent plusieurs sujets différents selon des relations particulières. Les mettre dans un modèle en base de données se solde par des limites de représentation de la réalité. Ainsi le choix des ontologies est mieux adapté pour décrire les concepts rencontrés.
Le choix des ontologies n'est pas une nouveauté en sécurité, et a évolué de plusieurs manières pour satisfaire les besoins ou les objectifs. La première recherche visant à appliquer des ontologies en sécurité revient à *Raskin et al.* [55], dans laquelle les chercheurs expriment que la connaissance sur la classification des menaces et les techniques de défense nécessitent d'être formalisées. Certains avantages ont été discutés sur ce choix des ontologies [55]. Le niveau de détail qu'on peut atteindre et la facilité d'expansion en intégrant de nouvelle connaissance.
Les premières ontologies étaient plus une description ou classification de la taxonomie du domaine lié à la sécurité. *Booth et Turner* [56] ont proposé un framework pour décrire les vulnérabilités sous forme d'ontologie (VDO) dans le but de partager la connaissance entre les professionnels en sécurité.
Li et al. [57] présentent Extended Semantic Web Rule Language *(XSWRL)* comme extension à *SWRL* pour représenter le modèle de corrélation d'alerte d'intrusion proposé aussi dans [57]. Cette extension vient définir plus de règles en XSRWL dans le but d'exprimer le scénario d'une attaque. Un système de détection d'intrusion pour détecter et valider les attaques Web avec l'objectif de réduire le taux de faux positif et augmenter la précision [58]. *Granadillo et al.* [59] ont proposé de modéliser le SIEM en une Ontologie. Deux classes principales ont été proposées. Une pour exprimer lesinformations sur le réseau comme les machines, utilisateurs, vulnérabilités, etc. L'autre reflète les opérations que le SIEM doit faire pour gérer les évènements, assurer une corrélation et appliquer une contremesure adéquate pour stopper la menace. Les chercheurs dans [59] utilisent SWRL pour définir des règles au niveau du modèle.
Onwubiko [60] présente CoCoa, une ontologie qui reprend le processus du SOC dès la collecte des logs à la résolution de l'incident. Comme sources, on retrouve les journaux, les informations du réseau telles la topologie, la configuration sur le Configuration and Management Database *(CMDB)*, les données non structurées comme les courriers échangés et les indicateurs de compromission qui sont rapportés par les renseignements sur la menace. *Onwubiko* [60] propose de modéliser un incident de sécurité et les relations entre toutes les entités proposées sous la forme d'un graphe de connaissance basé sur les ontologies. Toutefois, le développement de cette ontologie n'a pas été clairement expliqué, et aucune évaluation n'a été faite pour voir le résultat d'une requête par exemple.
L. Obrst et al. [61] proposèrent une méthodologie de développement d'ontologie pour le

domaine de cybersécurité en choisissant l'approche intermédiaire. Alors qu'il existe trois approches.

Approche de bas en haut Cette analyse se base sur la compréhension des données qui arrivent pour l'intégration dans le modèle ontologique. Elle permet de définir les concepts à partir des sources d'information telles les journaux.

Approche de haut en bas Cette analyse démarre des utilisateurs finaux et les différentes requêtes qu'ils veulent lancer sur l'ontologie. Ces questions seront aussi considérées comme des tests de validation.

Approche de l'intermédiaire Cette approche tente de faire la liaison entre les deux techniques précédentes.

Baesso Moreira et al. [62] proposent Computer Security Incident Handling Ontology *(CSIHO)*, une ontologie pour la gestion des incidents de sécurité. Elle couvre le domaine de traitement des incidents dans le but d'améliorer le processus de réponse aux attaques. Les auteurs ont évalué les précédentes contributions et ont défini sept classes pour modéliser l'environnement à savoir : ***Computer_Asset*** qui représente un noeud sur le réseau, étrangement seulement la station et le serveur ont été définis. ***Course_of_Action*** donne l'action courante qui est adoptée. ***Incident*** la classe principale associée avec tous les autres éléments de l'ontologie. ***Person***, cette classe fait référence aux analystes de différents niveaux. ***Security_Event*** définit les évènements de sécurité déclenchés divisés en sous-classes selon l'origine. L'évènement vient de l'antivirus, du pare-feu ou autre solution de sécurité déployée sur place. Un incident a un aspect temporel et passe par plusieurs étapes. En effet, les auteurs proposent alors la classe ***Timeline_Occurence*** pour indiquer la phase à laquelle l'incident se trouve parmi les phases principales de réponse à incident.
En vue d'évaluer **CSIHO**, les questions posées pour définir le cadre de l'ontologie dans l'étape de développement sont traduites en des requêtes *SPARQL*. Comme cas d'utilisation, les auteurs ont opté pour le maliciel WanaCry [7]

2.8 Technologies utilisées pour la détection de la menace interne

De nombreux types de produit permettent d'avoir une protection contre les menaces qui viennent de l'intérieur du réseau de l'entreprise. Chacun offrant une approche différente. Dans la famille des DLP, on retrouve deux catégories représentées dans la figure 2.6. Les solutions DLP pour entreprise sont utilisées pour avoir une vue globale incluant les machines, le réseau, la navigation Web, le courriel et les services nuages. Cette catégorie permet d'avoir

7. https ://us-cert.cisa.gov/sites/default/files/FactSheets/NCCIC ICS_FactSheet _WannaCry_Ransomware_S508C.pdf

une seule console pour surveiller l'ensemble des activités des utilisateurs dans l'entreprise. Généralement, les produits inscrits dans ce type de produit incluent des techniques d'inspection avancées qui permettent de créer divers scénarios complexes. Par contre, la quantité de données est considérable dans ce genre de solution et parfois elle risque de jouer en défaveur de la détection [2]. La seconde famille comporte les solutions DLP intégrées qui sont destinées à surveiller un canal de communication en particulier. Les analystes consultent les consoles de chaque solution pour analyser les évènements enregistrés [2]. De nombreux vendeurs pro-

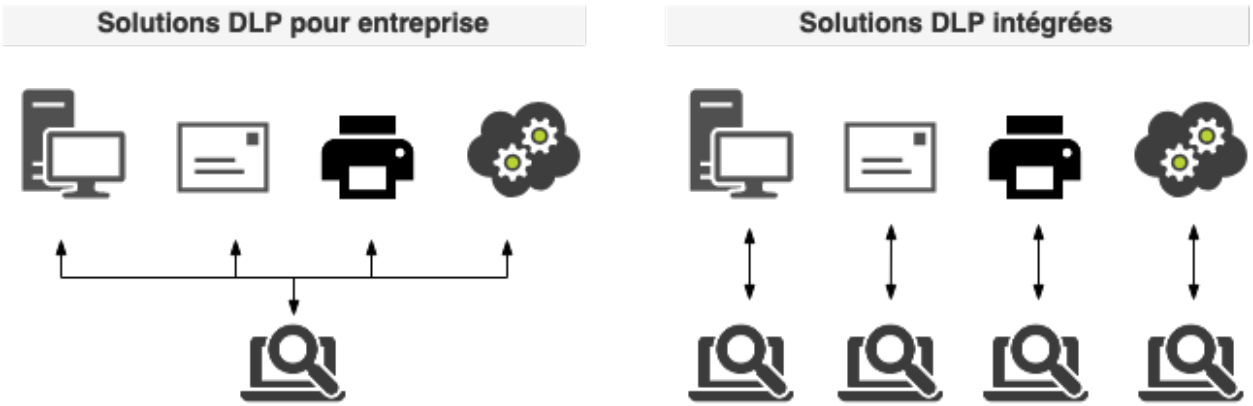

FIGURE 2.6 Les types du DLP [2]

posent des produits DLP pour aider les entreprises dans leur stratégie de protection contre la menace interne. Le tableau 2.3 énumère quelques produits par vendeur.

TABLEAU 2.3 Exemples de produits DLP

Barracuda	Barracuda DLP
Symantec	Symantec DLP
Cisco	Cisco DLP
Forcepoint	Forcepoint DLP
McAfee	McAfee *Total Protection for* DLP
Palo Alto Networks	Palo Alto *Networks Cloud-Delivered Data Protection*
Proofpoint	Proofpoint *Data Loss Prevention*
Zscaler	Zscaler DLP

Les solutions UEBA sont maintenant très déployées dans les entreprises. Ces solutions permettent d'observer le comportement des entités sur le réseau. Elles attribuent un score qui définit le risque d'une entité qui peut être un individu, d'où l'utilisation en menace interne. Au début, les solutions UEBA étaitent développées séparément des *Security Information and Event Management* (SIEM), tandis qu'actuellement les constructeurs de solutions de sécurité

s'orientent vers l'intégration des UEBA dans les SIEM pour définir de nouveaux scénarios ou cas afin d'accroitre le niveau de protection contre les menaces [63]. Des exemples de produits sont cités dans le tableau 2.4.

Tableau 2.4 Exemples de produits UEBA

Exabeam	*Advanced Analytics*
Forcepoint	UEBA *Behavioral Analytics*
LogRhythm	UEBA
RSA	Netwitness UEBA
Splunk	UBA

2.8.1 Conclusion du chapitre

Les sections discutées dans ce chapitre démontrent l'intérêt pour la menace interne du point de vue académique et industriel. Les différents travaux de recherche optent pour l'utilisation des techniques de l'apprentissage machine, car l'option de se baser sur la détection par règle engendre un travail répétitif lors de l'investigation. Les différents jeux de données CERT, Schonlau, RUU et TWOS nous permettent de nous orienter vers les évènements qu'il faut réunir afin d'avoir une visibilité globale sur l'activité de l'utilisateur sur le réseau. Le prochain chapitre donne un aperçu sur l'écosystème de détection de menace interne pour une grande entreprise dans le secteur des finances.

CHAPITRE 3 ÉCOSYSTÈME DE LE MENACE INTERNE

3.1 Introduction

Ce chapitre présente le programme de protection contre la menace interne mise en oeuvre par le partenaire industriel. Ce programme s'inscrit parmi les programmes de sécurité de l'information, il possède une vision et des objectifs bien définis et ne repose pas seulement sur une composante technologique. L'approche que nous proposons doit s'inscrire dans la même vision et contribuer aux mêmes objectifs.

3.2 Environement

La détection d'incident relative à la menace interne se base sur un écosystème de surveillance. Cet écosystème inclut une partie technologique qui collecte les évènements depuis des sources de données qui représentent les actifs de l'entreprise, puis les prépare au traitement avant de les proposer aux analystes pour une analyse éventuelle. En plus de la couche technologique, un processus d'investigation et de réponse à incident est mis en place. Les analystes suivent les étapes décrites dans ce processus pour s'assurer d'un passage entier sur les éléments de l'incident. D'habitude, l'investigation des incidents de sécurité se passe dans un SOC et les incidents relatifs aux intrus internes sont traités par les analystes qui traitent aussi les évènements liés à la tentative d'exploitation de vulnérabilité ou la propagation de programmes malveillants. L'environnement dans lequel ce travail est effectué est différent, une importance est donnée à la protection contre la menace interne et des équipes sont dédiées pour s'assurer d'identifier les actions douteuses qui se sont produites sur le réseau.
Introduire une nouvelle approche ou une amélioration nécessite de prendre en considération le système en place et répondre aux besoins exprimés. Souvent, des projets qui ont coûté de l'argent et du temps se retrouvent sans suite, car ils ne satisfaisaient pas les exigences ou ont du mal à s'adapter à l'environnement. lors de ce projet, nous proposerons une mise à jour qui s'intègre au fonctionnement du coté technologique et qui s'intègre dans le processus d'investigation.
Bien que l'intelligence artificielle apporte des solutions automatiques qui simplifient et réduisent la charge du travail, certaines approches se présentent comme des boites noires et les usagers ont du mal à les interpréter. Les solutions qui se présentent de cette manière sont critiquées tout le temps par les utilisateurs finaux, les analystes dans notre cas.

3.3 Équipes participantes dans la menace interne

Assurer une fluidité et continuité de la surveillance dans le cadre de la menace interne nécessite la collaboration de plusieurs équipes au sein de l'entreprise. Dans notre environnemnt, quatre équipes existent pour assurer le fonctionnement et l'évolution de l'écosystème en place. L'équipe **Stratégie et connaissance de la menace** s'occupe d'identifier, définir et prioriser les scénarios d'activité douteuse sur le réseau afin de les ajouter à la surveillance globale. Cette équipe communique avec les autres départements ou directions pour évaluer les potentielles menaces ou collecter les informations sur les actions jugées non appropriées pour les employés. Les données nécessaires à la surveillance de l'activité notée sont alors mentionnées à l'équipe **Outils et développement de la plateforme de surveillance** qui s'occupe de collecter les évènements liés au scénario défini. Si le scénario nécessite d'introduire une nouvelle règle sur une solution de protection, cette dernière sera implémentée et les évènements subséquents reçus. L'équipe Outils s'occupe aussi du maintien de la plateforme analytique en termes d'évolution et résolution de problème. Les analystes au sein de l'équipe **opérationnelle** réalisent des investigations sur les utilisateurs ayant les scores de risque les plus élevés et analysent chaque évènement enregistré afin de confirmer si le score est lié à une action dangereuse. Les investigations se transforment en des billets sur la solution de management des incidents et les analystes veillent au suivi du billet avec les parties prenantes jusqu'à la résolution de l'incident. Cette équipe joue un double rôle dans l'intégration de l'IA. Ses membres sont les consommateurs des recommandations produites par les modèles et participent à produire des retours quant aux incidents, ce qui permet d'avoir des éléments étiquetés, ou quant à la performance d'un modèle en production.
La dernière équipe **Valorisation et science des données** s'occupe de produire des changements à travers des requêtes aux données reçues et construire des modèles en apprentissage machine afin d'aider les analystes dans l'investigation et leur éviter une charge de travail sur de faux positifs et les orienter plus vers les cas les plus pertinents. L'objectif final est d'identifier des patrons qui permettront de prédire un comportement qui ne respecte pas les mesures de sécurité en place, avant sa production et éviter ainsi l'impact qui l'accompagne.

3.4 Architecture transversale

Acheminer de l'information depuis les sources d'évènement et les présenter aux analystes passe par une architecture technologique composée de quatre composants. La figure 3.1 reprend ces parties dans lesquelles les données transitent depuis les sources vers la publication.

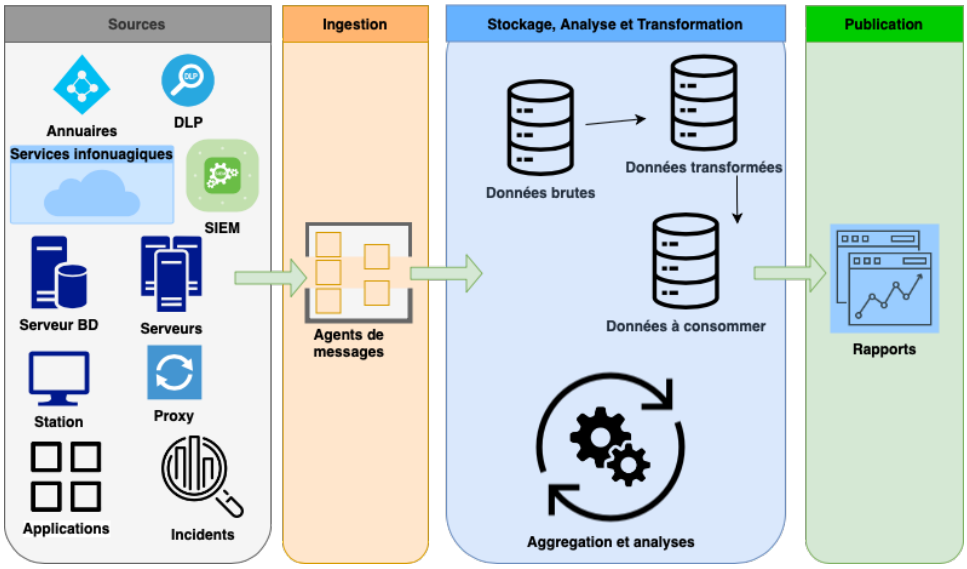

FIGURE 3.1 Architecture globale de la plateforme analytique pour la menace interne

3.4.1 Sources

La composante de sources regroupe l'ensemble des actifs nécessaires à interroger afin de récupérer les actions d'un utilisateur. Cependant ces sources dépendent des scénarios identifiés par l'équipe de stratégie et l'ensemble des actifs de l'entreprise n'est pas consulté. Accéder aux enregistrements diffère d'un actif à un autre. La plupart des sources d'évènement peuvent être configurées pour transmettre les fichiers journaux vers une destination spécifique à travers un canal de communication, mais ce n'est pas le cas pour toutes les sources, pour certaines les programmes du composant de l'ingestion se chargent de collecter les données depuis les sources.

Selon le scénario défini, des règles sont implémentées pour chercher une action bien précise que l'utilisateur a effectuée. Des signatures sont paramétrées sur les DLP pour signaler une donnée spécifique. Par exemple, un numéro de carte de crédit inclut dans un document envoyé en attaché sur un courriel sortant vers une boite qui semble être la boite personnelle de l'individu ayant fait l'action.

Les classes de source importantes sont énumérées dans le tableau 3.1. une description est associée à chaque classe et son intérêt dans la surveillance.

On observe deux types de données depuis les sources

Données structurées La majorité des sources produisent des journaux qui suivent un schéma prédéterminé par le vendeur de la solution ou de l'équipement. Les analyseurs réussissent à identifier chaque élément d'un enregistrement.

Données non structurées Certaines informations se présentent sous forme textuelle par

exemple, on retrouve souvent du texte dans les courriels. Ces données ont besoin de traitement pour extraire les éléments qui nous intéressent, mais le traitement diffère à cause de l'absence de contraintes qui régissent l'information. Par exemple, deux analystes reportent différemment à l'incident au niveau du journal.

3.4.2 Ingestion

Ce composant inclut des services de messagerie distribués pour acheminer les données des sources vers l'espace de stockage réservé à cet effet. Des travaux planifiés sont aussi exécutés pour récupérer les données.

3.4.3 Stockage, Analyse et Transformation

Toutes les données acheminées sont stockées dans un espace dédié pour le traitement de la menace interne. Une analyse est effectuée pour avoir certaines informations sur l'utilisateur et pour les corréler. À ce niveau, des indicateurs sont créés par rapport à un évènement ou un ensemble d'évènements.

3.4.4 Publication

La publication est représentée sous forme d'un tableau de bord qui contient la liste des utilisateurs avec leurs scores de risque qui dépend des indicateurs qu'ils ont déclenchés. Le tableau de bord est aussi classé par type de menace pour simplifier la recherche par les analystes sur des cas particuliers. Par exemple, le scénario qui implique l'exfiltration par courriel. En plus des scores, les évènements sont regroupés par groupe de techniques utilisées par le potentiel attaquant, ce qui permet aux analystes de voir à quelle étape de la chaîne d'attaque l'utilisateur est rendu. Cette chaîne d'attaque ressemble à celle discutée dans la section 2.4 du chapitre revue de littérature.

3.5 Indicateurs et calcul de score de risque

Les actions réalisées par les usagers sur le réseau de l'entreprise sont traduites en des indicateurs. Un indicateur est une information extraite suite à un évènement ou une corrélation entre plusieurs évènements. En fait, l'indicateur est le résultat de traitement d'une donnée reçue depuis les sources de données ou une information contextuelle pour un compte d'utilisateur.

Deux types d'indicateurs existent. Le premier type a un aspect dynamique et regroupe les indicateurs originaires d'évènements enregistrés pour un utilisateur qui se sont produits sur le réseau. Le second reflète le privilège octroyé, le privilège peut être le fait d'avoir accès au

Tableau 3.1 Tableau sur les sources les plus importantes

Source	Description
Annuaires	L'identité est un élément important et critique en menace interne, elle permet d'identifier l'individu au sein de l'organisation. Tous les évènements produits par un compte sont liés à un utilisateur. Toutefois dans une grande organisation, un utilisateur peut avoir plusieurs comptes sur le réseau par mesure de sécurité et certains administrateurs se voient attribuer un compte à haut privilège pour protéger le réseau si le compte de base se fait compromettre.
DLP	Les DLP sont déployés aux frontières des canaux de communications pour inspecter le contenu et effectuer une analyse contextuelle afin d'identifier des informations sensibles que l'utilisateur veut transmettre à l'extérieur de l'entreprise. La stratégie configurée sur le DLP permet de bloquer l'action ou juste la signaler. Depuis cette classe on reçoit les alertes qui servent à créer les indicateurs.
SIEM	Le SIEM sert à collecter les évènements des actifs de l'entreprise pour la surveillance, il a été utilisé longtemps dans notre evrionnement pour permettre l'investigation et l'édition d'alerte. Il représente un magasin d'anciens évènements qui ne sont plus disponibles sur les sources elles-mêmes. Il est aussi utilisé comme intermédiare, car la majorité des sources lui envoie les enregistrements.
Serveur BD	Les requêtes envoyées au serveur de base de données sont aussi récoltées pour voir si des utilisateurs consomment des données sensibles depuis la base de données.
Proxy Web	Toute l'activité de navigation sur les sites Web passe par la passrelle Web qui filtre les sites dangereux. Les utilisateurs malveillants ou négligents visitent surtout les sites de partage pour envoyer les données à l'extérieur de l'entreprise. Pour cette raison cette source est importante et doit être incluse.
Incidents	Cette source regroupe les anciens billets ouverts dans le contexte de menace interne et représente une source de données importante, car elle permet de retrouver les évènements qui ont participé à l'incident et c'est le seul endroit où on peut avoir l'information sur la classification de l'incident (vrai positif ou faux positif). Les incidents sont analysés pour comprendre les évènements qui se sont produits afin de développer des applications intelligentes qui peuvent prédire les actions.
Services infonuagiques	L'entreprise utilise des solutions infonuagiques et des utilisateurs manipulent les données sur ces services. Il est important d'avoir une visibilité sur l'activité de l'utilisateur sur les différentes plateformes accessibles.
Station	L'activité locale sur les machines des utilisateurs est aussi prise en compte pour voir par exemple les impressions ou les copies vers les clés amovibles.
Serveurs	Certains serveurs contiennent des informations sensibles et il est important de voir quel utilisateur y accède et quelles actions il effectue.

compte administrateur sur la machine locale. Ce type inclut aussi la classe de l'information accédée selon qu'elle soit stratégique ou technique. Un autre type d'indicateur a été proposé, mais son inclusion reste limitée du fait de la vie privée des employés. Un exemple serait une période personnelle difficile tels un divorce ou des problèmes financiers. Chaque indicateur possède une cote de risque dont la valeur est attribuée par l'équipe stratégie lors de la définition d'un scénario de menace. Cette valeur dépend de la criticité de l'action faite par l'utilisateur et elle est sujette à un changement si elle ne favorise pas l'investigation. Le tableau 3.2 reprend des exemples d'indicateurs. La cote de risque dépend beaucoup de la gravité ou du type d'information transmise.
Les indicateurs de privilèges ne sont cumulés qu'une seule fois, tandis que les indicateurs dynamiques le sont à chaque évènement. L'indicateur **Ind_3** sur le tableau 3.2 est inclus dans l'indicateur **Ind_4**, lors du calcul de score seule la contribution de l'indicateur Ind_4 est prise en compte.

Tableau 3.2 Exemple sur les indicateurs

Indicateur	Type	Description	Cote de risque
ind_1	Privilège	Privilège d'utilisations des ports USB	10
ind_2	Comportement	Évènement d'envoi de fichier avec information financière vers une boite personnelle	100
ind_3	Comportement	Évènement d'envoi de fichier avec trois occurrences d'information privée	150
ind_4	Comportement	Évènement d'envoi de fichier avec quinze occurrences d'information privée	400

Le score final de l'utilisateur est calculé avec l'ensemble des indicateurs enregistrés pour ses comptes sur le réseau. Le calcul est la somme des cotes de risque des indicateurs déclenchés en lien avec des évènements plus les côtes de risque des indicateurs de privilège (Voir équation 3.1).

$$score(individu) = \sum_{i=1}^{n} c_i \tag{3.1}$$

où : c = indicateur

3.6 Processus d'investigation

Les analystes dans l'équipe opération consultent le tableau de bord des scores de risque et procèdent à l'évaluation des actions effectuées par l'utilisateur. En fait, les analystes prennent les utilisateurs avec les scores les plus élevés et ouvrent sur la solution de management des

incidents à cet effet un billet de *hunt*. Ce billet sert à faire un suivi sur les utilisateurs sélectionnés. Les indicateurs de chaque utilisateur dans ce billet seront analysés afin de confirmer la menace. L'analyste retourne son analyse sur le billet et procède à la réduction des scores de tous les individus sous investigation pour ne pas les revoir d'aussitôt ou pour éviter qu'un autre analyste les traite.
Lorsque l'analyse est fructueuse et que le *hunt* est positif sur un utilisateur, un billet d'investigation est ouvert. Ce billet d'investigation concerne un seul individu et lorsque ce niveau d'enquête est atteint, on considère le billet comme étant associé à un vrai positif avec une probabilité de 80%. La suite sera de communiquer les artefacts soit au gestionnaire de l'individu ou à une équipe d'enquêteurs spécialisée. Une fois qu'un retour est reçu l'incident sera mise à jour et le billet fermé.

3.7 Analyse de la problématique d'investigation

La solution disponible au niveau de notre environnement propose un moyen pour détecter les menaces, mais ne prend pas en considération les anciennes investigations. Il s'ensuit que les mêmes utilisateurs reviennent assez souvent avec les mêmes indicateurs. Ceci impacte le travail des analystes sachant que la majorité des utilisateurs avec un haut score ont fait des actions légitimes et que l'investigation va produire de faux positifs. En outre, certains utilisateurs se trouvent en bas du tableau de bord ne seront pas analysés à temps ou jamais en raison des faux positifs qui prennent le haut de la liste avec des scores élevés.

Tableau 3.3 Exemple du tableau de bord

Ligne	**Utilisateur**	**Score**
01	user02	1200
02	user04	1200
03	user51	1000
04	user17	900
	• • •	
2000	user533	57

L'exemple présenté dans le tableau 3.3 reprend l'idée présentée dans cette section. Alors que pour les usagers des lignes une à quatre sont tous des cas de faux positifs, l'utilisateur **user533** avec un score de 57 se trouve à la ligne 2000 et il a moins de chance d'être mis sous investigation bien qu'il ait effectué une action qui viole la politique de sécurité.

3.7.1 Conclusion du chapitre

Dans ce chapitre, nous avons exploré l'écosystème mis en place pour la détection de menace interne chez le partenaire industriel. L'objectif est d'identifier l'endroit auquel l'amélioration doit être appliquée pour aider dans le processus de détection. Toutefois, l'approche proposée doit s'aligner avec le système global en place. Les solutions tels les UEBA qui se basent sur un calcul de score de risque ressemblent à ce qui a été défini dans la section 3.5. Elles n'incluent pas le contexte de l'utilisateur dans son environnement. L'approche que nous proposons dans la suite de ces travaux vise à inclure le contexte organisationnel des employés.

CHAPITRE 4 APPROCHE ET MODÈLES

4.1 Introduction du chapitre

L'écosystème présenté dans le chapitre précédent inclut deux informations majeures qui sont comme suit ; à chaque utilisateur est associé un ensemble d'indicateurs qu'il a déclenchés suite à son activité sur le réseau ou à ses privilèges. En plus, cet utilisateur évolue dans une organisation et possède des relations professionnelles avec d'autres collègues.
Notre objectif est de concevoir un modèle qui peut évoluer sur les vecteurs suivants.

— Propriétés de l'utilisateur qui est le vecteur des indicateurs qui lui est associé.

— Relations au sein de l'organisation telle la relation d'appartenance à la même équipe.

— Les informations remontées dans les billets d'investigation précédents.

Assez souvent les analystes rapportent que les membres d'une même équipe ou qui partagent des responsabilités ont tendance à avoir un comportement similaire et déclenchent les mêmes indicateurs. Bien que ce raisonnement soit connu, cette information n'est pas exploitée et les analystes se retrouvent à refaire l'investigation sur les indicateurs déclenchés par un utilisateur même si récemment l'activité d'un utilisateur de son entourage a été étiqutée faux positif pour les mêmes évènements. Notre approche vise à tirer profit de cette logique pour vérifier si les indicateurs d'un usager ressemblent aux indicateurs d'un autre de son milieu professionel dont les activités ont été analysées. L'approche ne se limite pas à la découverte des faux positifs seulement, mais sert aussi à identifier les vrais positifs.
Dans notre étude,le nombre d'utilisateurs considérés est infime comparé au nombre d'employés dans l'entreprise. Ce constat nous pousse alors à nous orienter vers un apprentissage semi-supervisé qui utilise des données étiquetées et non étiquetées. Cette classe d'apprentissage se situe entre l'apprentissage supervisé dans lequel toutes les données sont étiquetées et l'apprentissage non supervisé ou aucune donnée n'est étiquetée.

4.2 Plan du chapitre

Ce chapitre inclut principalement des expérimentations. En premier lieu, nous présentons la méthode de graphe de convolution. Ensuite, nous proposons de modéliser les relations pour construire le graphe d'entrée à notre GCN. La première évaluation inclut le jeu de données global et a été réalisée sur les conceptions de graphes proposées. Puisque, le jeu de données globale est grand, il est assez difficile d'analyser les résultats sachant en plus qu'il n'existe pas de références antérieurs pour comparer. À cet effet, nous avons exploré un autre jeu de données pour lequel nous avons déjà les résultats pour les modèles supervisés qui vont nous

permettre d'évaluer notre nouveau modèle. Afin de produire un graphe d'entrée qui reflète la réalité de la structure au sein de l'entreprise, nous avons exploré les marches aléatoire avec node2vec.
Nous avons utilisé le graphe obtenu grâce à node2vec comme entrée à notre GCN pour le jeu de données restreint et nous avons comparés aux résultats des autres méthodes. La dernière expérimentation comporte le fait qu'on augmente les attributs du jeu de données avec les vecteurs obtenu grâce à node2vec, puis on réévalue les résultats.

4.3 Graphe de données

Les données récoltées sont projetées sur un graphe G, dont les noeuds N sont les usagers et les liens dans l'ensemble E sont les relations organisationnelles. le graphe G est défini par 4.1.

$$G = (N, E) \tag{4.1}$$

La figure 4.1 montre une représentation graphique de la projection des données. À chaque usager $n_i \in N$ est associé un vecteur d'indicateurs $vi \in V$.

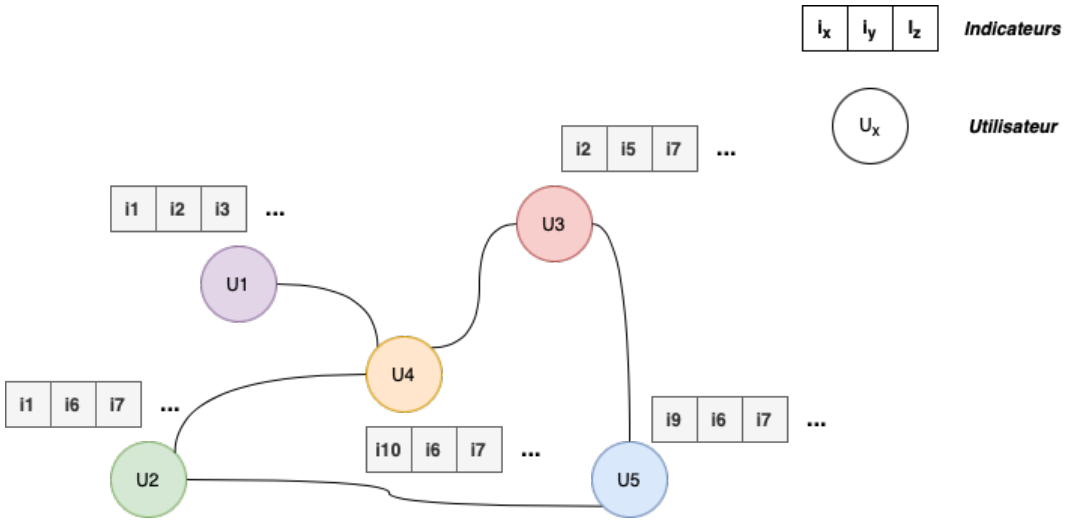

FIGURE 4.1 Représentation graphique des données

4.4 Réseau de graphe de convolution

Le graphe résultant de la projection des données est la donnée en entrée de notre modèle de graphe de convolution dans lequel notre but est de classifier les noeuds du graphe. Le modèle GCN est un réseau de neurones qui opère sur les graphes dont la règle de propagation générale entre les couches est définie en 4.3. Il correspond parfaitement à nos conditions citées sur les

vecteurs sur lesquelles notre modèle évolue, à savoir le vecteur des indicateurs, les relations dans l'entreprise et la prise en compte des incidents précédents. Il répond ainsi à notre objectif d'apprentissage semi-supervisé.
Ce modèle construit la fonction $f(H, A)$ où H représente la matrice des indicateurs des usagers et A la matrice d'adjacence.

$$H^{(l+1)} = f(H^{(l)}, A) \tag{4.2}$$

$$f(H^{(l)}, A) = \sigma(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}) \tag{4.3}$$

où :

$H^{(l+1)}$: représente la matrice des features de la couche suivante

σ : la fonction d'activation qui définit comment la somme pondérée de l'entrée est transformée en une sortie d'un ou plusieurs nœuds d'une couche du réseau de neurones.

$\tilde{D}$: $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ la matrice des degrés des noeuds du graphe.

$\tilde{A}$: $\tilde{A} = A + I_N$ la matrice d'adjacence du graphe plus la matrice d'identité.

$H^{(l)}$: $H^{(l)} \in \mathbb{R}^{N \times F}$ est la matrice d'activation de taille N*F (ou N est le nombre de noeuds et F le nombre d'indicateurs par noeud) à la couche l du réseau de neurones. $H^{(0)} = X$ où X est la matrice des vecteurs de features pour les noeuds. Dans notre cas les vecteurs d'indicateurs des utilisateurs.

$W^{(l)}$: est une matrice de poids entraînable spécifique à la couche du réseau de neurones.

Cette équation réalise des transformations sur chaque vecteur d'indicateurs d'un usager pour inclure les vecteurs d'indicateurs de ses voisins selon la matrice d'adjacence et les degrés des noeuds. L'exemple suivant donne le processus de transformation appliqué avant l'exécution de la première couche et en général avant l'exécution de chaque couche avec les résultats de la couche précédente. L'équation 4.4 construit la matrice d'adjacence augmentée $\tilde{A}$ par la somme de la matrice d'adjacence originale A et la matrice d'identité I qui nous permet lors du calcul de garder les caractéristiques du noeud lui-même.
Afin d'éviter le problème de diminution ou d'explosion de gradient au niveau du réseau de neurones [64] et garder un contrôle sur les valeurs de sortie du modèle, la matrice d'adjacence $\tilde{A}$ est multipliée par la matrice inverse $\tilde{D}$. Les auteurs dans l'article d'origine de *Kips et Welling* [65] ajoutent $\tilde{D}^{\frac{1}{2}}$ à droite et à gauche de $\tilde{A}$ pour normaliser par ligne et par colonne. La matrice d'adjacence A dans l'équation 4.4 représente les liaisons du graphe de la figure 4.2. I_N est une matrice identité de taille (5×5). La matrice de degré D dans l'équation 4.5 donne le nombre de liaisons pour chaque noeud sur la matrice $\tilde{A}$. La matrice $\tilde{D}^{-1}$ est la

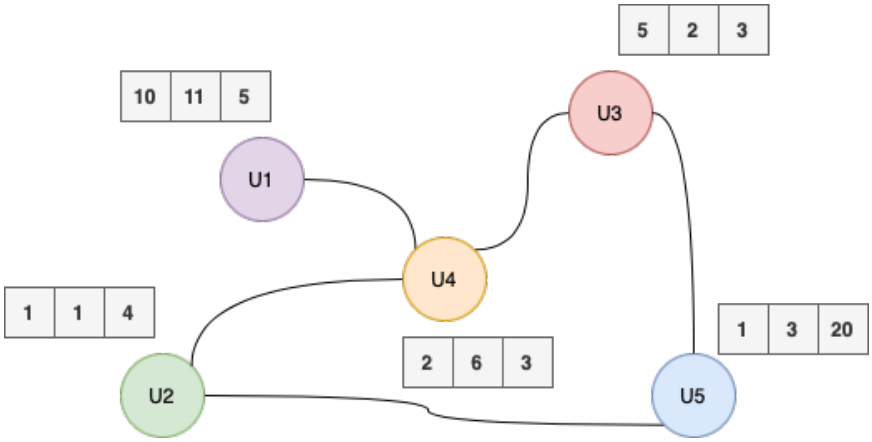

FIGURE 4.2 Exemple de graphe

matrice inverse de la matrice des degré D. La matrice X de l'équation 4.6 est l'ensemble des vecteurs de chaque noeud, la première ligne est le vecteur associé au noeud U1 et la deuxième pour U2 et ainsi de suite pour tous les noeuds.

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix} = \overbrace{\begin{bmatrix} & U1 & U2 & U3 & U4 & U5 \\ U1 & 0 & 0 & 0 & 1 & 0 \\ U2 & 0 & 0 & 0 & 1 & 1 \\ U3 & 0 & 0 & 0 & 1 & 1 \\ U4 & 1 & 1 & 1 & 0 & 0 \\ U5 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}}^{A} + \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}^{I_N} \tag{4.4}$$

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \quad \tilde{D}^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} \end{bmatrix} \quad \tilde{D}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \tag{4.5}$$

$$X = \begin{bmatrix} 10 & 11 & 5 \\ 1 & 1 & 4 \\ 5 & 2 & 3 \\ 2 & 6 & 3 \\ 1 & 3 & 20 \end{bmatrix} \tag{4.6}$$

$$\underset{\tilde{D}^{\frac{1}{2}}}{\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} \end{bmatrix}} \underset{\tilde{A}}{\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}} \underset{\tilde{D}^{-\frac{1}{2}}}{\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} \end{bmatrix}} \underset{X}{\begin{bmatrix} 10 & 11 & 5 \\ 1 & 1 & 4 \\ 5 & 2 & 3 \\ 2 & 6 & 3 \\ 1 & 3 & 20 \end{bmatrix}} \tag{4.7}$$

La figure 4.3 reprend l'exemple cité dans l'article [3] dans lequel le modèle GCN utilise deux couches pour l'apprentissage. La règle de propagation de l'équation 4.3 est envoyée comme paramètre pour les deux fonctions d'activation $ReLU$ et $Softmax$.

ReLU La fonction d'activation *Rectified Linear activation Unit* (ReLU) définie en 4.8 est une fonction non linéaire utilisée dans l'apprentissage profond. Cette fonction retourne zéro si l'entrée est négative. Elle se comporte comme la fonction identité si l'entrée est positive (elle retourne la valeur de l'entrée) [66].

Softmax définie par la relation 4.9, elle calcule la distribution de probabilité d'un vecteur de nombre réel dont la somme est un. Elle est utilisée dans les modèles de classification pour retourner la probabilité de chaque classe [66].

$$f(x) = max(0, x) \tag{4.8}$$

$$softmax(x_i) = \frac{e_i^x}{\sum_i e_i^x} \tag{4.9}$$

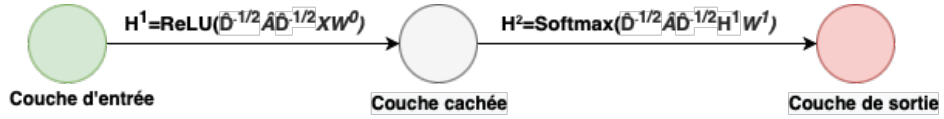

FIGURE 4.3 Exemple avec deux couches du GCN [3]

4.5 Relations dans le graphe

Certes les employés partagent des relations au sein de l'entreprise, mais les modéliser et les inclure dans notre graphe reste un défi. Chaque relation ou couple de relations choisies

risque d'impacter notre apprentissage. Par exemple, la relation de supérieur direct peut être interprétée de différentes manières. On peut la considérer entre l'employé et son manager et créer un lien entre les deux, ainsi le graphe formera alors une structure hiérarchique (voir la structure du graphe avec la relation de manager en (a) dans la figure 4.4). Cependant, cette même notion de relation peut se modéliser par la création de liens entre les employés qui possèdent le même supérieur direct (voir la structure du graphe avec ce point de vue en (b) dans la figure 4.4). Dans la dernière structure en (c) dans la figure 4.4, on considère les deux relations citées précédemment. Le choix des relations est important, car l'agrégation

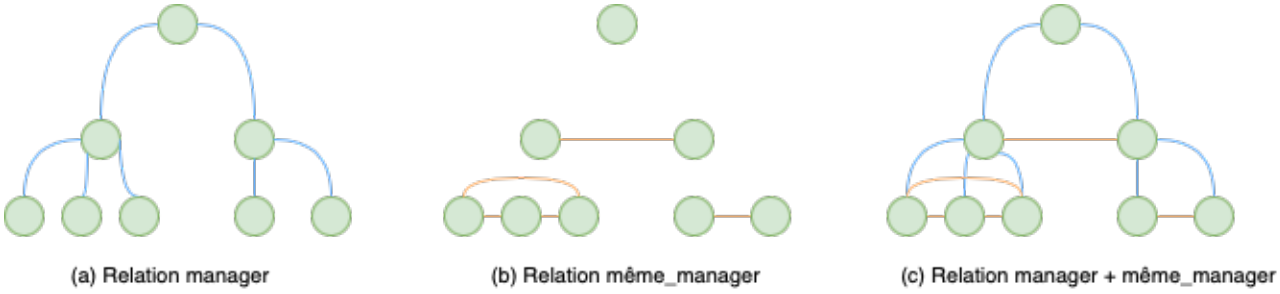

FIGURE 4.4 Schéma de graphe avec différentes relations

par la règle de propagation des vecteurs d'indicateurs se fait entre les entités liées et chaque relation sélectionnée, et génère une matrice d'adjacence différente. Par ailleurs, cela dépend aussi du nombre des couches dans notre réseau de neurones. L'agrégation peut atteindre, par exemple, le voisin du voisin si le nombre des couches est de deux. Si la relation de manager est maintenue seule avec deux couches, l'agrégation permet de récupérer les vecteurs des personnes ayant le même manager. Par contre, si la relation du même_manager est maintenue, le voisin est atteint deux fois ou plus par liaison directe et par liaison indirecte à travers une personne de la même équipe.

On observe les matrices d'adjacences selon les différentes relations de la figure 4.4 dans l'équation 4.10. Les deux matrices relation manager et relation même_manager sont différentes et lorsqu'on fait la somme entre ces deux matrices on retrouve une autre matrice encore différente.

(a) relation manager

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(b) relation même_manager

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.10}$$

(a) relation manager + même_manager

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.11}$$

4.5.1 Expériences avec le GCN

Nous avons réalisé plusieurs expérimentations sur le modèle en GCN avec plusieurs idées de conception de graphe. Ces conceptions tiennent compte des différentes relations qu'on peut observer au sein de l'entreprise. La configuration du modèle est détaillée dans le tableau 4.1 et a été maintenue pour l'ensemble des expériences sur les sections suivantes. Nous avons utilisé la librairie Deep Graph Library (DGL) [67] qui regroupe plusieurs algorithmes d'apprentissage

profond destinés à faire des calculs sur les structures de graphe. Elle a été proposée à NYU [1] Shanghai par les deux Professeurs Zheng Zhang et Quan Gan et le site officiel [68] indique d'autres contributeurs qui se sont joints à son développement du milieu académique représenté par l'université de New York, NYU Shanghai et l'université de Fudan ainsi que du milieu industriel par Intel et NVIDIA.

Tableau 4.1 Paramètres pour le GCN

Nombre de couches	2
Nombre de neurones pour la couche cachée	40
Nombre d'*epoch*	100
Optimiseur Adam, taux d'apprentissage	0.01

Le nombre d'*epoch* ou itérations représente le nombre de fois où le modèle a eu l'opportunité de faire un passage sur les données d'entrainement et de mettre à jour les poids internes du modèle. L'optimiseur Adaptive Moment Estimation (Adam) [69] est un algorithme d'optimisation du gradient utilisé pour mettre à jour les poids du réseau de neurones. Cet algorithme est efficace en termes de calcul et nécessite peu de mémoire. Nous avons opté pour cette configuration afin de suivre les configurations sur l'article de la détection de fraude avec les GCN [45].

Graphe sans relation

Le graphe sans relation qu'on fournit comme donnée d'entrée à notre modèle GCN sera notre base de référence pour observer l'évolution en présence des différentes relations.

Graphe contrsuit avec la relation de manager

On construit notre graphe en créant une relation entre l'employé et son manager direct. La figure 4.5 donne la forme du graphe avec la relation de manager.

Graphe contrsuit avec la relation du même manager

Le graphe dans cette section, à la différence de la section précédente, est construit en créant un lien entre les usagers qui possède le même responsable. La figure 4.6 montre la différence dans le graphe lorsqu'on utilise la relation du même manager.

1. *New York University*

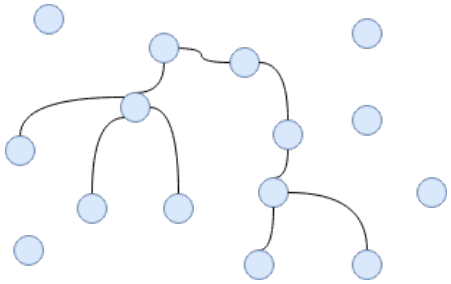

FIGURE 4.5 Graphe avec la relation de manager

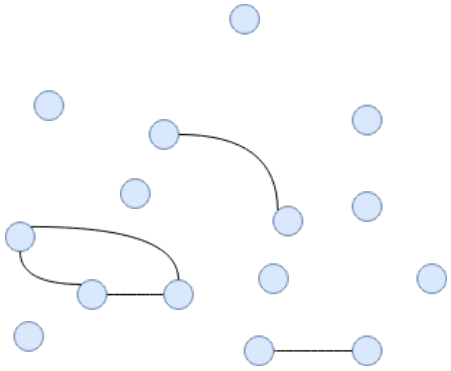

FIGURE 4.6 Graphe avec la relation de même manager

Graphe contrsuit avec la relation du même première vice-présidence

Ici on réunit les usagers qui font partie de la même première vice-présidence dans la compagnie, cette relation permet de lier plus d'individus. La figure 4.7 montre les relations entre chaque noeud sur chacune des deux pvp (PVP1 et PVP2). On remarque que le nombre de liens est plus grand que dans les graphes avec les relations des sections précédentes (manager et même manager). Il existe dans ce graphe un lien entre chaque noeud pour chaque pvp.

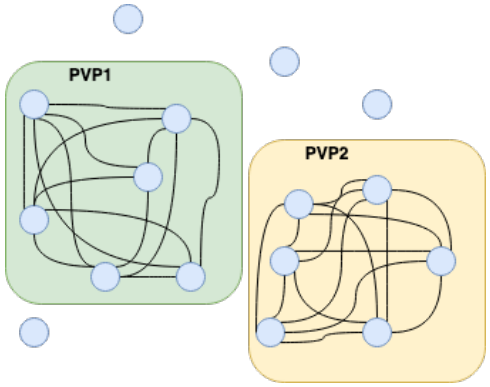

FIGURE 4.7 Graphe avec la relation de la même PVP

TABLEAU 4.2 Détails et résultats des diffétentes expériences

	GCN			
	Sans relation	Relation manager	Relation même manager	Relation même pvp
Nombre de noeuds	51965			
Nombre d'attributs	102			
Nombre de relation	N/A	42310	4240910	142680542
Noeuds isolés	N/A	9655	6687	23764
Vrai positif	233			
Faux positif	123			
Vrai positif - entrainement	86			
Vrai positif - test	37			
Faux positif - entrainement	152			
Faux positif - test	81			
F1 score	0,163	0.16	0,125	0,09
Précision	0,30	0.31	0,27	0,28
Spécificité	0,61	0.64	0,64	0,66
Rappel	0,11	0.11	0,08	0,05
ROC	0,63	0.49	0,52	0,59

Les métriques *F1-score*, précision, spécéficité, rappel et *Receiver Operating Characteristic* (ROC) sont définies dans l'annexe A

Avec chaque relation qu'on évalue, on remarque dans le tableau 4.2 que le nombre de noeuds isolés varie. Cette situation survient, car pour certains individus, les relations avec d'autres sont absentes ou n'existent pas comme pour l'exemple des consultants qui n'ont pas de gestionnaire direct ou les directeurs qui ne possèdent pas de supérieur. Parfois la donnée utilisée pour créer la relation est absente suites aux facteurs suivants :

— Erreur dans la saisie des informations

— Le réseau s'est agrandi en accueillant d'autres entreprises qui ne respectent pas forcément le schéma organisationnel de l'environement dans lequel ce travail a été réalisé.

— Erreurs survenues lors du traitement au niveau de la composante de transformation de l'architecture expliqué dans le chapitre 3 ou lors de l'acheminement des données depuis les sources.

Le nombre de relations dans le graphe évolue avec chaque choix de relation, lorsqu'on prend la relation de vice-présidence on remarque que le nombre de liens est très élevé, car les pvp regroupent de nombreux individus et cette option va créer un lien entre chaque individu.
Les résultats des différentes métriques utilisées du tableau 4.2 montrent que le modèle n'arrive pas à bien attribuer la bonne classe pour les deux classes - vrai positif / faux positif -. Sur l'ensemble des utilisateurs, le ratio des classifications est très faible. La classification tend à se rapprocher de résultats aléatoires. Il est assez difficile de les interpréter avec ce jeu de données, car le nombre des utilisateurs est important et la convolution évolue sur un grand nombre de relations.
Au niveau des ROC, tel montré sur les figures 4.8, 4.9, 4.10 et 4.11, la courbe obtenue sur le ratio de vrais positifs par rapport aux faux positifs est proche de 50% sur l'ensemble des variations du seuil de classification sur les résultats du modèle.

4.6 Scénario d'exfiltration par courriel

Continuer à travailler sur le jeu de données global avait des limites au niveau de la vérification des résultats obtenus, car le nombre d'instances est très élevé. En plus, dans notre cas nous ne pouvons pas réaliser des évaluations par rapport à d'autres démarches comportant d'autres modèles. Le scénario relatif aux alertes sur les courriels est très souvent observé par les analystes et la majorité des incidents sont liés à une activité comportant le courriel. Ce scénario exploite seulement les indicateurs développés dans le besoin d'observer les évènements possibles d'exfiltration par courriel.
Le jeu de données résultant étant étiqueté a permis d'utiliser les modèles d'apprentissage supervisés décrits sur l'annexe A à savoir la régression logistique, les forêts d'arbres aléatoires et le perceptron multicouche comme méthodes de référence pour évaluer notre approche en GCN comme l'approche de comparaison suivie par*Jiang et al.* [45]. Le tableau 4.3 détaille les caractéristiques du jeu de données de ce scénario. Nous avons reproduit les mêmes expériences conduites précédemment, sauf que la création du graphe basé sur la relation de manager ne peut être retenue, car les managers ne figurent pas dans le jeu de données.
Le tableau 4.4 montre la configuration de chaque graphe dans notre scénario d'exfiltration par courriel. Le meilleur résultat obtenu est avec la modélisation de graphe sans relation qui se rapproche des résultats des modèles supervisés. Lors de l'utilisation du graphe sans relation aucun calcul lié à la convolution n'est effectué et ceci explique le même résultat obtenu avec le perceptron multicouches au niveau du ROC dans la figure 4.15. Sans convolution les modèles GCN et le perceptron multicouches sont équivalents, ils évoluent sur les attributs d'origine pour le réseau de neurones et non les attributs mis à jour par la convolution.
Les résultats obtenus pour le GCN sur les graphes avec les relations montrent que le modèle

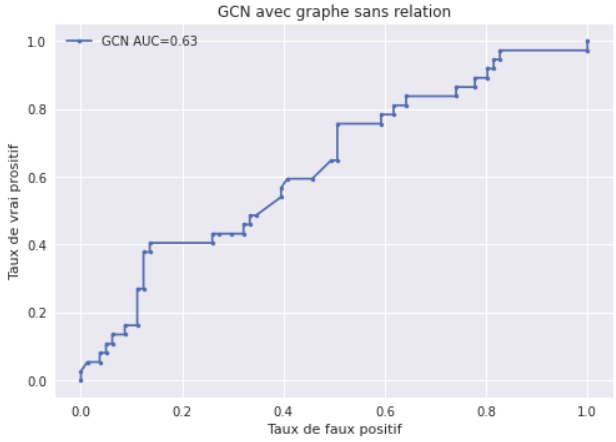

FIGURE 4.8 Courbe ROC - GCN sans relation

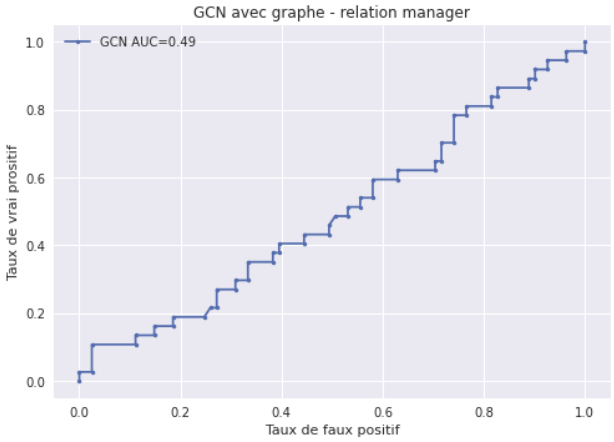

FIGURE 4.9 Courbe ROC - GCN avec relation manager

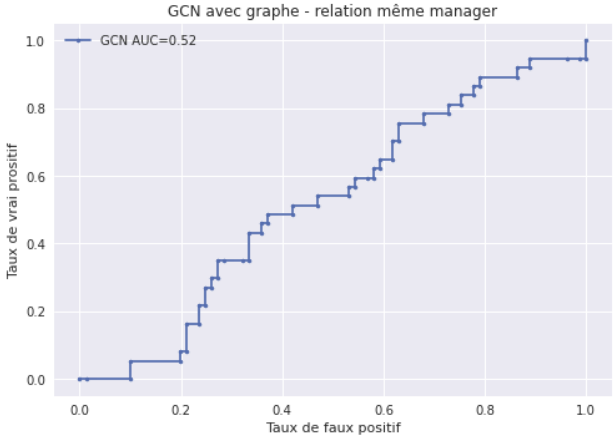

FIGURE 4.10 Courbe ROC - GCN avec relation même manager

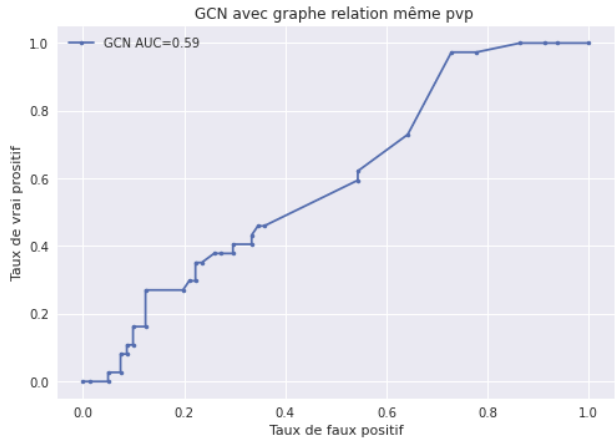

FIGURE 4.11 Courbe ROC - GCN avec relation même pvp

est impacté par la convolution. L'ensemble des métriques décroissent avec l'augmentation du nombre de relations et le modèle a du mal à bien faire la classification.

4.6.1 Marche aléatoire avec Node2vec

Les précédentes expériences démontrent que le choix de relation a un impact sur les capacités d'apprentissage du GCN et les résultats obtenus. En effet, pour chaque choix de relation, nous nous retrouvons avec des noeuds isolés. Au début, nous avons pensé à intégrer des relations disjointes pour créer plus de connectivité au niveau du graphe. Par exemple, nous avons choisi de mettre la relation de manager en parallèle avec la relation de consultant qui vient pour créer des liens entre les consultants. Cependant, les relations que nous avions à ce

Tableau 4.3 Attributs pour le jeu de données exfiltration par courriel

Feature	Description
Indicateur 001	Courriels compressés envoyés en dehors du réseau.
Indicateur 002	Courriel envoyé à l'externe avec attachement à un destinataire qui semble correspondre à son adresse personnelle.
Indicateur 003	Courriel envoyé à l'externe avec attachement à un destinataire qui semble correspondre l'adresse personnelle détecté par une solution de sécurité.
Indicateur 004	Courriel envoyé à un domaine considéré comme domaine à risque.
Indicateur 005	Courriel envoyé à l'externe avec un fichier sensible.
Indicateur 006	Courriel envoyé à l'externe avec un fichier sensible bien particulier.

Tableau 4.4 Détails et résultats des diffétentes expériences - Scénario d'exfiltration par courriel

	GCN			**Modèles supervisés**		
	Sans relation	Relation même manager	Relation même pvp	Forêts aléatoires	Regression logistique	Perceptron Multicouches
Nombre de noeuds	413					
Nombre d'attributs	6					
Nombre de relation	N/A	749	12907	N/A		
Noeuds isolés	N/A	251	6	N/A		
Vrai positif	179					
Faux positif	234					
Vrai positif - entrainement	124					
Vrai positif - test	55					
Faux positif - entrainement	165					
Faux positif - test	69					
F1 score	0,51	0,41	0,06	0.52	0,59	0,68
Précision	0,77	0,63	0.5	0,78	0,89	0,79
Spécificité	0,68	0,72	0,55	0,68	0,73	0,75
Rappel	0,38	0,29	0,03	0,39	0,44	0,60
ROC	0,73	0,64	0,50	0,78	0,787	0,786

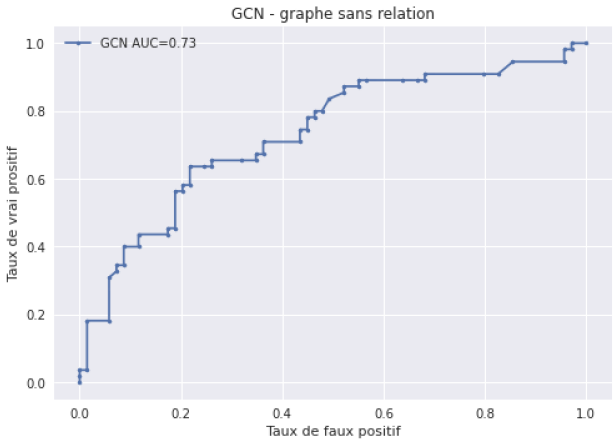

FIGURE 4.12 Courbe ROC - GCN sans relation

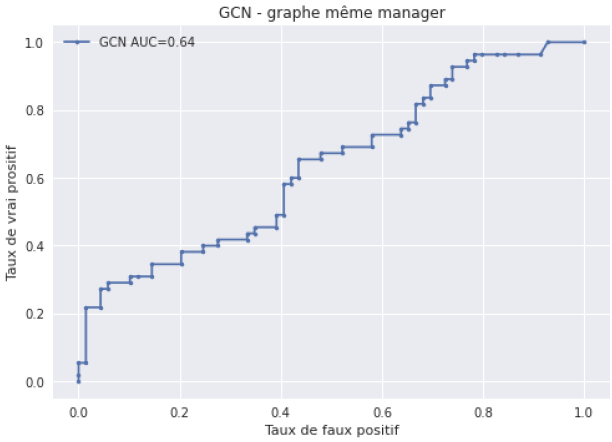

FIGURE 4.13 Courbe ROC - GCN avec relation même manager

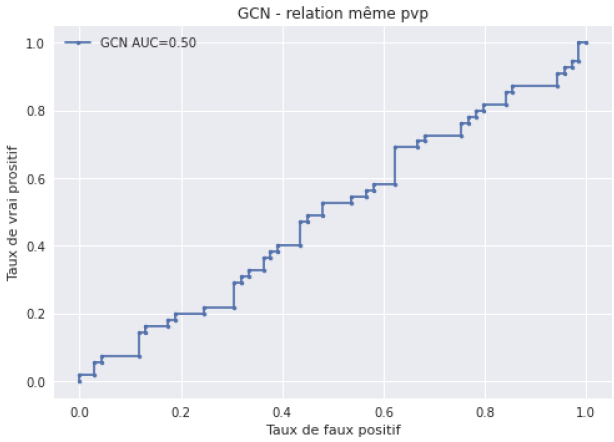

FIGURE 4.14 Courbe ROC - GCN avec relation même pvp

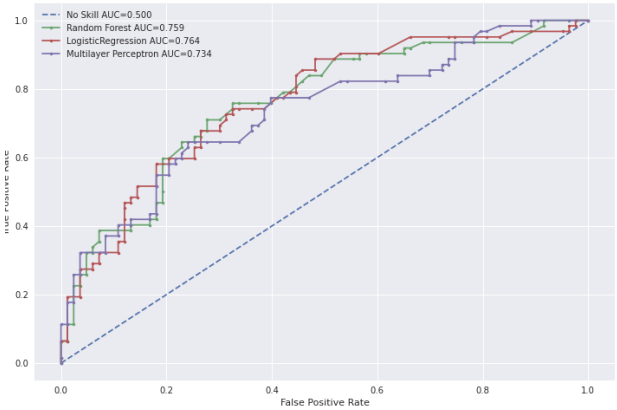

FIGURE 4.15 Courbe ROC - modèles apprentissage supervisés

moment n'exprimaient pas le niveau de liaison entre les individus dans l'entreprise. En effet, la liaison qui lie les consultants est équivalente à celle qui lie les utilisateurs dans la relation de manager. Un autre exemple si on choisissait la relation de manager et celle de la première vice présidence PVP. Les individus qui ont le même manager auront deux liaisons et ceux qui sont dans la même pvp auront une seule liaison, mais aucun indicateur nous permet de dire qu'une relation est plus forte qu'une autre ou que deux individus sont plus proches.
Pour tenter de produire un graphe plus représentatif des relations organisationnelles, nous avons utilisé des données contextuelles pour affiner les relations entre individus. Nous avons extrait depuis les données de l'annuaire les données relatives à la position de l'utilisateur dans la compagnie. Puis nous avons exploité la technique des marches aléatoires avec Node2vec [70]

pour construire les vecteurs qui représentent les noeuds du graphe. Le choix de node2vec est basé sur les performances obtenues de ce dernier sur les différents travaux précédents [70] à propos de la transformation de noeud de graphe vers un vecteur *node embedding*. *Dehghan-kooshkghazi et al.* se sont focalisés sur node2vec, car il présente de meilleures performances sur les graphes réels et synthétiques avec des résultats stables [71].
La similarité cosinus exploite alors les vecteurs produits pour les usagers pour déterminer si les vecteurs pointent vers la même direction [72].

$$sim(X,Y) = \frac{X.Y}{\|X\| \, \|Y\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}} \tag{4.12}$$

où :

X et Y : les vecteurs correspondant respectivement au noeud X et au noeud Y

$\|X\|$: la norme euclidienne du vecteur $X = (x_1, x_2, ..., x_n)$

Ainsi le score de similarité obtenu permet de passer d'une matrice d'adjacence binaire (1 s'il y a une relation, 0 sinon) à une matrice d'adjacence avec le score de similarité comme poids de la relation.
En résumé, on crée un second graphe $G_2 = (V, E)$ qui intègre le schéma organisationnel de l'entreprise. Les noeuds seront hétérogènes incluant les employés, l'unité organisationnelle, la direction, le type d'emploi (employé interne ou consultant), le partenaire ....
Dans la figure 4.16, les relations relatives à quatre employés sont illustrées. U1,U5,U33 sont des employés réguliers. U1 et U5 font partie de la direction 1 et de l'unité organisationnelle 1. U33 de l'unité organisationnelle 3 et de la direction 3, le consultant U12 réalise un mandat à la direction 3 et il travaille pour le partenaire 2.

La figure 4.17 montre un exemple du résultat qu'on estime obtenir. Dans le graphe organisationnel les employés U2 et U1 sont des conseillers et ont le même supérieur U4, ces derniers font partie de l'unité 1 et de la direction 1. Ces relations se traduisent alors par une similarité de 0.99 entre U1 et U2 et 0.90 entre U1, U4 /U2, U4. Puisque U3 partage seulement la relation de direction, sa similarité avec les autres est 0.20.
Node2vec exploite la technique de marche aléatoire pour créer des séquences pour chaque noeud à partir des noeuds qui lui sont accessibles dans le graphe. La fonction qui s'occupe de faire le mappage entre le noeud vers sa représentation vectorielle de dimension d est défini par, $f : V \to \mathbb{R}^d$.
Étant donné un noeud source $u \in V$, on simule des parcours aléatoires de longueur fixe. Les noeuds à visiter sont générés par la distribution sur l'équation 4.13.

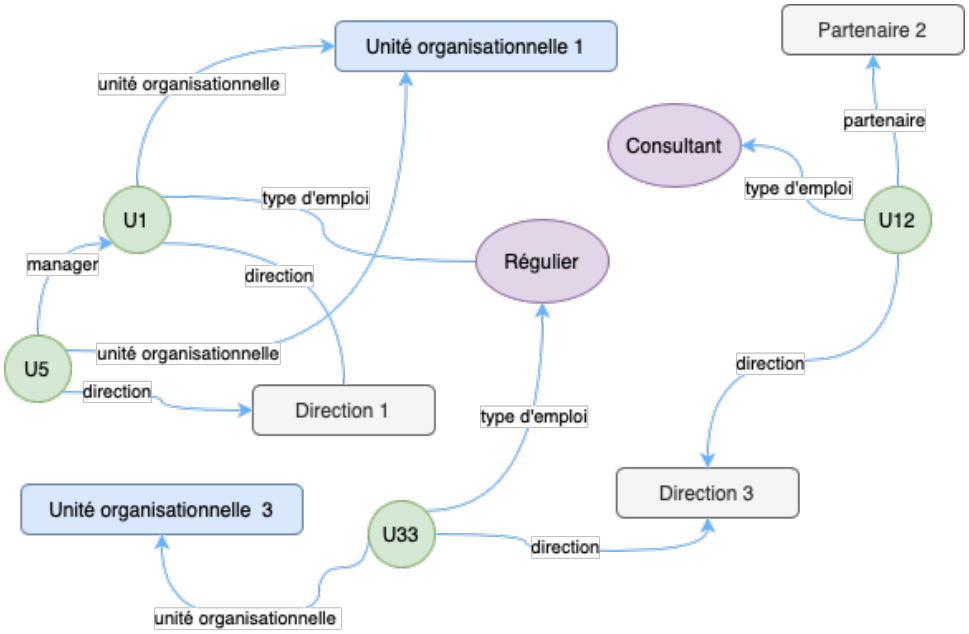

FIGURE 4.16 Abstraction du graphe organisationnel

$$P(c_i = x \mid c_{i-1} = v) = \begin{cases} \frac{\pi_{vx}}{Z} & Si\ (v,x) \in E \\ 0 & sinon \end{cases} \tag{4.13}$$

où :

c_i : Le $i^{ème}$ noeud dans le parcours, c_0 est le noeud de départ u

$\frac{\pi_{vx}}{Z}$: La probabilité de transition non normalisée entre les noeuds v et x, Z est une constante de normalisation. Afin d'ajouter un biais sur les marches aléatoires, les deux paramètres, p et q sont définis pour orienter le prochain noeud à visiter. Supposons que le parcours a traversé l'arrête (t, v) dans la figure 4.18 et que maintenant on est sur le noeud v. L'algorithme doit décider de la prochaine étape, alors il évalue la probabilité de transition π_{vx} sur la liaison (v, x) originaire de v.

$$\pi_{vx} = \alpha_{p,q}(t,x) \cdot w_{vx}. \tag{4.14}$$

$$\alpha(t,x) = \begin{cases} \frac{1}{p} & Si\ d_{tx} = 0 \\ 1 & Si\ d_{tx} = 1 \\ \frac{1}{q} & Si\ d_{tx} = 2 \end{cases} \tag{4.15}$$

Dans le cas où la relation (v, x) n'a pas de poids, $w_{vx} = 1$ sinon w_{vx} prend le poids de la

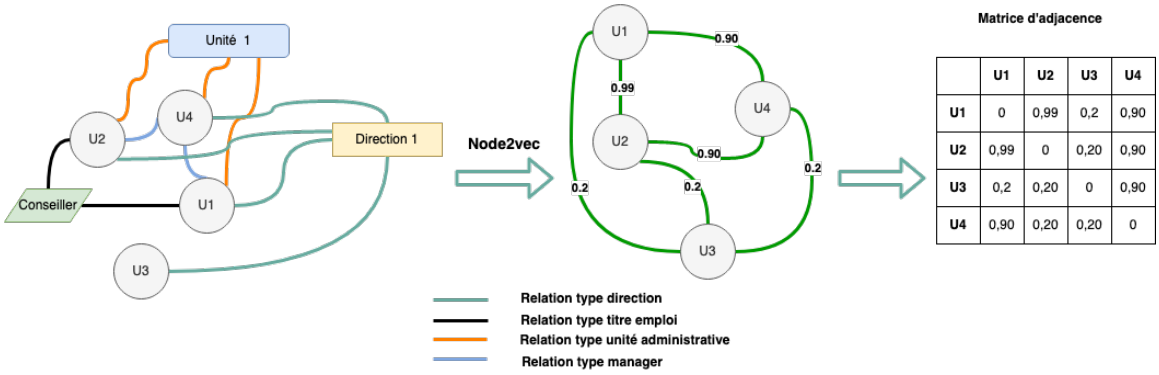

FIGURE 4.17 Abstraction du graphe organisationnel

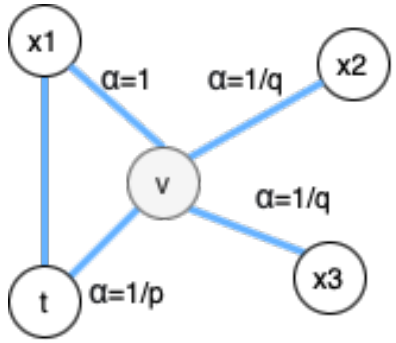

FIGURE 4.18 Transition dans node2vec

relation. d_{tx} désigne le plus court chemin entre t et x. La fonction f est une matrice de taille $|V| \times d$. Le réseau de voisins $N_S(u) \subset V$ pour chaque noeud source $u \in V$, est défini par une stratégie d'échantillonnage S. L'objectif est d'optimiser la fonction qui maximise la probabilité d'observer le réseau des voisins $N_S(u)$ pour un noeud u sachant la représentation vectorielle du noeud u.

$$\max_f \sum_{u \in V} log\ Pr(N_S(u)|f(u)). \tag{4.16}$$

$$Pr(N_S(u)|f(u)) = \prod_{n_i \in N_S(u)} Pr(ni|f(u)). \tag{4.17}$$

$$Pr(ni|f(u)) = \frac{e^{(f(n_i)\cdot f(u))}}{\sum_{v \in V} e^{(f(v)\cdot f(u))}}. \tag{4.18}$$

Node2vec utilise le modèle **skip-gram** [73, 74] pour transformer les séquences obtenues par les marches aléatoires en des vecteurs. Cette technique a été développée pour le traitement de langage naturel qui s'attend à recevoir une suite linaire de séquences de mots qui est tout simplement la nature du texte. Cependant, le graphe ne respecte pas cette nature et le

processus de marche aléatoire présenté en haut nous permet d'avoir les séquences en noeud qui ressemble à celui en mot. À partir du graphe 4.20 dont les informations sont présentées dans le tableau 4.5 nous avons conduit l'algorithme node2vec.

Tableau 4.5 Statistiques pour node2vec

Attribut	Nombre d'entrée unique	Commentaires
Transit ou code entreprise	48	31 usagers ne possèdent pas cette information
PVP	10	101 usagers ne possèdent pas cette information
Type d'emploi	9	10 usagers ne possèdent pas cette information
Unité administrative	182	101 usagers ne possèdent pas cette information
Partenaire	4	N/A
Compagnie	75	30 usagers ne possèdent pas cette information
Manager	251	11 usagers ne possèdent pas cette information
Titre d'emploi	151	17 usagers ne possèdent pas cette information

Nous avons choisi d'utiliser l'algorithme de regroupement hiérarchique [75] pour voir si le résultat de l'approche avec node2vec permettait de regrouper les individus ayant les mêmes caractéristiques organisationnelles. La figure 4.21 montre une visualisation avec t-SNE [76] des groupes d'individus dans l'entreprise. t-SNE transforme un enregistrement avec un vecteur à grande dimension en un vecteur de deux ou trois dimensions pour permettre de projeter les données sur un plan à deux ou trois dimensions pour des buts de visualisation. Au total, à partir de nos données, nous avons obtenu sept groupes qui reflètent les employés au sein de l'organisation dans notre jeu de données.

Nous avons gardé la même configuration pour le GCN sauf que nous avons défini quatre noeuds pour la couche cachée. L'utilisation de node2vec pour construire le graphe permet d'avoir un niveau de connectivité assez élevé comme le montre le nombre de relations dans le tableau 4.6, mais les résultats n'étaient pas aussi performants, comparés à ceux obtenus avec les modèles supervisés. L'évolution de la convolution sur le graphe apporte des changements sur le vecteur d'attributs de chaque noeud. Le jeu de données d'origine est mis à jour avec chaque passage de convolution.

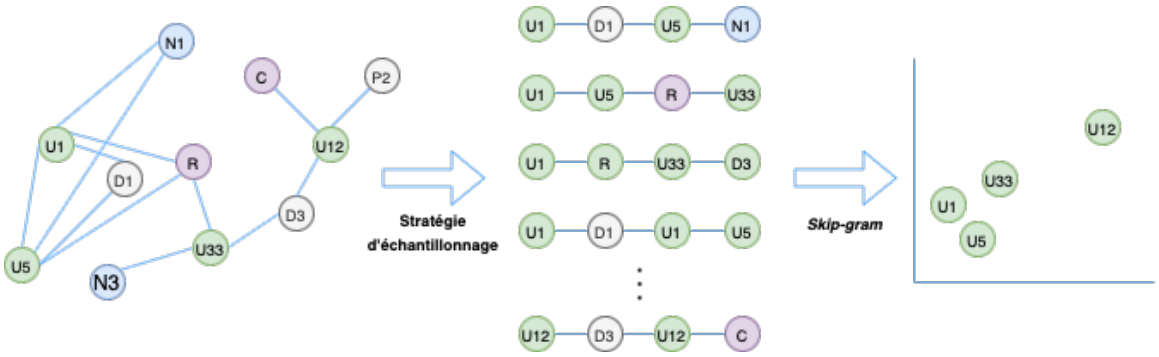

FIGURE 4.19 Processus node2vec

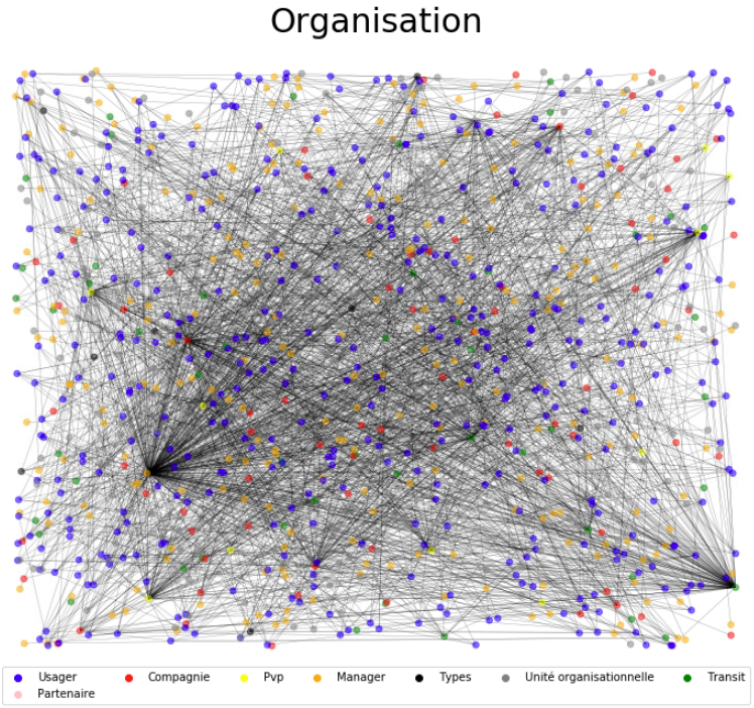

FIGURE 4.20 Extrait du graphe organisationnel en production

Nous avons voulu par la suite contrôler le niveau de connectivité et ainsi l'évolution de la convolution sur le graphe en définissant des seuils sur les scores de similarité pour garder les relations fortes. Les résultats sont mentionnés dans le tableau 4.7 : lorsqu'on augmente le seuil on remarque que le nombre de relations diminue et que le nombre de noeuds isolés augmente. Toutefois, tous les résultats obtenus pour l'ensemble des métriques étaient équivalents avec les résultats pour le graphe avec toutes les relations, mais les résultats étaient loin d'être performants que ceux obtenus grâce aux modèles supervisés.

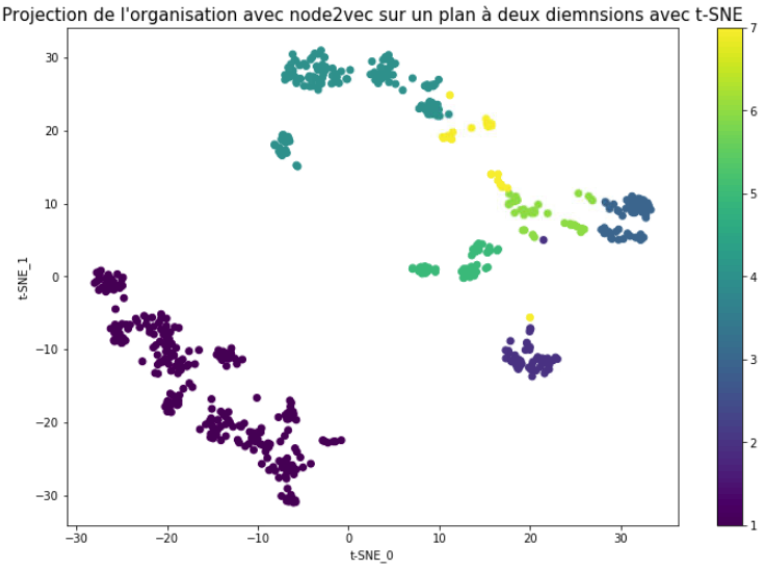

FIGURE 4.21 Projection du graphe organisationnel avec node2vec sur un plan à deux dimensions

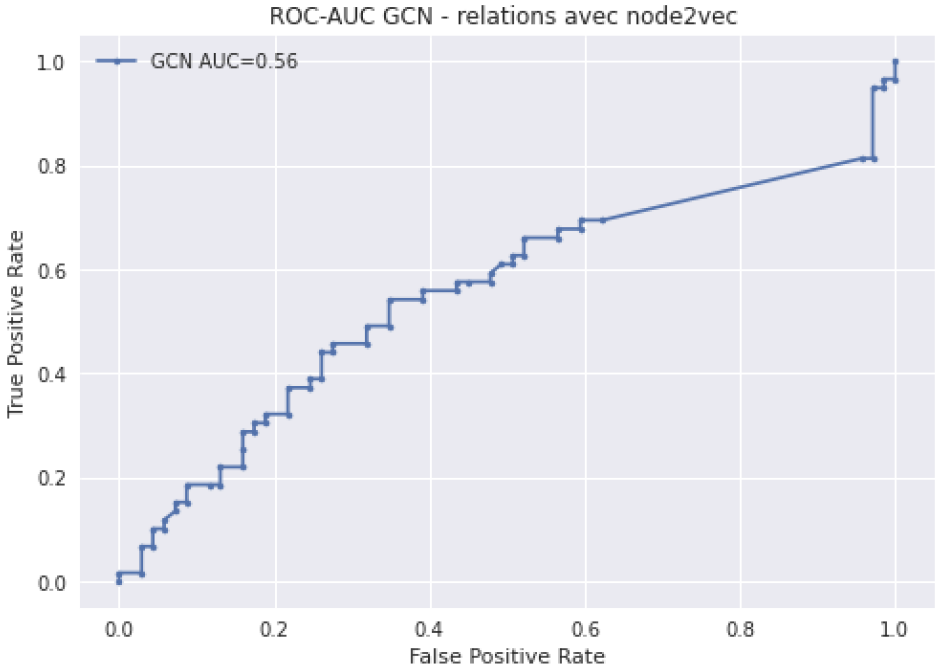

FIGURE 4.22 Courbe ROC GCN avec relations node2vec

4.6.2 Limite de modélisation pour le GCN

Une limite de conception du graphe pour l'approche en GCN a été notée. En effet, lorsqu'on observe les données, on s'aperçoit que certains utilisateurs figurent plusieurs fois dans de

TABLEAU 4.6 Détails et résultats de l'utilisation de GCN avec la matrice de node2vec

	GCN	**Modèles supervisés**		
	relation avec node2vec	Forêts aléatoires	Regression logistique	Perceptron Multicouches
Nombre de noeuds	426	N/A		
Nombre d'attributs	6	N/A		
Nombre de relation	13655	N/A		
Noeuds isolés	43	N/A		
Vrai positif	179			
Faux positif	247			
Vrai positif - entrainement	120			
Vrai positif - test	59			
Faux positif - entrainement	178			
Faux positif - test	69			
F1 score	0,03	0.52	0,59	0,68
Précision	0,50	0,78	0,89	0,79
Spécificité	0,53	0,68	0,73	0,75
Rappel	0,017	0,39	0,44	0,60
ROC	0,56	0,78	0,787	0,786

TABLEAU 4.7 Résultats des expériences avec la variation de la matrice d'adjacence

Score de similarité	**Nombre de relations**	**noeuds isolés**	**ROC AUC**	**F1-score**	**Sensibilité**	**Précision**	**Rappel**
>= 0.80	13529	60	0,46	0,17	0,56	0,66	0,10
>= 0.85	13373	68	0,40	0,21	0,54	0,53	0,13
>= 0.90	12841	83	0,09	0,53	0,5	0,53	0,05
>= 0.95	10883	111	0,60	0,09	0,53	0,5	0,05
>= 0.97	7219	172	0,50	0,30	0,53	0,5	0,22
>= 0.98	4040	239	0,50	0,30	0,53	0,5	0,22
>= 0.99	1146	334	0,54	0,06	0,53	0,5	0,03

multiples incidents avec les deux mentions (vrai positif et faux positif). Si on les maintient, les mêmes relations se retrouvaient entre les noeuds liés aux noeuds représentant ces utilisateurs.

Lors de la convolution, les noeuds liés se retrouvent à échanger de l'information à propos de leurs attributs. Cet échange se multiplie pour les noeuds qui représentent le même utilisateur, car ils héritent des relations que l'utilisateur possède avec les autres utilisateurs.
La figure 4.23 reprend cette limite de modélisation. Les deux utilisateurs U1 et U2 ont une relation dans l'entreprise. Nous avons trois instances de U1 selon les incidents IST-001, IST-002 et IST-003 (IST pour *Information security Ticket*) et deux instances de U2 dont une a été déjà rapportée dans l'incident IST-004 et nous essayons de prédire la classe de l'autre. Si on remarque bien, les instances héritent des relations des utilisateurs. Les incidents mentionnés en vert sont des faux positifs alors que l'incident en rouge correpond à un vrai positif. Les indicateurs dans le noeud U2 sans incident ressemblent à ceux déclenchés par l'usager U1 lors de l'incident IST-001 qui est un vrai positif.
Afin de simplifier les calculs lors de la convolution, nous avons choisi dans cet exemple de faire la moyenne entre les attributs reçus pendant la convolution. Lorsqu'on exécute seulement un seul passage, chacun des noeuds va recevoir de l'information des autres noeuds auxquels il est lié, ensuite la moyenne est calculée pour créer le nouveau vecteur pour chaque noeud. On remarque alors que les vecteurs pour l'ensemble des noeuds sont identiques et on s'éloigne de l'information de départ.

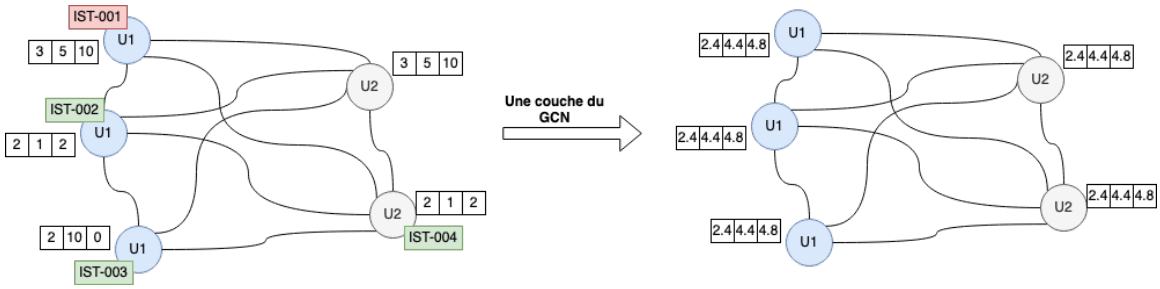

FIGURE 4.23 Limite de conception pour le GCN

On a reproduit les expériences précédentes en supprimant de manière aléatoire les instances multiples. À partir du jeu de données du scénario d'exfiltration par courriel nous avons obtenu 377 sur 413 individus soit une réduction de 9%.
Les résultats obtenus consignés dans le tableau 4.8 montrent que le nombre de relations diminue. Comparés aux résultats précédents, on obtient une amélioration pour les résultats du GCN qui intègre la relation de même pvp et les relations construites grâce à node2vec. En fait, ces deux idées pour la création de graphes permettent d'avoir un graphe assez connecté, mais les instances multiples viennent ajouter plus de relation et plus de convolutions. 2500 relations de plus pour le cas de relation de pvp et 12321 pour le cas des relations avec node2vec.

TABLEAU 4.8 Détails et résultats des diffétentes expériences - Scénario d'exfiltration par courriel - sans multiple instances

	GCN - Sans relation	**GCN - Relation même manager**	**GCN - Relation même pvp**	**GCN - Relation avec node2vec**
Nombre de noeuds	377			371
Nombre d'attributs	6			
Nombre de relation	N/A	567	10461	1334
Noeuds isolés	N/A	246	11	34
Vrai positif	178			175
Faux positif	199			196
Vrai positif - entrainement	121			126
Vrai positif - test	57			49
Faux positif - entrainement	142			133
Faux positif - test	57			63
F1 score	0,55	0,57	0,23	0,40
Précision	0,60	0,70	0.72	0,48
Spécificité	0,62	0,64	0,54	0,55
Rappel	0,49	0,49	0,14	0,34
ROC	0,66	0,67	0,66	0,55

4.6.3 Ajout du résultat de vecteurs de node2vec comme features

Au final, l'approche avec Noe2vec construit un vecteur représentatif du noeud d'une dimension qu'on définit comme paramètre. Nous avons composé un jeu de données (voir tableau 4.9) basé sur les vecteurs sortants et les indicateurs considérés pour le scénario d'exfiltration par courriel. Les variables **c1** et **c2** sont des colonnes des vecteurs de l'utilisation de node2vec. Nous avons par la suite réévalué nos modèles supervisés avec ce nouveau jeu de données. Nous nous sommes assurés de garder les mêmes ensembles d'entrainement et test en présence et non des attributs apportés par node2vec.
Les résultats mentionnés pour la métrique ROC sur les figures 4.24 et 4.25 montrent qu'on a eu une amélioration après l'ajout du contexte organisationnel pour l'ensemble des techniques.

TABLEAU 4.9 Jeu de données exfiltration par courriel avec les vecteur de node2vec

	Indicateurs courriels			Vecteur de node2vec		
Usager	**Ind_mail1**	**Ind_mail2**	...	**c1**	**c2**	...
user01	20	10	...	0.5	1.19	...
user02	2	100	...	0.46	2.01	...
user03	20	0	...	1.04	0.73	...

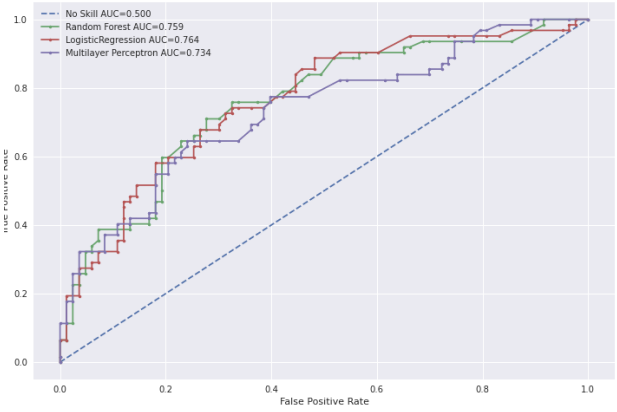

FIGURE 4.24 Courbes ROC pour les modèles sans contexte organisationnel

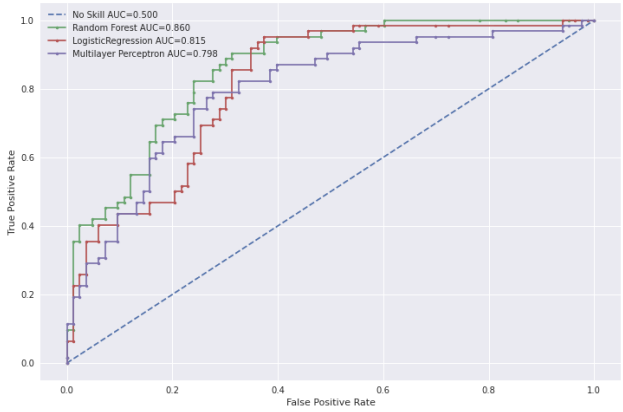

FIGURE 4.25 Courbes ROC pour les modèles avec contexte organisationnel

4.7 Intégration des résultats

Si on revient sur le processus d'investigation, nous nous rappellons que les analystes explorent les évènements pour les utilisateurs avec un haut score dans le tableau de bord. nous proposons d'ajouter une colonne pour dire que cet usager tend plus vers un faux positif ou un vrai

positif. Avec cette nouvelle colonne, l'analyste évitera d'explorer les usagers des lignes [1-4] dans le tableau 4.10 et explorera l'usager de ligne 2000.

Tableau 4.10 Exemple du tableau de bord avec l'ajout du résultat du modèle

Ligne	Utilisateur	Score	Résultat du modèle
01	user02	1200	faux positif
02	user04	1200	faux positif
03	user51	1000	faux positif
04	user17	900	faux positif
	• • •		
2000	user533	57	vrai positif

4.8 Discussion de chapitre

À travers nos expériences avec le modèle en GCN, nous avons proposé un moyen de construire un graphe d'entrée plus représentatif de l'organisation des employés et de leurs relations dans l'entreprise pour le GCN. Cependant, nous n'avons pas obtenu de meilleurs résultats comparés aux résultats des modèles supervisés pour notre jeu de données d'exfiltration par courriel. L'utilisation de la convolution permet d'explorer les attributs du voisinage, mais son fonctionnement apporte des mises à jour sur les attributs d'origine qui change complètement le jeu de données de départ, surtout en présence de la limite de modélisation présentée dans la section 4.6.2. Toutefois, l'intégration du contexte sur le jeu de données a fait évoluer les résultats pour les modèles supervisés. Le grand facteur affectant nos modèles est surtout le jeu de données, nous nous retrouverons rapidement dans les limites de l'utilisation des indicateurs comme attributs.

Le chapitre suivant présente la manière d'intégrer une ontologie afin d'ajouter de la coordination sur l'ensemble des données qu'on reçoit. Fréquemment beaucoup de traitements sont effectués avant le passage vers les modèles d'apprentissage machine et on s'est aperçu que chaque scientifique de données les gère d'une manière indépendante. La qualité des données pose un sérieux problème qui nous oriente à trouver une solution. En plus d'avoir une base solide, on a choisi d'extraire des données plus pertinentes ; on pense qu'elles seront plus efficaces pour l'entrainement des différents modèles dans le futur.

CHAPITRE 5 ONTOLOGIE POUR LA MENACE INTERNE DÉVELOPPEMENT ET UTILISATION

5.1 Introduction

L'idée présentée dans la revue de littérature sur les problèmes rencontrés lors des jointures entre les tables de données a aussi été notée. En effet, nous avons rencontré beaucoup de défis à surmonter pour pouvoir appliquer notre modèle. Plusieurs traitements sont en conséquence effectués avant de passer aux modèles d'apprentissage machine. La règle 80/20 accompagne toujours les scientifiques de données *"Most data scientists spend only 20 percent of their time on actual data analysis and 80 percent of their time finding, cleaning, and reorganizing huge amounts of data, which is an inefficient data strategy"* [77]. Les stratégies trouvées pour nettoyer les données ne sont pas normalisées et risquent d'impacter le modèle et les décisions prises suite aux résultats du modèle. Dans le contexte de menace interne, toutes les stratégies mises en place pour améliorer la posture des données doivent prendre en considération les deux points suivants :

— Les décisions suite aux résultats du modèle impactent les employés.

— Le nombre de données étiquetées est très petit comparé au jeu de données global.

Les informations de la menace interne sont gérées par quatre équipes, en l'absence d'un schéma global qui réunit l'ensemble des données, plusieurs incohérences font surface et sont parfois difficiles à gérer. Tous les modèles construits autour des données sont impactés et leurs performances influencées. L'identification de ce genre d'incohérence risque de ne pas être possible lors de la manipulation de grande quantité de données. Additionnellement, la mise en place de contrôle pour vérifier l'état des données n'est pas possible s'il n'existe pas un schéma de référence. Nous nous intéressons, non seulement à nous assurer de la conformité des données, mais de récupérer et concevoir les données pour qu'elles correspondent à notre modèle.

5.2 Plan du chapitre

Ce chapitre présente deux phases majeures qu'on aimerait ajouter au processus de détection de la menace interne dans l'espérance de l'améliorer.
La première phase consiste à concevoir et produire une ontologie pour remédier aux erreurs observées lors de la manipulation du jeu de données et des tables qui le composent et de valider les données avant utilisation. Une expérience est ainsi réalisée sur une nouvelle version du jeu

de données comportant l'expérience sur l'ajout des vecteurs de node2vec aux attributs qui a donné le meilleur résultat sur le chapitre précédent.
La deuxième phase quant à elle est orientée à produire des nouveaux jeux de données pour entrainer les différents modèles. Les attributs de ces jeux de données se baseront sur l'activité de l'investigation plutôt que sur les indicateurs. Afin d'offrir aux différentes personnes qui participent dans le programme de menace interne la possibilité d'interagir avec l'ontologie, on propose un assistant virtuel qui sera une interface entre les utilisateurs et l'ontologie. Cet assistant prend comme entrée des requêtes en langage naturel puis les transforme en requêtes sparql.
Les différentes requêtes émises permettront dans le futur d'extraire des attributs pour les nouveaux jeux de données.

5.3 Défis rencontrés lors de la manipulation de données

Dans les sous-sections suivantes, des exemples dérivés des données réelles en production sont énumérés et classifiés par niveau de défi.

5.3.1 Duplication des données

La duplication des données risque de multiplier les entrées sur le jeu de données que nous utilisons pour entrainer les modèles. Tout traitement et modèle construit autour peut être impacté et les résultats deviennent imprévisibles. L'exemple dans le tableau 5.1 montre les conséquences de l'absence de contrôles adéquats. Deux occurrences de l'usager1 figurent dans le tableau de l'annuaire. Au moment de la jointure par la colonne usager l'information sur l'incident inc1 sera dupliquée sur la table de sortie.
Le fait d'avoir des duplicatas sur le jeu de données implique que les paramètres du modèle vont être recalculés pour chaque instance. Le terme *overfitting* s'applique mieux dans ce cas, car le modèle surapprend à cause des duplicatas.

5.3.2 Sensibilité à la casse et encodage des caractères

La sensibilité à la casse apparait comme un problème simple. Il risque toutefois d'engendrer des casse-têtes sur les tables de sortie utilisées dans les traitements. Dans notre environnement et sur le jeu de données d'exfiltration par courriel, nous avons signalé une perte de 55 lignes sur 426 lors de la jointure entre les deux sources incident et annuaire.
L'exemple illustré dans le tableau 5.2 montre la perte de l'information concernant l'incident inc1 lors de la jointure entre les deux tables, car l'usager1 est saisi en majuscule sur la table de l'annuaire.

Tableau 5.1 Illustration de l'exemple de duplication de données

(a) Table des incidents

incident	**usager**	**ind1**	**ind2**	**...**
inc1	usager1	40	20	...
inc2	usager2	20	10	...

(b) Table de l'annuaire

usager	**manager**	**unité**	**direction**	**...**
usager1	manager1	3	finance	...
usager1	manager1	3	finance	...
usager2	manager3	6	marketing	...

(c) Table sortie de la jointure

incident	**usager**	**ind1**	**ind 2**	**...**	**manager**	**unité**	**direction**	**...**
inc1	usager1	40	20	...	manager1	3	finance	...
inc1	usager1	40	20	...	manager1	3	finance	...
inc2	usager2	20	10	...	manager3	6	marketing	...

L'encodage des caractères pose des problèmes similaires à celui de la casse et peut engendrer des résultats de recherche incohérents ou incomplets. Les caractères avec accent (é,è,à) ou tout autre caractère d'une langue particulière doivent être normalisés dans le jeu de données considéré avant toute utilisation.
La conséquence dans cet exemple sur le modèle est le fait qu'on va perdre les instances étiquetées.

5.3.3 Duplication des clés de jointure

Cette erreur survient lorsque la colonne avec laquelle on lance la jointure contient des duplicatas. Ce problème peut être le résultat de deux hypothèses. La première est que l'usager a occupé plusieurs positions au sein de l'entreprise à des périodes différentes. À la différence de l'exemple de l'utilisateur dans le tableau reftab :jointure, ce dernier est dupliqué. La deuxième est peut être une conséquence d'une jointure ayant le problème discuté dans la section 5.3.1 ou tout autre problème non observé. L'usager1 dans le tableau 5.3 apparait sur l'annuaire avec deux entrées, car il a occupé des positions différentes dont la plus récente est celle à la direction de finance. Lors de la jointure selon la colonne considérée, l'information de l'incident sera dupliquée dans la table de sortie.
La même conséquence sur le jeu de données de l'entrainement s'applique pour ce défi que celui de la duplication des données.

Tableau 5.2 Illustration de l'exemple de la sensibilité à la casse

(a) Table des incidents

incident	**usager**	**ind1**	**ind2**	**...**
inc1	usager1	40	20	...
inc2	usager2	20	10	...

(b) Table de l'annuaire

usager	**manager**	**unité**	**direction**	**...**
USAGER1	manager1	3	finance	...
usager2	manager3	6	marketing	...

(c) Table sortie de la jointure

incident	**usager**	**ind1**	**ind 2**	**...**	**manager**	**unité**	**direction**	**...**
inc2	usager2	20	10	...	manager3	6	marketing	...

5.3.4 Absence des données

L'absence de données est le plus grand problème qu'on a rencontré pour lequel nous n'avions pas réellement trouvé une solution concrète. Alors que pour les autres défis, nous sommes arrivé à présenter des solutions de traitement pour corriger les erreurs comme celles présentées dans les sections précédentes, pour l'absence de données notre champ de manipulation était quasiment nul.

Le choix des données du modèle ou le modèle influence nettement les données nécessaires. Par exemple les modèles de régression logistique, des forêts d'arbres aléatoires ou le perceptron multicouche ne prenaient pas en considération l'aspect d'organisation et le manque d'information de ce côté-ci n'avait aucune incidence, mais le modèle en GCN se base sur le graphe construit avec les relations qu'on retrouve dans l'entreprise. Plusieurs raisons sont à l'origine de cette situation. Le réseau dans lequel le travail est effectué a subi des changements majeurs au niveau organisation. En effet, l'entreprise s'est vue s'agrandir en accueillant d'autres entreprises dans son réseau. Chaque entreprise possédait son système informatique qui n'adhérait pas nécessairement aux mêmes conditions. Le résultat de cette transition crée un environnement avec des données diverses.

En plus de ces cas plus généraux, les équipes participantes dans le contexte de menace interne ont parfois des soucis de manipulation de données. Si on revient sur la manière de la construction de la table des incidents, on remarque que l'information sur l'usager est extraite depuis le journal de l'incident qui est sous format textuel qui en partie donne lieu aux problèmes cités plus haut de ce chapitre. Les champs texte sur la plateforme de gestion

Tableau 5.3 Illustration de l'exemple de duplication de clé de jointure

(a) Table des incidents

incident	**usager**	**ind1**	**ind2**	**...**
inc1	usager1	40	20	...
inc2	usager2	20	10	...

(b) Table de l'annuaire

usager	**manager**	**direction**	**début**	**...**
usager1	manager1	finance	10-01-2021	...
usager1	manager2	comptabilité	20-06-2017	...
usager2	manager3	marketing	03-04-2019	...

(c) Table sortie de la jointure

incident	**usager**	**ind1**	**ind 2**	**...**	**manager**	**direction**	**début**	**...**
inc1	usager1	40	20	...	manager1	finance	10-01-2021	...
inc1	usager1	40	20	...	manager2	comptabilité	20-06-2017	...
inc2	usager2	20	10	...	manager3	marketing	03-04-2019	...

des incidents sont libres de contraintes de validation. Chaque analyste se trouve alors libre de les utiliser de la meilleure manière qu'il juge importante pour la résolution de l'incident, mais pas nécessairement pour l'exploitation ou la consommation des informations utilisées par des modèles automatiques ou d'autres équipes.

Tableau 5.4 Exemple de journal pour trois incidents

Incident	Analyste	Titre	Résumé
Inc1	Analyste 1	Envoi d'un document confidentiel à une boite personnelle	l'usager **usager1@domain** a tranféré un document confidentiel à sa boitre personnelle.
Inc2	Analyste 2	(**usager2**) Envoi d'un document confidentiel à une boite personnelle	l'usager **usager2** a tranféré un document confidentiel à sa boitre personnelle.
Inc3	Analyste 1	Envoi d'un document confidentiel à une boite personnelle	l'usager **user1** a tranféré un document confidentiel à sa boitre personnelle.

Le tableau 5.4 donne un extrait des situations rencontrées pour un même type d'incident. Sur le premier incident Inc1, l'analyste a identifié l'utilisateur dans le résumé de l'incident avec la mention **usager1@domain**, alors que l'analyste 2 dans l'incident Inc2 a identifié

l'utilisateur sur le titre entre parenthèses et a fait mention du nom de l'utilisateur **usager2** dans le résumé. L'analyste 1 revient sur l'incident Inc3 avec un autre format de réponse, sur cet incident, l'utilisateur est toujours mentionné dans le résumé, mais avec un autre format d'identification à savoir **user1**. Du côté de la gestion des incidents, les informations sont valides, mais du côté du consommateur des données, chaque incident qui se présente doit être traité.

Tableau 5.5 Statistiques de classe de défis sur la jointure entre la table incident et annuaire

Classe de problème	**Nombre d'occurence observé**
Duplication des données	Huit occurences sur la table de l'annuaire
Sensibilité à la casse et encodage des caractères	55 occurences
Duplication sur la clé de jointure	5923 sur la table de l'annuaire
Absence des données	17 occurences

Les tables utilisées sont conformes et respectent leur propre schéma, mais les résultats après la jointure ne le sont pas forcément.

5.4 Ontologie pour résoudre les problèmes

L'écart de conceptualisation entre les requêtes des personnes participant dans le projet de protection contre la menace interne et les différentes tables donne naissance à de multiples erreurs. Afin de résoudre les défis cités plus haut et tout défi qui peut se présenter suite à une incompréhension entre les experts et les données, on propose de construire une ontologie de domaine qui servira par la suite à formaliser le vocabulaire et imposer certaines conditions sur les données pour assurer leur fiabilité avant utilisation. La méthode *Pay-as-you-go* présentée dans [78] correspond à notre état d'avancement et le besoin actuel, cette méthodologie consiste à transformer les tables de base de données en une ontologie. Les auteurs ont choisi d'exprimer les notions de l'ontologie avec d'autres termes pour communiquer plus aisément. Comme eux nous utiliserons le terme Concept pour parler de classe, Attribut pour remplacer la propriété de la donnée, et finalement Relation pour le prédicat.

5.4.1 La méthode *Pay-as-you-go*

La méthodologie *Pay-as-you-go* est composée de trois phases principales. La première sert à capturer la connaissance sur les questions que les experts posent et les données nécessaires pour y répondre. La connaissance acquise se transforme alors en un rapport. La deuxième phase sert à implémenter l'ontologie et faire la correspondance depuis le rapport de la phase

une. L'expert du domaine dans la phase trois interroge directement l'ontologie sans communiquer avec d'autres équipes.

5.4.2 Cas d'utilisation - exfiltration par courriel

Ce cas d'utilisation a été étudié dans le chapitre 4 pour implémenter un modèle en apprentissage machine afin d'aider les analystes à identifier rapidement ce scénario. Nous l'explorons ici également puisque nous possédons le jeu de données réelles correspondant.
Une certaine complexité existe dans ce scénario lors de la manipulation. Les analystes tentent de trouver les utilisateurs concernés par les mauvaises pratiques.

Phase 1 - capture de la connaissance

Dans la première phase, on pose les questions présentés sur le tableau 5.6 en lien avec le problème ciblé pour bien le délimiter et identifier les sources d'informations nécessaires pour y répondre.

Tableau 5.6 Questions de la phase 1

Quoi :	Qui sont les usagers impliqués dans les incidents liés à l'exfiltration par courriel.
Pourquoi :	Parfois l'information sur l'usager n'est pas complète et pose un problème lors de l'investigation
Qui :	Les analystes dans l'équipe opération consomment l'information sur le déclenchement des indicateurs directement et les scientifiques de données les utilisent pour développer des modèles d'apprentissage machine.
Comment :	Les analystes et les scientifiques de données demandent aux ingénieurs de l'équipe plateforme de leur mettre à disposition les informations sur les indicateurs, les individus concernés par les incidents qui ont activé ces indicateurs et les incidents en question.
Quand :	Les données sont consommées automatiquement en temps réel et sont poussées vers le tableau de bord des analystes de l'équipe opération et pendant l'entrainement des modèles.

Les concepts

Dans cette section nous décrivons les classes de notre ontologie. Certains concepts ont été pris directement depuis l'ontologie proposée par le CERT pour la menace interne [33], qui est une ontologie solide incluant plusieurs concepts qui viennent d'incidents réels. En plus nous retrouvons souvent ces concepts dans les entreprises. Nous nous sommes basé sur cette ontologie.

Les concepts repris depuis l'ontologie CERT sont :

Tableau 5.7 Concepts repris dapuis l'ontologie CERT pour la menace interne

Action	Cette classe et ses sous-classes définissent l'action faite par un individu.
Action digitale	C'est est une sous-classe de la classe Action. Elle regroupe tous les évènements qui se sont produits dans l'environnement digital de l'entreprise.
Action-email	Sous-classe de Action, elle représente les évènements relatifs aux courriels.
Actif	Cette classe regroupe les actifs de l'entreprise.
Actif digital	Cette classe est une sous-classe de la classe Actif. Elle regroupe les actifs du réseau de l'entreprise comme les serveurs, les postes de travail, les solutions de sécurité comme pare-feu, les DLP, etc.

Nous avons créé plusieurs rapports de connaissance pour les concepts de notre ontologie, dont deux sont illustrés dans ce chapitre dans les tableaux 5.8 et 5.9 et les autres dans l'annexe B. La figure 5.1 donne un aperçu des différentes classes nécessaires pour modéliser notre scénario d'exfiltration par courriels.

Tableau 5.8 Rapport de connaissance - Concept indicateur

Nom du concept	Indicateur
id du concept	indicateur
Nom de la table ou requête SQL	select ind from table_indicateur

Tableau 5.9 Rapport de connaissance - Concept Incident

Nom du concept	Incident
id du concept	incident
Nom de la table ou requête SQL	select inc from incident

Les attributs

Chaque instance d'une classe possède des attributs. Le tableau 5.10 donne un aperçu sur le rapport de connaissance des attributs de la classe Action-email, chaque instance de cette classe doit inclure les attributs définis, pour chaque classe, on a défini un rapport de connaissance pour ses attributs et ils sont disponibles dans l'annexe B. La figure 5.2 montre les attributs pour chaque classe de notre ontologie.

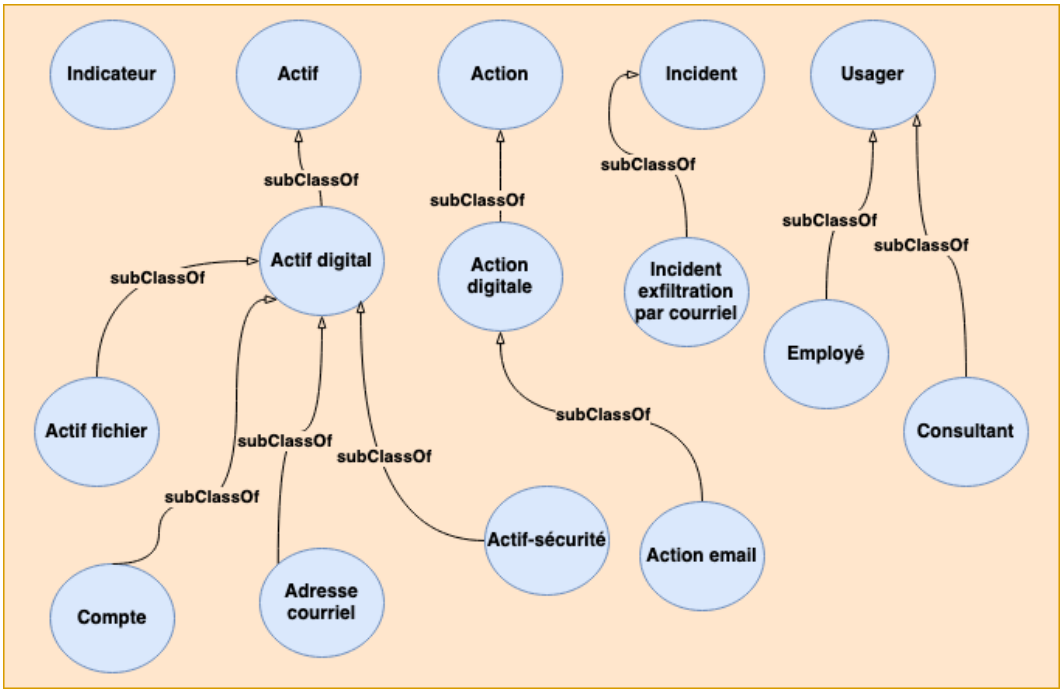

FIGURE 5.1 Classes de l'ontologie

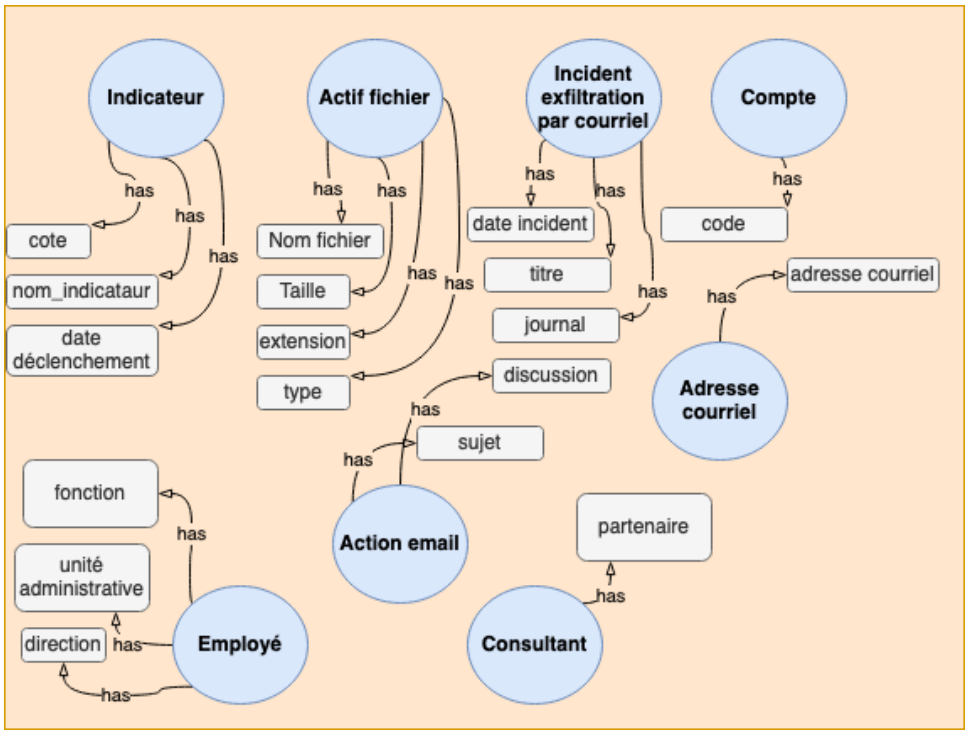

FIGURE 5.2 Attributs des classe de l'ontologie

Tableau 5.10 Rapport de connaissance - Attributs de Action-email

Nom de l'attribut	Sujet
id de l'attribut	sujet
Appliqué au concept	Action-email
Nom de la table ou requête SQL	select action-email,sujet from email_table
Nom de la colonne	sujet
Type de données	Chaine caractères
Possible de valeur nulle	Oui
Valeur si nulle	N/A
Nom de l'attribut	Discussion
id de l'attribut	discussion
Appliqué au concept	Action-email
Nom de la table ou requête SQL	select action-email,discussion from email_table
Nom de la colonne	discussion
Type de données	Chaine caractères
Possible de valeurnulle	Oui
Valeur si nulle	N/A

Relations

Les relations entre les classes sont aussi des rapports de connaissance, le tableau 5.11 donne la relation qui lie les concepts Usager et Incident. Les relations de notre ontologie sont montrées dans la figure 5.3, la suite des rapports pour les relations sont mis dans l'annexe B.

Tableau 5.11 Rapport de connaissance - Relation usager considéré dans l'incident

Nom de la relation	usager considéré
Définition de la relation	Un incident d'exfiltration concerne un usager
Id de la relation	usager_considéré
Depuis le concept	Incident
Nom de la table ou requête SQL	Incident
Vers le concept	Usager

Phase 2 - Implémentation de l'ontologie

À travers les rapports de connaissance qu'on a produits, on entame la phase de correspondance entre les éléments des rapports et les concepts de l'ontologie. Le tableau 5.12 donne la correspondance entre chaque entité et son équivalent en langage OWL. Deux des classes de notre ontologie sont présentes sur les *listings* B.1 5.1.
Nous avons introduit des exemples basés sur le schéma d'ontologie qu'on a proposé. Ces

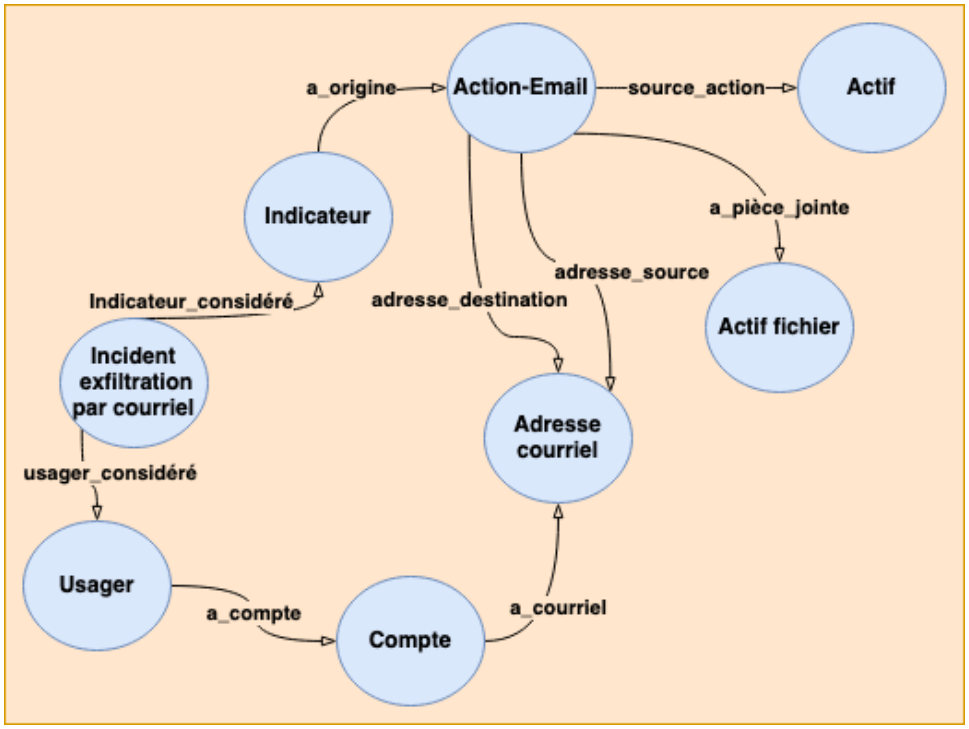

FIGURE 5.3 Relations de l'ontologie

Listing 5.1 Classe Action courriel en OWL

```
mi:Action-Email rdf:type owl:Class ; rdfs:label "Action courriel";
    rdfs:subClassOf mi:Action-digitale.
mi:sujet rdf:type owl:DatatypeProperty ; rdfs:domain mi:Action-Email;
    rdfs:range xsd:string.
mi:discussion rdf:type owl:DatatypeProperty ; rdfs:domain mi:Action-Email;
    rdfs:range xsd:string.
mi:source_action rdf:type owl:ObjectProperty; rdfs:label "origine de
    l'action" rdfs:doamin mi:Action-Email; rdfs:range mi:Actif.
mi:a_piece_jointe rdf:type owl:ObjectProperty; rdfs:label "Piece jointe"
    rdfs:doamin mi:Action-Email; rdfs:range mi:Actif-fichier.
mi:adresse_source rdf:type owl:ObjectProperty; rdfs:label "Adresse source"
    rdfs:doamin mi:Action-Email; rdfs:range mi:Adresse courriel.
mi:adresse_destination rdf:type owl:ObjectProperty; rdfs:label "Adresse
    destination" rdfs:doamin mi:Action-Email; rdfs:range mi:Actif-fichier.
```

TABLEAU 5.12 Correspondance entre les rapports de connaissance et les concepts de l'ontologie

Rapport de connaissance	**Concept OWL**	Description des concepts de OWL
Concept	owl :Class	La classe dans OWL.
Nom du concept	rdfs :label	Permet de donner un nom lisible pour les humains.
Définition du Concept	owl :comment	Fournit une desctiption de l'entité
Id du concept	/	Utilisé pour créer l'URI de la classe.
Attribut	owl :DatatypeProperty	Propriété d'une instance d'une classe.
Nom de l'attribut	rdfs :label	Permet de donner un nom lisible pour les humains.
Définition de l'attribut	rdfs :comment	Fournit une desctiption de la propiété.
Id de l'attribut	/	Utilisé pour créer l'URI de la propiété.
Depuis le concept	rdfs :domain	Le domaine permet d'avoir le type de notre source dans la relation ou le sujet
Vers le concept	rdfs :range	La portée permet d'avoir l'objet impliqué dans la relation
Relation	owl :ObjectProperty	La relation dans OWL.
Nom de la relation	rdfls :label	Permet de donner un nom lisible pour les humains.
Défintion de la relation	rdfs :commentaire	Donne des commentaires ou une description sur la relation.
Id de la relation	/	

données sont des données fictives qui reflètent des informations de la réalité. Ensuite, nous avons émis des requêtes SPARQL pour valider l'ontologie. Dans notre requête, on essaye de trouver les usagers ayant envoyé des courriels à des adresses appartenant au domaine *"pastbin"*. La figure 5.4 donne le résultat. On peut observer deux actions courriel.

— Un courriel envoyé depuis l'adresse *Pierce.Dalton@Beta.com* vers l'adresse de destination *exfiltration2@pastbin.co*. L'objet du courriel était *Protected* et ce courriel avait le fichier *file3* en pièce jointe. Cette action courriel a déclenché l'indicateur **Ind_005**.

— Un courriel envoyé depuis l'adresse *Roger.Niven@alpha.com* vers l'adresse *exfiltration@pastbin.co*. L'objet du courriel était *Project3*, cependant deux fichiers sont

envoyés en pièce jointe (*file2* et *file8*). De même, cette action courriel a déclenché l'indicateur **Ind_005** et aussi l'indicateur **Ind_007**

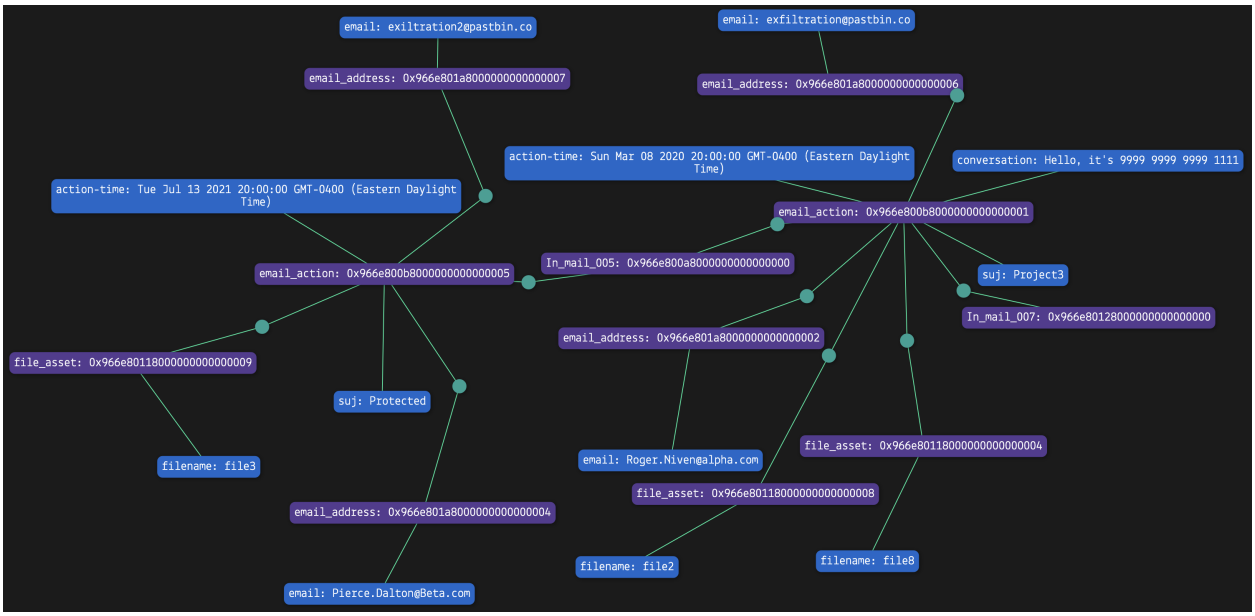

FIGURE 5.4 Résultat de la requête SPARQL dans 5.2

5.4.3 L'ontologie pour ajouter de la sémantique aux données

Puisque nous avons transformé notre ensemble de données vers une ontologie, on peut désormais proposer des contrôles sur l'état de la donnée. Par exemple, l'unicité ou la possibilité d'avoir une valeur nulle. En plus, on peut proposer des règles de logique pour répondre au défi de l'absence de l'information. Par exemple, si un usager ne possède pas l'information sur son unité administrative et sa vice-présidence, on peut les récupérer à partir des informations disponibles du côté de son superviseur, si elles sont toutefois disponibles. La raison est que l'employé travaille au sein de la même unité administrative et la pvp que son manager. Dans l'exemple exprimé par la règle logique en SWRL en 5.1, nous avons deux relations. La première est une relation de manager dans laquelle on spécifie que l'usager U4 est le manager de l'usager U1. La seconde nous informe que l'usager U4 appartient à l'unité N1. Ces deux relations impliquent que l'usager U1 appartient lui aussi à l'unité N1.

$$hasManager(?U1, ?U4) \wedge hasunit(?U4, ?N1) \implies hasunit(?U1, ?N1) \qquad (5.1)$$

Cette règle fonctionne aussi dans l'autre sens, si l'information est manquante du côté du manager telle qu'exprimé dans la section 5.2. Les deux règles subséquentes 5.3 et 5.4 sont

Listing 5.2 Trouver les personnes qui ont envoyé par courriel des fichiers au domaine pastbin

```
PREFIX mi:<http://xmlns.com/mi>
SELECT ?usager ?fichier ?action ?indicateur
WHERE {
   ?x a mi:actif_fichier.
   ?x mi:nom_fichier ?fichier.
   ?usager a mi:usager.
   ?usager mi:a_compte ?c.
   ?c a mi:compte.
   ?c mi:a_courriel ?l.
   ?l a mi:adresse_courriel.
   ?indicateur a mi:indicateur.
   ?indicateur mi:a_origine ?action.
   ?action a mi:action_email.
   ?action mi:adresse_source ?l.
   ?d a mi:adresse_courriel.
   ?action mi:adresse_destination.
   ?d mi:domain ?m.
  filter contains (?m,"pastbin").
}
```

aussi possibles pour le cas de la pvp.

$$hasManager(?U1, ?U4) \wedge hasunit(?U1, ?N1) \implies hasunit(?U4, ?N1) \tag{5.2}$$

$$hasManager(?U1, ?U4) \wedge haspvp(?U1, ?P1) \implies haspvp(?U4, ?P1) \tag{5.3}$$

$$hasManager(?U1, ?U4) \wedge haspvp(?U1, ?P1) \implies haspvp(?U4, ?P1) \tag{5.4}$$

Le tableau 5.13 montre le nombre d'instances où la donnée sur l'unité administrative est absente suivi du tableau 5.14 qui énumère le nombre d'utilisateurs ou la pvp est absente. Les deux tableaux dépendent de la position de l'utilisateur dans l'entreprise pour notre jeu de données - exfiltration par courriel qui contient 377 utilisateurs. On remarque que certains usagers internes qui travaillent en tant que régulier ou agent d'assurance ne possèdent pas l'information sur l'unité administrative et la pvp.

Tableau 5.13 Nombre d'instance dont la pvp est absente par type d'emploi

Type d'emploi	**nombre d'instance**
fournisseur	1
agents d'assurance	22
temporaire	1
partenaire	8
consultant	5
régulier	57

Tableau 5.14 Nombre d'instance dont l'unité administrative est absente par type d'emploi

Type d'emploi	**nombre d'instance**
fournisseur	1
agents d'assurance	22
temporaire	1
partenaire	8
consultant	5
régulier	56

Grâce aux règles logiques, le nombre d'instances pour lesquelles l'information est manquante a été réduit. Les deux tableaux 5.15 et 5.16 donnent le résultat final, notamment on remarque que les usagers qui travaillent en tant qu'agent d'assurance ne figurent plus parmi les usagers des tableaux précédents. Bien qu'on ait amélioré de 9% l'ensemble des données, nos résultats d'expérience n'ont pas beaucoup varié comparés aux résultats illustrés dans le chapitre 4. Nous avons déroulé les expériences avec la nouvelle version du jeu de données. Cependant, nous n'avons pas pu observer un grand changement au niveau des résultats montrés sur les figures 5.5 et 5.6.

Tableau 5.15 Nombre d'instance dont la pvp est absente par type d'emploi

Type d'emploi	**nombre d'instance**
fournisseur	1
temporaire	1
partenaire	8
consultant	5
régulier	57

TABLEAU 5.16 Nombre d'instance dont l'unité administrative est absente par type d'emploi

Type d'emploi	nombre d'instance
fournisseur	1
temporaire	1
partenaire	8
consultant	5
régulier	56

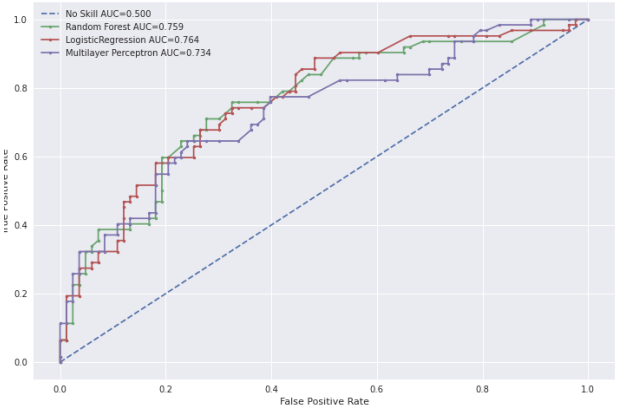

FIGURE 5.5 Courbes ROC pour les modèles sans contexte organisationnel

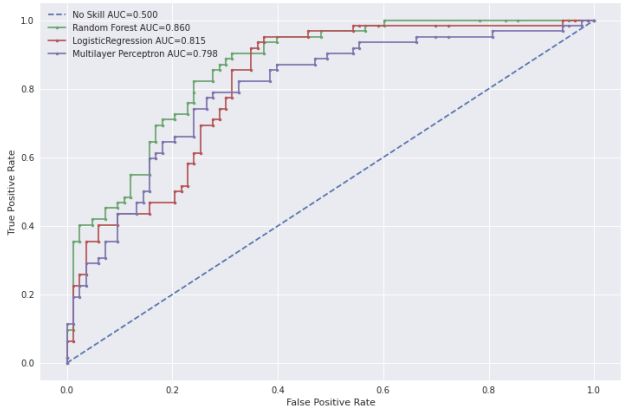

FIGURE 5.6 Courbes ROC pour les modèles avec contexte organisationnel

5.4.4 Un assistant virtuel pour l'investigation

Le processus de gestion des incidents en menace interne présenté dans le chapitre 3 n'expose pas les actions qui se déroulent durant l'investigation. En réalité ces activités sont tous les jours réalisées par les analystes, mais ne sont malheureusement pas inscrites dans un endroit où on peut les exploiter. Les jeux de données utilisés pour l'entrainement des modèles se basent sur les indicateurs qui sont des informations de haut niveau que nous ne pouvons pas manipuler facilement. Lorsqu'on utilise un indicateur comme attribut, deux choix nous sont offerts.

— La cote de risque de l'indicateur.

— Le nombre de fois que l'utilisateur a declenché l'indicateur.

Afin de créer un jeu de données plus consistent, nous avons choisi d'observer les informations de bas niveau qui sont recherchées par les analystes. Alors, nous proposons de transformer les différentes actions qui sont en réalité des requêtes manuelles ou automatiques vers une source

d'information donnée par laquelle l'analyste confirme une hypothèse en une conversation avec un assistant virtuel. L'orientation vers un assistant virtuel vient du fait qu'en plus que les usagers ne maitrisent pas SPARQL, ils ne connaissent pas toutes les entités de l'ontologie pour faire des requêtes. À la réception d'une requête, cet assistant émet une demande vers notre base de connaissances. Toutes les requêtes effectuées par l'analyste seront accessibles et nous permettent d'avoir une visibilité sur l'investigation faite à propos d'un incident. Par le biais de l'assistant, on peut récupérer les informations pertinentes pour un analyste durant l'exploration de l'incident. Ces caractéristiques seront, dans le futur, les données qu'on pousse à nos différents modèles en apprentissage machine.

La figure 5.7 illustre le processus qui nous permettra de tirer avantage de notre assistant pour construire des modèles. Nous avons toujours nos sources de données qui sont variées et qui alimentent en ce moment notre ontologie. Puisque cette ontologie a été construite avec l'idée d'un schéma commun, les données sont toujours validées avant consommation. Ensuite les analystes lanceront des requêtes vers la base de connaissances à travers l'assistant virtuel. Finalement, les scientifiques de données examinent les conversations pour extraire des éléments susceptibles d'être employés dans un modèle. Un exemple de conversation est présenté sur les figures 5.8 et 5.9 avec [79]. L'analyste commence par voir les actions courriel effectuées par l'usager *SDalton*, puis vérifie si l'une des adresses de destination a déjà figuré sur un incident précédent. L'adresse *exfiltration2@pastbin.co* n'a jamais été vue, cependant l'adresse de destination de l'action deux *exfiltration@pastbin.co* a été déjà inscrite dans un incident d'exfiltration. Comme dernière action, l'analyste vérifie si le domaine *pastbin* est considéré comme domaine à risque. À partir de cette conversation, on peut ajouter deux caractéristiques à notre jeu de données d'exfiltration par courriel. L'une pour vérifier si l'adresse de destination figurait ou non dans un incident précédent, l'autre à propos du domaine s'il correspond à *pastbin*. Plusieurs opportunités s'offrent pour avoir d'autres caractéristiques : les informations sur le fichier sont aussi importantes dont la taille par exemple, la correspondance entre le type et l'extension, car les opérateurs malveillants tentent de passer les solutions de sécurité en ajoutant une extension supplémentaire. Les informations de l'action courriel sont aussi importantes, le destinataire peut se trouver soit comme destinataire direct, en copie carbone Carbon Copy (CC) ou en copie carbone invisible Blind Carbon Copy (BCC). L'objet du courriel et la conversation sont aussi à explorer. D'autres opérations complexes sont aussi réalisables. Par exemple, l'analyste peut à travers l'assistant virtuel faire appel à un *notebook* pour exécuter des tâches particulières.

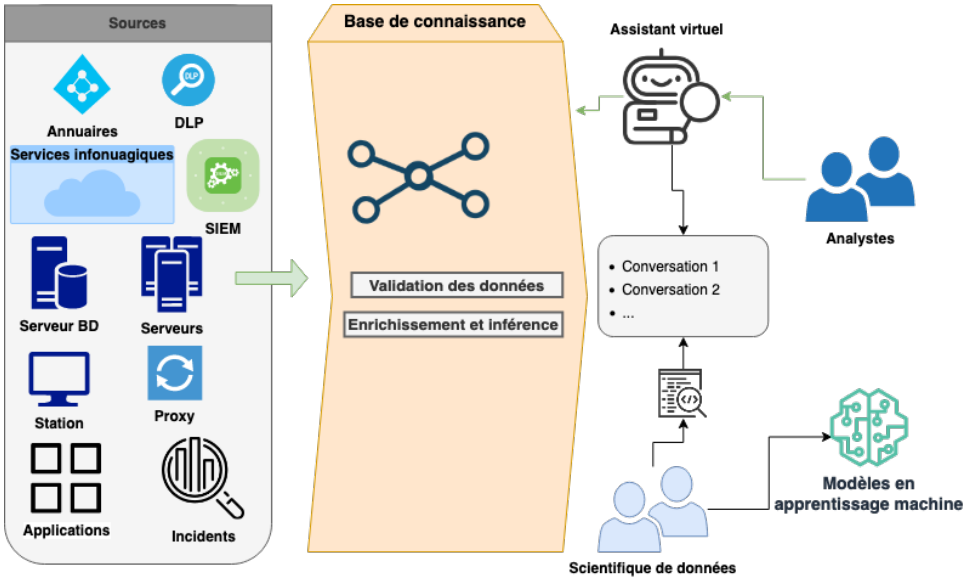

FIGURE 5.7 Intégration de l'assistant virtuel dans l'écosystème de la menace interne

5.4.5 Identification du crime organisé lié à la menace interne

Les attaques informatiques sont parfois liées à du crime organisé qui nécessitent des techniques de piratage évoluées incorporant des vulnérabilités inconnues et des utilisateurs très qualifiés. La menace interne quant à elle ne nécessite pas autant de ressources, il suffit d'un intervenant qui travaille dans une entreprise et qui souhaite faire des profits. Cela ouvre une voie rapide à l'élaboration d'un réseau d'activités illicites important. Identifier un usager comme étant malveillant dans une entreprise peut être seulement une partie visible de l'iceberg. On peut ajouter un niveau à notre base de connaissance pour unifier les informations de plusieurs acteurs dans le secteur financier pour leur permettre d'échanger de l'information. L'assistant virtuel devient un facteur solide dans une telle architecture, car sans connaitre les concepts en place les opérateurs peuvent poser des questions qui leur permettront d'avoir une visibilité plus profonde du crime en place.

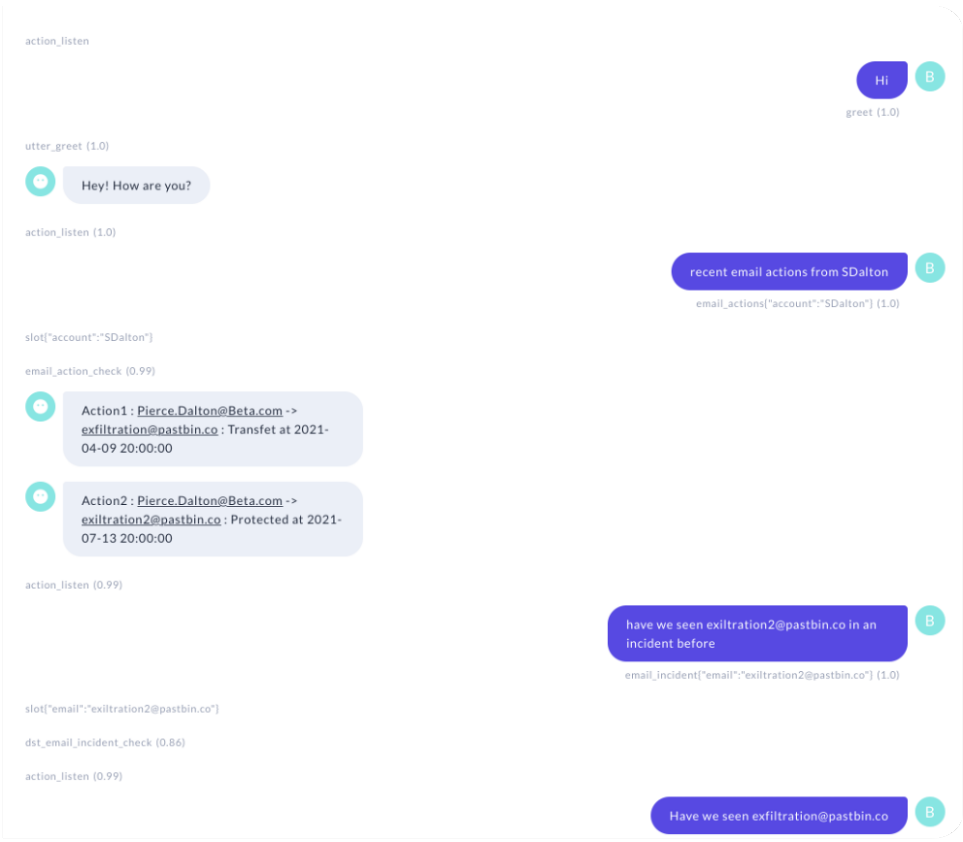

FIGURE 5.8 Conversation avec l'assistant 1

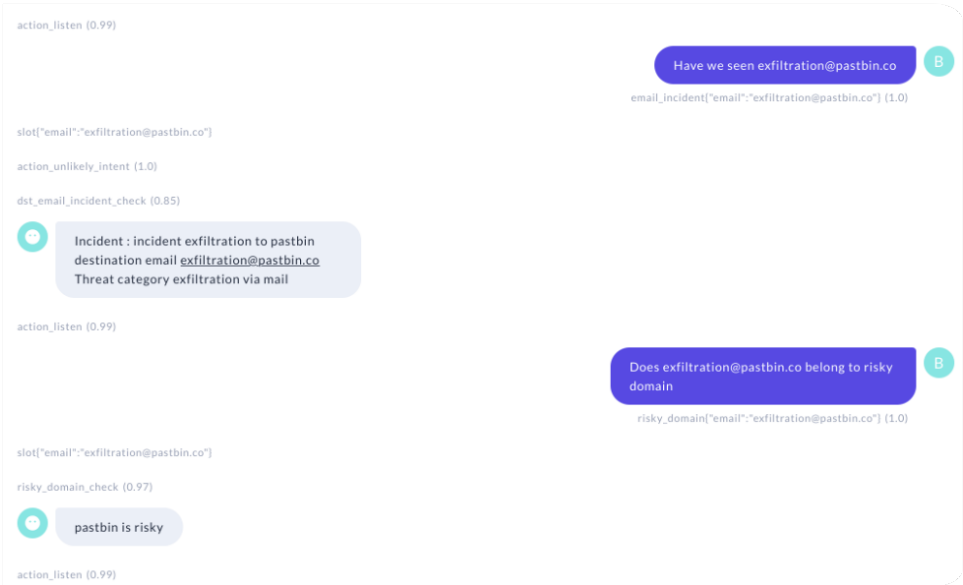

FIGURE 5.9 Conversation avce l'assistant 2

CHAPITRE 6 CONCLUSION

La menace interne représente un risque réel pour les entreprises. De plus, le nombre d'incidents dans ce contexte ne cesse de croitre au fil du temps. De nombreuses compagnies et leurs clients ont été victimes des abus réalisés par leurs employés.
Bien qu'il existe de nos jours des solutions qui permettent d'identifier les comportements suspects, les analystes ont du mal à confirmer qu'un évènement spécifique est un vrai positif ou un faux positif. Dans nos travaux comme premier objectif, nous avons étudié le programme de protection contre la menace interne dans une grande entreprise pour comprendre où la difficulté se situe réellement. Le réseau de cette entreprise a une complexité très élevée qui inclut une grande distribution de systèmes divers que les usagers manipulent. L'une des premières conséquences de cette complexité est l'énorme quantité de données qui est collectée partout sur le réseau pour analyser le comportement d'un individu. Le framework bâti inclut un tableau de bord qui aide les analystes à pointer les individus les plus dangereux. Néanmoins et maintes fois, des usagers avec des actions normales se retrouvent souvent dans le haut de la liste des personnes à surveiller. Notre approche avait pour but de pousser les utilisateurs malveillants ou ayant suivies des mauvaises pratiques vers le haut en profitant des anciens incidents et d'autre part, d'exploiter les relations dans l'entreprise pour éviter que les personnes proches se retrouvent être sous investigation si l'un d'eux a été déjà analysé pour les mêmes actes. Durant la réalisation de ce deuxième objectif, nous avons accompli plusieurs expérimentations principalement avec les graphes de convolution dans lequel nous avons exploré plusieurs tentatives de modélisation différentes de graphe qu'on introduit. L'ensemble des relations ont été transformées en une seule liaison grâce à node2vec, cependant le graphe obtenu n'a pas donné de résultats plus intéressants que les précédents.
Lors de la manipulation de données de sources différentes, nous avons eu beaucoup de défis à surmonter. Afin de répondre à ce genre de problématique qui revient assez souvent, on a pensé à construire une ontologie pour le scénario d'exfiltration par courriel. Comme dernier effort, nous avons proposé d'ouvrir un canal de communication simplifié pour interagir avec l'ontologie à travers un assistant virtuel.

6.1 Limitations de la solution proposée

Une limitation pour la conception du graphe pour le GCN a été abordée dans le chapitre 4. En plus, nous nous sommes focalisé sur le scénario d'exfiltration par courriel qui fait partie d'un scénario beaucoup plus grand. Le travail peut être incrémenté pour atteindre le scénario global soit pour la partie apprentissage machine ou pour la partie ontologie.

D'autres expérimentations de requêtes SPARQL sont indispensables pour vérifier et confirmer l'utilisation de l'assistant virtuel sur une longue durée de temps afin de collecter l'ensemble de données voulu. L'entrainement et la manipulation de l'assistant virtuel sont nécessaires pour permettre un passage en production.

6.2 Recherches futures

Nous avons atteint dans nos recherches un début d'amélioration en proposant l'idée de l'assistant virtuel. En effet, on propose d'extraire de l'information depuis les requêtes des analystes pour l'utiliser plus tard dans le développement de modèles plus robustes en apprentissage machine. Cette orientation nécessite toutefois comme première phase, de réunir et assembler une base de connaissance solide. La deuxième étape est de collecter un ensemble de jeu de données qu'on extrait des questions que les analystes lancent à travers l'assistant.
En réalité, le souci ne se situe pas dans les modèles ou les différents algorithmes de l'apprentissage machine, mais il réside dans les données qu'on utilise pour l'entrainement.
Le marché de l'information pour les entreprises devient de plus en plus rentable et des acteurs bien organisés tentent d'en tirer profit. L'identification d'un usager comme acteur malveillant dans un établissement financier peut aider à avoir une porte d'entrée aux activités illicites conduites par l'organisation criminelle sur l'ensemble des institutions financières. Une idée est de concevoir une ontologie plus globale qui réunit plusieurs acteurs économiques et tirer avantage de l'assistant virtuel comme une interface d'accès simple. En effet, les personnes qui souhaitent faire des requêtes vers la base de connaissance communiqueront directement avec l'assistant sans pour autant connaître les concepts sous-jacents.

RÉFÉRENCES

[1] "A user and entity behavior analytics scoring system explained," https://www.exabeam.com/ueba/user-entity-behavior-analytics-scoring-system-explained/, en ligne, dernier accès le 17/04/2021.

[2] Ravisha Chugh. (2020) How to choose between enterprise dlp and integrated dlp approaches. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://www.gartner.com/document/3983684?ref=authbottomrec&refval=4002997

[3] Jonathan Hui. (2020) Graph convolutional networks (gcn) & pooling. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://jonathan-hui.medium.com/graph-convolutional-networks-gcn-pooling-839184205692

[4] A. B. Shareth Ben. (2020, mai) 2020 securonix insider threat report, highlights of behaviors, detection techniuqes, and key takeaways from the field. [En ligne]. Disponible : https://pages.securonix.com/rs/179-DJP-142/images/Insider-Threat-Report-May-2020-Securonix.pdf

[5] T. C. P. Tara Deschamps, "Shopify says it notified privacy commissioner of breach involving 'rogue' staff," September 2020, en ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://globalnews.ca/news/7352778/shopify-data-breach/

[6] Securities et E. Commission, "Sec charges amazon finance manager and family with insider trading," September 2020. [En ligne]. Disponible : https://www.sec.gov/news/press-release/2020-228

[7] P. institue. (2020) 2020 cost of insider threats global report. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://www.proofpoint.com/sites/default/files/gtd-pfpt-uk-tr-ponemon-institute-2020-cost-of-insider-threats.pdf

[8] M. Maybury, P. Chase, B. Cheikes, D. Brackney, G. Fort, Meade, S. Matzner, B. Wood, T. Longstaff, T. Hetherington, C. Sibley, J. Marin, L. Spitzner, J. Copeland, S. Lewandowski et J. Haile, "Analysis and detection of malicious insiders," 03 2005.

[9] B. Aleman-Meza, P. Burns, M. Eavenson, D. Palaniswami et A. Sheth, "An ontological approach to the document access problem of insider threat," dans *Intelligence and Security Informatics*, P. Kantor, G. Muresan, F. Roberts, D. D. Zeng, F.-Y. Wang, H. Chen et R. C. Merkle, édit. Berlin, Heidelberg : Springer Berlin Heidelberg, 2005, p. 486–491.

[10] Q. Althebyan et B. Panda, "A Knowledge-Base Model for Insider Threat Prediction," dans *2007 IEEE SMC Information Assurance and Security Workshop*.

West Point, NY, USA : IEEE, juin 2007, p. 239–246. [En ligne]. Disponible : http://ieeexplore.ieee.org/document/4267567/

[11] I. Homoliak, F. Toffalini, J. Guarnizo, Y. Elovici et M. Ochoa, "Insight into insiders and it : A survey of insider threat taxonomies, analysis, modeling, and countermeasures," *ACM Comput. Surv.*, vol. 52, n°. 2, avr. 2019. [En ligne]. Disponible : https://doi.org/10.1145/3303771

[12] S. Pfleeger, J. B. Predd, J. Hunker et C. Bulford, "Insiders behaving badly : Addressing bad actors and their actions," *IEEE Transactions on Information Forensics and Security*, vol. 5, p. 169–179, 2010.

[13] S. Sinclair et S. W. Smith, *Preventative Directions For Insider Threat Mitigation Via Access Control*. Boston, MA : Springer US, 2008, p. 165–194. [En ligne]. Disponible : https://doi.org/10.1007/978-0-387-77322-3_10

[14] M. Bishop, S. Engle, S. Peisert, S. Whalen et C. Gates, "Case studies of an insider framework," 2009.

[15] M. Theoharidou, S. Kokolakis, M. Karyda et E. Kiountouzis, "The insider threat to information systems and the effectiveness of iso17799," *Comput. Secur.*, vol. 24, n°. 6, p. 472–484, sept. 2005. [En ligne]. Disponible : https://doi.org/10.1016/j.cose.2005.05.002

[16] M. N. Al-Mhiqani, R. Ahmad, Z. Z. Abidin, W. Yassin, A. Hassan, K. H. Abdulkareem, N. S. Ali et Z. Yunos, "A review of insider threat detection : Classification, machine learning techniques, datasets, open challenges, and recommendations," *Applied Sciences*, vol. 10, p. 5208, 2020.

[17] E. Cole et S. Ring, "Insider threat : Protecting the enterprise from sabotage, spying, and theft," 2005.

[18] L. Liu, O. D. Vel, Q. Han, J. Zhang et Y. Xiang, "Detecting and preventing cyber insider threats : A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, p. 1397–1417, 2018.

[19] C. Probst, J. Hunker, D. Gollmann et M. Bishop, "Insider threats in cyber security," dans *Insider Threats in Cyber Security*, 2010.

[20] B. Lindauer, "Insider Threat Test Dataset," 9 2020. [En ligne]. Disponible : https://kilthub.cmu.edu/articles/dataset/Insider_Threat_Test_Dataset/12841247

[21] J. Glasser et B. Lindauer, "Bridging the gap : A pragmatic approach to generating insider threat data," dans *2013 IEEE Security and Privacy Workshops*, 2013, p. 98–104.

[22] E. K. Zeilenga, "Lightweight Directory Access Protocol (LDAP) : Technical Specification Road Map," Internet Requests for Comments, HJP, RFC 4510, June 2006. [En ligne]. Disponible : https://www.hjp.at/doc/rfc/rfc4510.html

[23] M. Schonlau et M. Theus, “Detecting masquerades in intrusion detection based on unpopular commands,” *Inf. Process. Lett.*, vol. 76, n°. 1–2, p. 33–38, nov. 2000. [En ligne]. Disponible : https://doi.org/10.1016/S0020-0190(00)00122-8

[24] M. B. Salem et S. Stolfo, “Modeling user search behavior for masquerade detection,” dans *RAID*, 2011.

[25] A. Harilal, F. Toffalini, I. Homoliak, J. Castellanos, J. Guarnizo, S. Mondal et M. Ochoa, “The wolf of sutd (twos) : A dataset of malicious insider threat behavior based on a gamified competition,” *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, vol. 9, p. 54–85, 2018.

[26] “Mitre attack framework,” https://attack.mitre.org, en ligne, dernier accès le 08/11/2021.

[27] H. Wang, H. Xu, B. Lu et Z. Shen, “Research on Security Architecture for Defending Insider Threat,” dans *2009 Fifth International Conference on Information Assurance and Security*. Xi’An China : IEEE, 2009, p. 30–33. [En ligne]. Disponible : http://ieeexplore.ieee.org/document/5283457/

[28] G. Ali, N. A. Shaikh et Z. A. Shaikh, “Towards an automated multiagent system to monitor user activities against insider threat,” dans *2008 International Symposium on Biometrics and Security Technologies*. Isalambad, Pakistan : IEEE, avr. 2008, p. 1–5. [En ligne]. Disponible : http://ieeexplore.ieee.org/document/4547660/

[29] Q. Yaseen et B. Panda, “Tackling insider threat in cloud relational databases,” dans *Proceedings of the 2012 IEEE/ACM Fifth International Conference on Utility and Cloud Computing*, ser. UCC ’12. USA : IEEE Computer Society, 2012, p. 215–218. [En ligne]. Disponible : https://doi.org/10.1109/UCC.2012.18

[30] W. R. Claycomb, C. L. Huth, B. Phillips, L. Flynn et D. McIntire, “Identifying indicators of insider threats : Insider it sabotage,” dans *2013 47th International Carnahan Conference on Security Technology (ICCST)*, 2013, p. 1–5.

[31] “Knowledgegraph conference 2019,” https://www.knowledgegraph.tech/conference-2019/speakers/, en ligne, dernier accès le 08/11/2021.

[32] J. F. Sequeda, *Integrating Relational Databases with the Semantic Web : A Reflection*. Cham : Springer International Publishing, 2017, p. 68–120. [En ligne]. Disponible : https://doi.org/10.1007/978-3-319-61033-7_4

[33] D. Costa, M. Albrethsen, M. Collins, S. Perl, G. Silowash et D. Spooner, “An insider threat indicator ontology,” Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, Rapport technique CMU/SEI-2016-TR-007, May 2016. [En ligne]. Disponible : http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=454613

[34] A. Ambre et N. Shekokar, “Insider threat detection using log analysis and event correlation,” *Procedia Computer Science*, vol. 45, p. 436–445, 12 2015.

[35] E. T. Axelrad, P. Sticha, O. Brdiczka et J. Shen, “A bayesian network model for predicting insider threats,” *2013 IEEE Security and Privacy Workshops*, p. 82–89, 2013.

[36] W. Liu, L. Ci et L. Liu, “Research on behavior trust based on bayesian inference in trusted computing networks,” dans *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*, 2015, p. 1134–1138.

[37] E. Santos, H. Nguyen, F. Yu, K. J. Kim, D. Li, J. T. Wilkinson, A. Olson, J. Russell et B. Clark, “Intelligence analyses and the insider threat,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A : Systems and Humans*, vol. 42, n°. 2, p. 331–347, 2012.

[38] F. Janjua, A. Masood, H. Abbas et I. Rashid, “Handling insider threat through supervised machine learning techniques,” *Procedia Computer Science*, vol. 177, p. 64–71, 01 2020.

[39] J. Yao, S. Zhao et l. fan, “An enhanced support vector machine model for intrusion detection,” 09 2006, p. 538–543.

[40] D. S. Kim et J. Park, “Network-based intrusion detection with support vector machines,” *Lecture Notes in Computer Science*, vol. 2662, p. 747–756, 01 2003.

[41] B. Bhati et C. Rai, “Analysis of support vector machine-based intrusion detection techniques,” *Arabian Journal for Science and Engineering*, vol. 45, 07 2019.

[42] P. Parveen, Z. R. Weger, B. Thuraisingham, K. Hamlen et L. Khan, “Supervised learning for insider threat detection using stream mining,” dans *2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, 2011, p. 1032–1039.

[43] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor et J. Platt, “Support vector method for novelty detection,” dans *Proceedings of the 12th International Conference on Neural Information Processing Systems*, ser. NIPS’99. Cambridge, MA, USA : MIT Press, 1999, p. 582–588.

[44] T. Rashid, I. Agrafiotis et J. R. C. Nurse, “A new take on detecting insider threats : Exploring the use of hidden markov models,” *Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats*, 2016.

[45] J. Jiang, J. Chen, T. Gu, K.-K. R. Choo, C. Liu, M. Yu, W. Huang et P. Mohapatra, “Anomaly detection with graph convolutional networks for insider threat and fraud detection,” dans *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)*, 2019, p. 109–114.

[46] C. Legg, “Ontologies on the semantic web,” *Annual Review of Information Science and Technology*, vol. 41, n°. 1, p. 407–451, 2007. [En ligne]. Disponible : https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/aris.2007.1440410116

[47] N. Noy et D. Mcguinness, “Ontology development 101 : A guide to creating your first ontology,” *Knowledge Systems Laboratory*, vol. 32, 01 2001.

[48] M. Ehrig, *Ontology Alignment : Bridging the Semantic Gap*, ser. Semantic Web and Beyond. New York : Springer, 2007, vol. 4.

[49] M. Taye, “Understanding semantic web and ontologies : Theory and applications,” *Journal of Computing*, vol. 2, 06 2010.

[50] T. Berners-Lee, J. Hendler et O. Lassila, “The semantic web,” *Scientific American*, vol. 284, n°. 5, p. 28–37, 2001.

[51] Wikipedia contributors, “Ramanathan v. guha — Wikipedia, the free encyclopedia,” 2021, [Online ; accessed 29-December-2021]. [En ligne]. Disponible : https://en.wikipedia.org/w/index.php?title=Ramanathan_V._Guha&oldid=1010280962

[52] W3C. (2021) Rdf schema 1.1. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://www.w3.org/TR/rdf-schema/

[53] M. O’Connor, R. Shankar, C. Nyulas, S. Tu et A. Das, “Developing a web-based application using owl and swrl,” 01 2008, p. 93–98.

[54] V. Fortineau, T. Paviot, L. Louis-Sidney et S. Lamouri, “Swrl as a rule language for ontology-based models in power plant design,” vol. 388, 07 2012.

[55] V. Raskin, C. Hempelmann, K. Triezenberg et S. Nirenburg, “Ontology in information security : a useful theoretical foundation and methodological tool.” 01 2001, p. 53–59.

[56] C. T. Harold Booth, “Vulnerability description ontology (vdo),” National Institute of Standards and Technology, Rapport technique, 2016.

[57] W. Li, Y. Zhu et S. Tian, “Intrusion Alerts Correlation Model Based on XSWRL Ontology,” dans *2008 Second International Symposium on Intelligent Information Technology Application*. Shanghai, China : IEEE, déc. 2008, p. 894–898. [En ligne]. Disponible : http://ieeexplore.ieee.org/document/4739700/

[58] H. A. Karande et S. S. Gupta, “Ontology based intrusion detection system for web application security,” dans *2015 International Conference on Communication Networks (ICCN)*. Gwalior, India : IEEE, nov. 2015, p. 228–232. [En ligne]. Disponible : http://ieeexplore.ieee.org/document/7507454/

[59] G. G. Granadillo, Y. B. Mustapha, N. Hachem et H. Debar, “An ontology-driven approach to model SIEM information and operations using the SWRL formalism,”

International Journal of Electronic Security and Digital Forensics, vol. 4, n°. 2/3, p. 104, 2012. [En ligne]. Disponible : http://www.inderscience.com/link.php?id=48412

[60] C. Onwubiko, "CoCoa : An Ontology for Cybersecurity Operations Centre Analysis Process," dans *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. Glasgow : IEEE, juin 2018, p. 1–8. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8551486/

[61] L. Obrst, P. Chase et R. Markeloff, "Developing an ontology of the cyber security domain," dans *STIDS*, 2012.

[62] G. B. Moreira, V. M. Calegario, J. C. Duarte et A. F. P. D. Santos, "Csiho : An ontology for computer security incident handling," 2018.

[63] G. Sadowski, J. Care, N. McDonald et H. Teixeira. (2019) Market guide for user and entity behavior analytics. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://www.gartner.com/document/3917096?ref=solrResearch&refval=293215666

[64] Kurtis Pykes. (2020) The vanishing/exploding gradient problem in deep neural networks. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://towardsdatascience.com/the-vanishing-exploding-gradient-problem-in-deep-neural-networks-191358470c11

[65] T. N. Kipf et M. Welling, "Semi-supervised classification with graph convolutional networks," 2017.

[66] C. Nwankpa, W. Ijomah, A. Gachagan et S. Marshall, "Activation functions : Comparison of trends in practice and research for deep learning," *ArXiv*, vol. abs/1811.03378, 2018.

[67] M. Wang, D. Zheng, Z. Ye, Q. Gan, M. Li, X. Song, J. Zhou, C. Ma, L. Yu, Y. Gai, T. Xiao, T. He, G. Karypis, J. Li et Z. Zhang, "Deep graph library : A graph-centric, highly-performant package for graph neural networks," 2020.

[68] "Why dgl ?" https://www.dgl.ai/pages/about.html, en ligne, dernier accès le 16/09/2021.

[69] D. P. Kingma et J. Ba, "Adam : A method for stochastic optimization," 2017.

[70] A. Grover et J. Leskovec, "Node2vec : Scalable feature learning for networks," dans *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA : Association for Computing Machinery, 2016, p. 855–864. [En ligne]. Disponible : https://doi.org/10.1145/2939672.2939754

[71] A. Dehghan-Kooshkghazi, B. Kamiński, Łukasz Kraiński, P. Prałat et F. Théberge, "Evaluating node embeddings of complex networks," 2021.

[72] J. Han, M. Kamber et J. Pei, “2 - getting to know your data,” dans *Data Mining (Third Edition)*, third edition éd., ser. The Morgan Kaufmann Series in Data Management Systems, J. Han, M. Kamber et J. Pei, édit. Boston : Morgan Kaufmann, 2012, p. 39–82. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/B9780123814791000022

[73] T. Mikolov, K. Chen, G. S. Corrado et J. Dean, “Efficient estimation of word representations in vector space,” 2013. [En ligne]. Disponible : http://arxiv.org/abs/1301.3781

[74] Manish Chabiani. (2017) Word2vec (skip-gram model) : Part 1 - intuition. En ligne, dernier accès le 08/11/2021. [En ligne]. Disponible : https://towardsdatascience.com/word2vec-skip-gram-model-part-1-intuition-78614e4d6e0b

[75] D. Müllner, “Modern hierarchical, agglomerative clustering algorithms,” *ArXiv*, vol. abs/1109.2378, 2011.

[76] L. van der Maaten et G. Hinton, “Viualizing data using t-sne,” *Journal of Machine Learning Research*, vol. 9, p. 2579–2605, 11 2008.

[77] “Why dgl ?” https://www.infoworld.com/article/3228245/the-80-20-data-science-dilemma.html, en ligne, dernier accès le 16/09/2021.

[78] J. F. Sequeda, W. J. Briggs, D. P. Miranker et W. P. Heideman, “A pay-as-you-go methodology to design and build enterprise knowledge graphs from relational databases,” dans *The Semantic Web – ISWC 2019*, C. Ghidini, O. Hartig, M. Maleshkova, V. Svátek, I. Cruz, A. Hogan, J. Song, M. Lefrançois et F. Gandon, édit. Cham : Springer International Publishing, 2019, p. 526–545.

[79] Rasa Technologies, “Rasa : Open source conversational ai.” [En ligne]. Disponible : https://rasa.com/

[80] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, n°. 1, p. 5–32, 2001. [En ligne]. Disponible : http://dx.doi.org/10.1023/A%3A1010933404324

[81] S. W. Kwok et C. Carter, “Multiple decision trees,” dans *Uncertainty in Artificial Intelligence*, ser. Machine Intelligence and Pattern Recognition, R. D. SHACHTER, T. S. LEVITT, L. N. KANAL et J. F. LEMMER, édit. North-Holland, 1990, vol. 9, p. 327–335. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/B9780444886507500305

[82] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, n°. 8, p. 832–844, août 1998. [En ligne]. Disponible : https://doi.org/10.1109/34.709601

[83] T. G. Dietterich, “An experimental comparison of three methods for constructing ensembles of decision trees : Bagging, boosting, and randomization,” *Mach. Learn.*, vol. 40, n°. 2, p. 139–157, août 2000. [En ligne]. Disponible : https://doi.org/10.1023/A:1007607513941

[84] Y. Amit et D. Geman, “Shape quantization and recognition with randomized trees,” *Neural Comput.*, vol. 9, n°. 7, p. 1545–1588, oct. 1997. [En ligne]. Disponible : https://doi.org/10.1162/neco.1997.9.7.1545

[85] M. El Sanharawi et F. Naudet, “Comprendre la régression logistique,” *Journal Français d’Ophtalmologie*, vol. 36, n°. 8, p. 710–715, 2013. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0181551213002490

[86] J. Peng, K. Lee et G. Ingersoll, “An introduction to logistic regression analysis and reporting,” *Journal of Educational Research - J EDUC RES*, vol. 96, p. 3–14, 09 2002.

[87] P. Common, “Classification supervisée par réseaux multicouches,” *Traitement du Signal*, vol. 8, p. 387–407, 06 1991.

[88] “Chapitre 2. réseaux de neurones et approximation de fonction, 3. le perceptron multicouche,” https://www.becoz.org/these/memoirehtml/ch06s04.html, en ligne, dernier accès le 10/10/2021.

[89] M. Hossin et S. M.N, “A review on evaluation metrics for data classification evaluations,” *International Journal of Data Mining & Knowledge Management Process*, vol. 5, p. 01–11, 03 2015.

[90] T. Fawcett, “Introduction to roc analysis,” *Pattern Recognition Letters*, vol. 27, p. 861–874, 06 2006.

ANNEXE A APPRENTISSAGE MACHINE

Cette section donne une breve introduction aux différentes techniques utilisées pour comparer les résultats obtenus avec notre modèle basé sur le graphe de convolution.

Forêts aléatoires

Les forêts aléatoires - *Random forests* en Anglais - est un algorithme d'apprentissage machine supervisé qui a été conçu et présenté par *Brieman* en 2001 [80] comme une méthode de classification et de regression à l'aide des arbres de décision *Classification and Regression Trees* (CART). Le développement s'est basé sur plusieurs travaux antérieurs [81–84]. Chaque arbre de la forêt est entrainé séparement avec un échantillon aléatoire de données et avec un sous ensemble des caractéristiques présentes dans le jeu de données. Ces deux principes assure une diversité entre les arbres dans la forêt. Le résultat final est calculé selon le résultat de chaque arbre.

Régression logistique

La régression logitisque est une méthode d'apprentissage supervisé utilisée dans les problèmes de classification. Cette technique lie une variabale dépendante dichotomique (Y) à des variables explicatives continues ou catégorielles ($X_1, X_2, X_3, ...X_n$) [85].
Le modèle en régression logistique s'apparente à la régression linéaire qui permet de lier une variable (Y) quantitative et des variables explicatives ($X_1, X_2, X_3, ...X_n$). Parfois il est difficile de décrire la distribution des éléments avec une équation linéaire, car les données ne suivent pas une tendance linéaire et ils ne suivent pas une distribution normale ou constante. La régression logisitique permet de remédier à ces problèmes en applicant une transformation logit sur la variable dépendante (Y) afin d'obetnir la prédiction [86]. La régréssion logistique est définie par les équations A.1 et A.2.

$$logit(Y) = ln(\frac{\pi}{1-\pi}) = \beta X \tag{A.1}$$

où :

π : dénote la probabilité de succès (Y=1) pour un individu.

β : représente le vecteur des coefficients de régression.

X : dénote le vecteur des variables pour l'individu X.

$$\pi = \frac{e^{\beta X}}{1 + e^{\beta X}} \tag{A.2}$$

Perceptron MultiCouche

Un Perceptron MultiCouche (PMC) est un type de réseau de neurones artificiel organisé en une architecture stratiéfié de neurones formels [87], dans laquelle l'information transite depuis la couche d'entrée vers la couche de sortie. Une neurone de perceptron dans un PMC ajoute un biais b au produit scalaire de son vectuer d'entrée x et le vecteur de poids w. $y = f(x \cdot w + b)$, f représente une fonction d'activation [88].

Métriques

Précision

Cette métrique est utilisée pour mesurer les vrais positifs correctement idéntifiés du total des vrais positifs disponibles dans le jeu de données [89].

$$Précision = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux positifs}} \tag{A.3}$$

Spécéficité

Cette métrique est utilisée pour mesurer la fraction des vrais négatifs correctement identifés par le modèle [89].

$$Spécéficité = \frac{\text{Vrais négatifs}}{\text{Vrais négatif} + \text{Faux négatifs}} \tag{A.4}$$

Rappel

Cette métrique est utilisée pour mesurer la fraction des vrais positifs correctement identifés par le modèle [89].

$$Rappel = \frac{\text{Vrai positifs}}{\text{Vrai positifs} + \text{Vrai négatifs}} \tag{A.5}$$

F1 score

Cette métrique représente la moyenne harmonique entre les valeurs de rappel et de précision [89].

$$Rappel = 2 \times \frac{Précision \times Rappel}{Précision + Rappel} \tag{A.6}$$

ROC

La métrique ROC mesure la capacité d'un modèle à identifier les taux des classes du jeu de données en variant le seuil obtenu comme résultat par le modèle pour une classe. [90]

ANNEXE B CONCEPTS DE L'ONTOLOGIE

Concepts de l'ontologie

Concepts, attributs et relations pour la méthode *pay-as-you-go*

TABLEAU B.1 Rapport de connaissance - Concept Incident exfiltration par courriel

Nom du concept	Incident_exfiltration
Sous concept de	Incident
id du concept	incident_exf
Nom de la table ou requête SQL	select inc from incident where type_menace='exfiltration par courriel'

TABLEAU B.2 Rapport de connaissance - Concept Usager

Nom du concept	Usager
id du concept	usager
Nom de la table ou requête SQL	select usager from user_table

Listing B.1 Classe indicateur en OWL

```
mi:Indicateur rdf:type owl:Class ; rdfs:label "Indicateur" .
mi:nom_indicateur rdf:type owl:DatatypeProperty ; rdfs:domain
    mi:Indicateur; rdfs:range xsd:string.
mi:cote rdf:type owl:DatatypeProperty ; rdfs:domain mi:Indicateur;
    rdfs:range xsd:integer; rdfs:label "Cote de risque".
mi:date_ind rdf:type owl:DatatypeProperty ; rdfs:domain mi:Indicateur ;
    rdfs:range xsd:dateTime ; rdfs:label "Date de declenchement" .
mi:a_origine rdf:type owl:ObjectProperty; rdfs:label "origine de
    l'indicateuer" rdfs:doamin mi:Indicateur; rdfs:range mi:action-email.
```

Tableau B.3 Rapport de connaissance - Concept Employé

Nom du concept	Employé
Sous concept de	Usager
id du concept	employé
Nom de la table ou requête SQL	select usager as employé from user_table where type_emploi='régulier'

Tableau B.4 Rapport de connaissance - Concept Consultant

Nom du concept	Consultant
Sous concept de	Usager
id du concept	consultant
Nom de la table ou requête SQL	select usager as consultant from user_table where type_emploi='partenaire'

Tableau B.5 Rapport de connaissance - Concept Action-email

Nom du concept	Action-email
Sous concept de	Action-digitale
id du concept	action-email
Nom de la table ou requête SQL	select email-event as action-email from email_table

Tableau B.6 Rapport de connaissance - Concept Actif solution de sécurité

Nom du concept	Actif-sécurité
Sous concept de	Actif digital
id du concept	actif-sécurité
Nom de la table ou requête SQL	DLP,SIEM,Proxy Web ...

Tableau B.7 Rapport de connaissance - Concept Actif-fichier

Nom du concept	Actif-fichier
Sous concept de	Actif-digital
id du concept	actif-fichier
Nom de la table ou requête SQL	select attachement as actif-fichier from email_table

Tableau B.8 Rapport de connaissance - Concept adresse-courriel

Nom du concept	Adresse courriel
Sous concept de	Actif-Digital
id du concept	adresse-courriel
Nom de la table ou requête SQL	select usager, adresse-courriel from user_table | select adresse-source from email_table | select adresse-destination from email_table

Tableau B.9 Rapport de connaissance - Concept Compte

Nom du concept	Usager
id du concept	usager
Nom de la table ou requête SQL	select compte from user_table

Tableau B.10 Rapport de connaissance - Attributs de l'employé

Nom de l'attribut	Fonction
id de l'attribut	fonction
Appliqué au concept	Employé
Nom de la table ou requête SQL	select usager,fonction from user_table where type_emploi='régulier'
Nom de la colonne	fonction
Type de données	Chaine de caractères
Possible de valeur nulle	Non
Valeur si nulle	N/A
Nom de l'attribut	Unité organisationelle
id de l'attribut	unité
Appliqué au concept	Employé
Nom de la table ou requête SQL	select usager,unité from user_table where type_emploi='régulier'
Nom de la colonne	unité
Type de données	Chaine de caractère
Possible de valeur nulle	Non
Valeur si nulle	N/A
Nom de l'attribut	Direction
id de l'attribut	direction
Appliqué au concept	Employé
Nom de la table ou requête SQL	select usager,direction from user_table where type_emploi='régulier'
Nom de la colonne	direction
Type de données	Chaine de caractère
Possible de valeur nulle	Non
Valeur si nulle	N/A

Tableau B.11 Rapport de connaissance - Attributs du consultant

Nom de l'attribut	Partenaire
id de l'attribut	partenaire
Appliqué au concept	Consultant
Nom de la table ou requête SQL	select usager,partenaire from user_table where type_emploi='partenaire'
Nom de la colonne	partenaire
Type de données	Chaine de caractères
Possible de valeur nulle	Non
Valeur si nulle	N/A

Tableau B.12 Rapport de connaissance - Attributs de l'indicateur

Nom de l'attribut	Nom de l'indicateur
id de l'attribut	nom_indicateur
Appliqué au concept	Indicateur
Nom de la table ou requête SQL	select ind,nom_indicateur from table_inidcateur
Nom de la colonne	nom_indicateur
Type de données	Chaine de caractères
Possible de valeur nulle	Non
Valeur si nulle	N/A
Nom de l'attribut	Cote de risque
id de l'attribut	cote
Appliqué au concept	Indicateur
Nom de la table ou requête SQL	select ind,cote from table_inidcateur
Nom de la colonne	cote
Type de données	Entier positif
Possible de valeur nulle	Oui
Valeur si nulle	0
Nom de l'attribut	Date de déclanchement
id de l'attribut	date_ind
Appliqué au concept	Indicateur
Nom de la table ou requête SQL	select ind,date_ind from table_inidcateur
Nom de la colonne	date
Type de données	Date
Possible de valeur nulle	Non
Valeur si nulle	N/A

Tableau B.13 Rapport de connaissance - Attributs de l'incident

Nom de l'attribut	Date incident
id de l'attribut	date_inc
Appliqué au concept	Incident
Nom de la table ou requête SQL	select inc,date_inc from incident
Nom de la colonne	date_inc
Type de données	Date
Possible de valeur nulle	Non
Valeur si nulle	N/A
Nom de l'attribut	Titre
id de l'attribut	titre
Appliqué au concept	Incident
Nom de la table ou requête SQL	select inc,titre from incident
Nom de la colonne	titre
Type de données	Chaine de caratère
Possible de valeur nulle	Non
Valeur si nulle	N/A
Nom de l'attribut	Journal
id de l'attribut	journal
Appliqué au concept	Incident
Nom de la table ou requête SQL	select inc,titre from incident
Nom de la colonne	titre
Type de données	Chaine de caratère
Possible de valeur nulle	Non
Valeur si nulle	N/A

Tableau B.14 Rapport de connaissance - Relation indicateurs dans l'incident exfiltration par courriel

Nom de la relation	indicateur considéré
Définition de la relation	Un incident d'exfiltration concerne plusieurs indicateurs de surveillance de courriel sortants pour un usager
Id de la relation	indicateur_considéré
Depuis le concept	Incident
Nom de la table ou requête SQL	select inc,usager,ind from incident i join user_table u join table_indicateur t where i.usager=o.usager and i.usager=t.usager and i.date_inc > t.date_ind and t.nom_indicateur like '%mail%'
Vers le concept	Indicateur

Tableau B.15 Rapport de connaissance - Relation possède_email

Nom de la relation	a l'adresse courriel
Définition de la relation	Un compte possède au plus une adresse courriel
Id de la relation	a_corriel
Depuis le concept	Compte
Nom de la table ou requête SQL	user_table
Vers le concept	Adresse courriel

Tableau B.16 Rapport de connaissance - Relation usager dans l'incident exfiltration par courriel

Nom de la relation	usager considéré
Définition de la relation	Un incident d'exfiltration concerne un usager
Id de la relation	usager_considéré
Depuis le concept	Incident
Nom de la table ou requête SQL	select inc,usager from incident i join user_table u where i.usager=o.usager
Vers le concept	Usager

Tableau B.17 Rapport de connaissance - Attributs du Actif-fichier

Nom de l'attribut	Nom du fichier
id de l'attribut	nom-fichier
Appliqué au concept	Actif-fichier
Nom de la table ourequêSQL	select attachement as actif-fichier , nom-fichier from email_table
Nom de la colonne	nom-fichier
Type de données	Chaine caractères
Possible de valeur nulle	Oui
Valeur si nulle	N/A
Nom de l'attribut	Taille du fichier
id de l'attribut	taille-fichier
Appliqué au concept	Actif-fichier
Nom de la table ourequêSQL	select attachement as actif-fichier , taille-fichier from email_table
Nom de la colonne	taille-fichier
Type de données	Chaine caractères
Possible de valeurnulle	Oui
Valeur si nulle	N/A
Nom de l'attribut	Extension
id de l'attribut	extension
Appliqué au concept	Actif-fichier
Nom de la table ourequêSQL	select attachement as actif-fichier ,extension from email_table
Nom de la colonne	extension
Type de données	Chaine caractères
Possible de valeurnulle	Oui
Valeur si nulle	N/A
Nom de l'attribut	Type du fichier
id de l'attribut	type-fichier
Appliqué au concept	Actif-fichier
Nom de la table ourequêSQL	select attachement as actif-fichier , type-fichier from email_table
Nom de la colonne	type-fichier
Type de données	Chaine caractères
Possible de valeurnulle	Oui
Valeur si nulle	N/A

Tableau B.18 Rapport de connaissance - Relation source de l'action

Nom de la relation	source Action
Définition de la relation	Une ation est récupérer depuis une source donnée.
Id de la relation	source_action
Depuis le concept	Action
Nom de la table ou requête SQL	Incident
Vers le concept	Actif sécurité

Tableau B.19 Rapport de connaissance - Relation entre l'indicateur et l'évènement d'origine

Nom de la relation	évènement d'origine
Définition de la relation	Un indicateur d'exfiltration par courriel découle d'une action émail
Id de la relation	a_origine
Depuis le concept	Indicateur
Nom de la table ou requête SQL	select ind,action-email from from Indicateur i join email_table e where i.action-email=e.action-email
Vers le concept	Action-Email

Tableau B.20 Rapport de connaissance - Relation Pièce jointe

Nom de la relation	Pièce jointe
Définition de la relation	Une action courriel peut contenir une ou plusieurs pièces jointes
Id de la relation	a_pièce_jointe
Depuis le concept	Action-email
Nom de la table ou requête SQL	email_table
Vers le concept	Actif-fichier

Tableau B.21 Rapport de connaissance - Relation adresse source

Nom de la relation	Adresse Source
Définition de la relation	Une action courriel doit avoir une adresse courriel source
Id de la relation	adresse_source
Depuis le concept	Action-email
Nom de la table ou requête SQL	select email-event,adresse-source from email_table
Vers le concept	Adresse courriel

Tableau B.22 Rapport de connaissance - Relation adresse destination

Nom de la relation	Adresse Destination
Définition de la relation	Une action courriel doit avoir une ou plusieurs adresses courriel de destination
Id de la relation	adresse_source
Depuis le concept	Action-email
Nom de la table ou requête SQL	select email-event, adresse-destination from email_table
Vers le concept	Adresse courriel

Tableau B.23 Rapport de connaissance - Relation possède_compte

Nom de la relation	a le compte
Définition de la relation	Un usager possède un ou plusieurs comptes
Id de la relation	a_compte
Depuis le concept	Usager
Nom de la table ou requête SQL	select user from user_table
Vers le concept	Compte