Mémoire présenté devant l'Université de Paris-Dauphine
pour l'obtention du Certificat d'Actuaire de Paris-Dauphine
et l'admission à l'Institut des Actuaires
le 14 octobre 2024

Par : Nicolas ROSAL
Titre : Utilisation de données de cybersécurité pour une quantification dynamique du risque cyber

Confidentialité : ☑ Non ☐ Oui (Durée : ☐ 1 an ☐ 2 ans)

Les signataires s'engagent à respecter la confidentialité ci-dessus

Membres présents du jury de l'Institut des Actuaires :

Entreprise :
Nom : Exiom Partners
Signature :

Directeur de Mémoire en entreprise :
Nom : Peyrat Thomas
Signature :

Membres présents du Jury du Certificat d'Actuaire de Paris-Dauphine :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels (après expiration de l'éventuel délai de confidentialité)

Secrétariat :

Signature du responsable entreprise

Bibliothèque :

Signature du candidat

Résumé

Classé en première place de la cartographie des risques de France Assureurs depuis 2017, le risque cyber inquiète aujourd'hui plus que jamais. L'assurance cyber, en plein développement, se heurte aux spécificités de ce risque et à une pénétration inégale du marché des entreprises, particulièrement faible pour les PME. Ce contexte, accentué par un manque de données de sinistre fiables, rend la quantification complexe. De nombreux acteurs demandent alors une meilleure connaissance du risque et une quantification en temps réel pouvant se détacher de la nécessité d'une base de données de sinistres. Ce mémoire explore alors sous forme de pistes de recherche l'utilisation de données de cybersécurité pour la quantification dynamique du risque cyber dans un contexte d'assurance. Une modélisation basée sur des graphes et des modèles graphiques bayésiens est proposée. L'étude se concentre sur les PME, particulièrement vulnérables, et développe une approche permettant d'intégrer la volatilité des risques et l'évolution des vulnérabilités spécifiques à chaque entreprise, tout en se détachant du cadre classique. À travers un portefeuille fictif, la méthodologie est appliquée à un contexte de perte d'exploitation. Le mémoire démontre comment cette approche dynamique peut améliorer la quantification du risque en assurance cyber et comment elle peut être utilisée pour aider l'assureur à proposer l'accompagnement dont les PME ont besoin. L'approche est prometteuse, mais nécessite de futures recherches pour être applicable, notamment concernant les différents paramètres du modèle.

Mots-clés : *risque cyber* ; *modélisation dynamique* ; *PME* ; *graphes bayésiens* ; *tarification* ; *cybersécurité.*

Abstract

Ranked first in the risk mapping by France Assureurs since 2017, cyber risk is now more concerning than ever. Cyber insurance, which is rapidly developing, faces challenges due to the specific nature of this risk and uneven market penetration among companies, particularly low for SMEs. This context, compounded by a lack of reliable claims data, makes pricing complex. Many stakeholders are therefore calling for better risk understanding and real-time quantification that does not rely on a claims database. This thesis thus explores, in the form of research avenues, the use of cybersecurity data for the dynamic quantification of cyber risk in an insurance context. A modeling approach based on graphs and Bayesian graphical models is proposed. The study focuses on SMEs, which are particularly vulnerable, and develops a methodology that accounts for the volatility of risks and the evolution of vulnerabilities specific to each company, moving away from traditional frameworks. Using a hypothetical portfolio, the methodology is applied to a business interruption context. The thesis demonstrates how this dynamic approach can enhance cyber risk quantification in insurance and how it can help insurers provide the support SMEs need. While the approach is promising, it requires further research to become fully applicable, particularly regarding the various model parameters.

Keywords : *cyber risk*; *dynamic modeling*; *SMEs*; *bayesian graphs*; *pricing*; *cybersecurity.*

Note de Synthèse

Avant-propos

Il est à noter que ce mémoire s'inscrit dans une démarche exploratoire et vise à investiguer des pistes méthodologiques pour mieux quantifier le risque cyber. À ce titre, les approches proposées ne prétendent pas à une applicabilité immédiate, mais cherchent plutôt à lever certains verrous conceptuels dans un domaine où beaucoup reste à construire.

Contexte

La transformation numérique généralisée dans tous les secteurs a entraîné une forte dépendance aux systèmes d'information. Bien que source de connectivité et de productivité accrues, cette dépendance expose à des cybermenaces aux conséquences potentiellement dévastatrices.

Le risque cyber, impactant la confidentialité, l'intégrité et la disponibilité des données, connaît une croissance mondiale, qui touche particulièrement les entreprises. Les cyberincidents engendrent divers coûts (pertes d'exploitation, remédiation, etc.), rendant l'assurance cyber attrayante. Toutefois, cette couverture reste largement concentrée sur les grandes entreprises, avec moins de 15 % des ETI assurées contre 98 % des grandes entreprises (AMRAE, 2024). Cette faible adoption est d'autant plus notable que les PME et les ETI, plus vulnérables, manquent souvent de sensibilisation aux enjeux de la cybersécurité.

Le développement de l'assurance cyber est freiné par des difficultés de quantification du risque. Le secteur souffre d'un manque de données fiables en raison de la faible maturité des acteurs (risque émergent) et de la réticence des entreprises à fournir des informations sur leurs sinistres cyber, souvent stratégiques. La modélisation est également complexifiée par les interactions entre les différents acteurs, tant au niveau sectoriel qu'au niveau des systèmes d'information, rendant le risque systémique. Enfin, le risque cyber est un risque fortement évolutif, et ce, à une maille inférieure à la maille annuelle. L'ensemble de ces points complique l'utilisation d'une méthodologie actuarielle classique (du type coût fréquence).

Ce mémoire cherche à quantifier le risque cyber de manière dynamique dans le temps et à une maille fine, afin d'adapter la quantification aux spécificités des assurés et de capter le risque au niveau des systèmes d'information. En s'éloignant des modèles standards, cette méthodologie intègre des données de cybersécurité, ce qui permet d'obtenir une vision actualisée de l'exposition aux menaces cyber. Ainsi, cette modélisation ouvre la voie à de nouvelles méthodologies pour améliorer la prévention et renforcer l'accompagnement des PME.

Cette note commencera par présenter le cadre théorique, en détaillant les outils utilisés et la structure générale de la modélisation. Ensuite, un modèle assurantiel sur la perte d'exploitation sera introduit comme cadre applicatif du mémoire. Une base fictive, créée pour observer les résultats de ce modèle, sera ensuite présentée. Enfin, les principaux résultats seront exposés dans la dernière partie.

Cadre Théorique

Dans ce mémoire, la modélisation se concentre sur **l'attaque** plutôt que sur le sinistre. Cette approche permet de modéliser une attaque cyber et d'en estimer la fréquence, au lieu d'adopter le schéma classique coût-fréquence. Ce choix repose sur le fait que les données issues de la cybersécurité sur les attaques sont plus accessibles que celles sur les sinistres, souvent plus sensibles. Cette partie exposera le cadre de modélisation des déplacements d'un attaquant au sein d'une entreprise, tout en traduisant les pertes "physiques" (comme l'indisponibilité d'un serveur) en pertes économiques pour l'organisation.

Graphe d'attaque - De la vision descriptive à la probabilisation

Le réseau informatique d'une entreprise peut être représenté comme un graphe où chaque arête est une connexion entre deux éléments (ordinateurs, serveurs, etc.). L'attaquant se déplace sur ce réseau en exploitant des vulnérabilités *humaines* (phishing, ingénierie sociale, ...) ou *techniques* (se basant sur des vulnérabilités logicielles, des mauvaises configurations de pare-feu, ...). Le réseau peut être comparé à une demeure, où l'attaquant joue le rôle d'un voleur cherchant à atteindre un trésor. Pour y parvenir, il progresse en franchissant différentes barrières de sécurité, exploitant chaque vulnérabilité sur son passage, qu'il s'agisse d'une porte mal verrouillée ou de la crédulité du propriétaire.

Un formalisme appelé **graphe d'attaque** permet de modéliser l'espace d'évolution de l'attaquant. Représenté mathématiquement sous la forme d'un graphe orienté, il cartographie les combinaisons de vulnérabilités exploitables pour atteindre une cible au sein d'un réseau. Contrairement au réseau, qui représente l'ensemble des connexions possibles, le graphe d'attaque ne retient que les chemins **accessibles** à l'attaquant. Pour reprendre la comparaison avec une maison, toutes les portes ne sont pas mal fermées : seules celles qui le sont constituent des points d'entrée exploitables et sont donc représentées dans le graphe. Ce dernier peut être construit à différentes échelles : une échelle physique, illustrant les actifs de l'entreprise (ordinateurs, serveurs), ou une échelle plus abstraite, décrivant les privilèges acquis par l'attaquant. Ce mémoire se concentrera sur la première approche, axée sur la modélisation des actifs physiques.

Tatar et al., 2020 proposent l'utilisation de ce type de graphe dans un contexte assurantiel. En plus de l'approche descriptive fournie par le graphe d'attaque, une approche probabiliste est ajoutée en quantifiant la probabilité qu'a l'attaquant d'exploiter une vulnérabilité. Pour évaluer cette probabilité pour une vulnérabilité technique, l'article utilise les métriques d'exploitabilité du score CVSS, un indice mesurant la "dangerosité" de la faille. Toutes les failles techniques découvertes étant enregistrées dans la base CVE, ce score est rapidement calculé et disponible sur Internet.

La probabilité d'exploiter une vulnérabilité ne correspond pas directement à celle de la compromission d'un actif. Un actif peut avoir plusieurs vulnérabilités, augmentant son risque de compromission, et certaines vulnérabilités ne sont accessibles qu'à partir de nœuds spécifiques du réseau. Par analogie, une porte d'entrée non verrouillée ne peut être exploitée si le portail est fermé à clé. Pour évaluer cette interdépendance et cette structure, la théorie des **réseaux bayésiens** est utilisée. Pour calculer la probabilité conditionnelle au sein du réseau bayésien construit, une méthode basée sur le papier de Poolsappasit et al., 2012 est utilisée dans ce mémoire.

Il est ainsi possible de définir une structure de probabilité sur le graphe d'attaque. Cette structure permettra d'établir une méthodologie pour simuler une attaque sur le réseau de l'entreprise.

Graphe d'Impact - De la perte informatique à la perte économique

Une fois cet aspect pris en compte, il faut traduire la perte observée sur le système informatique en une perte économique pour l'entreprise. Cette transformation s'effectue à l'aide d'un **Graphe**

d'Impact, qui modélise les dépendances opérationnelles de l'entreprise sous forme de graphe orienté. Introduit par Jakobson, 2011 et adapté au domaine cyber par Tatar et al., 2020, ce graphe est structuré en trois couches. Parmi elles, la *couche des actifs*, représentant les nœuds du graphe d'attaque, et la *couche business*, permettant d'estimer les impacts économiques.

Une fois les actifs compromis identifiés lors de l'attaque, il devient possible, grâce au réseau de dépendances, de propager la perte d'opérabilité de la couche des actifs (où l'attaque a eu lieu) vers la couche business (où l'on peut quantifier les pertes économiques).

Modélisation Stochastique de la perte d'exploitation

La partie précédente a définie une méthodologie pour modéliser et probabiliser une attaque ainsi que le cadre permettant de transformer une perte informatique en perte économique pour l'entreprise. Cette partie proposera une application concrète de ces concepts au domaine de l'assurance dans le contexte de la perte d'exploitation cyber.

Le modèle de coût repose sur la simulation d'un grand nombre d'attaques sur l'entreprise. Il utilise comme base probabiliste un graphe d'attaque bayésianisé. À l'issue de chaque simulation, la perte physique est convertie en perte économique. Cette méthodologie permet, en bout de chaîne, d'obtenir une distribution de la perte d'exploitation de l'assuré en cas d'attaque. L'architecture du modèle créé est détaillée dans son intégralité dans la partie (2.3.2).

Pour se concentrer sur la perte d'exploitation, deux phases distinctes sont simulées :

- Une phase d'**attaque**, modélisant le déplacement de l'attaquant dans l'entreprise ainsi que la compromission des différents actifs,
- et une phase de **remédiation** durant laquelle l'entreprise va "réparer" les actifs compromis, période pendant laquelle sa productivité ne sera pas optimale.

La perte de l'entreprise n'est ainsi pas calculée sur un instant unique, mais bien sur une période complète, jusqu'à ce que l'ensemble des actifs compromis pendant l'attaque soit remédié.

Les étapes attaque/remédiation sont répétées de manière indépendante n fois afin d'obtenir une distribution des coûts de pertes d'exploitation pour une entreprise donnée, dont on connaît le graphe d'attaque et le graphe d'impact.

À cela pourrait s'ajouter un modèle d'estimation de la fréquence d'attaque afin d'obtenir une modélisation complète du risque. Cette partie n'est pas traitée dans ce mémoire, mais les données nécessaires à son estimation sont plus accessibles que celles sur la fréquence des sinistres, grâce aux études de cybersécurité portant sur les menaces actuelles.

Le Portefeuille fictif

La suite de ce mémoire cherche à observer les résultats de l'application du modèle de perte d'exploitation. Les données nécessaires à l'utilisation de celui-ci n'existant pas aujourd'hui à l'échelle d'un portefeuille (graphe d'attaque et graphe d'impact en particulier), il a été choisi d'élaborer un portefeuille fictif se concentrant sur les PME. Ce portefeuille a été construit selon une méthodologie bien spécifique permettant d'interpréter au mieux les résultats applicatifs du modèle.

La limitation du graphe d'impact sur la diversité des secteurs du portefeuille

Le choix des secteurs est important lors de la création d'un portefeuille, particulièrement en cyberassurance, car chaque domaine d'activité repose sur une architecture business et informatique spécifique, créant des sensibilités distinctes face aux cybermenaces. Le graphe d'impact, qui représente l'architecture de l'entreprise, est donc central dans la modélisation du portefeuille. Pour élaborer ces graphes dans le cadre du portefeuille fictif, des **entretiens approfondis avec des dirigeants de PME** ont permis de concevoir des graphes d'impact représentatifs de deux secteurs clés : l'industrie manufacturière et le commerce en ligne.

La création d'un graphe d'attaque est une tâche complexe et encore largement théorique. Ainsi, le choix a été fait de se concentrer sur deux secteurs afin d'analyser la sensibilité du modèle à ce graphe, tout en limitant le nombre de postulats nécessaires.

La simulation des graphes d'attaque

La création des graphes d'attaque pour le portefeuille fictif repose sur une méthodologie simplifiée mais inspirée de procédures réalistes. Cette démarche inclut la simulation d'un réseau d'entreprise, l'attribution de vulnérabilités (CVE) aux actifs, et la construction du graphe d'attaque à partir de ces données.

Chaque entreprise dispose d'un réseau généré aléatoirement en fonction de sa taille. Les actifs incluent des ordinateurs, serveurs, pare-feux et routeurs. La topologie adoptée est de type multi-étoiles, couramment utilisée pour sa gestion centralisée.

Une bibliothèque de vulnérabilités est construite en plusieurs étapes. D'abord, les actifs sont classés par type de programme, comme le système d'exploitation ou le navigateur pour un PC. Ensuite, chaque type de programme est associé aux principaux logiciels existants ainsi qu'à leur part de marché, par exemple Windows, macOS et Linux pour les systèmes d'exploitation. Enfin, les vulnérabilités (CVE) sont extraites via l'API de la NVD en filtrant par logiciel. Les logiciels sont ensuite répartis dans le réseau en fonction des parts de marché de chaque compétiteur. Chaque actif reçoit un nombre de failles aléatoire, dont la nature dépend de ses caractéristiques techniques.

Le graphe d'attaque est généré à partir du réseau construit en amont. Un lien est ajouté entre deux actifs si le nœud cible présente des vulnérabilités exploitables. La probabilité d'exploitation $\overline{p(e)}$ est calculée en cumulant les failles associées.

Cette méthodologie, bien que simplifiée, permet de simuler des graphes d'attaque cohérents avec les hypothèses formulées. Dans un cadre assurantiel réel, des outils comme MulVal et TVA remplaceraient ces simulations pour obtenir des graphes plus denses et précis.

Résultats et Analyses

La création de ce portefeuille permet l'application du modèle. Cent entreprises ont ainsi été simulées en respectant les proportions relatives du commerce et de l'industrie. L'ensemble des tests présentés a été réalisé sur ce portefeuille fictif.

Le tableau récapitulatif (1) illustre les résultats du modèle à l'échelle du portefeuille.

Ces résultats mettent en évidence la capacité du modèle à capter les différences sectorielles des entreprises. Cette variation, liée à la construction du graphe d'impact propre à chaque secteur, confirme l'aptitude du modèle à s'adapter aux spécificités des entreprises. Cependant, ces résultats sont influencés par les écarts de chiffre d'affaires entre les deux secteurs. Le mémoire propose une seconde analyse en exprimant les coûts en pourcentage du chiffre d'affaires afin de valider cette hypothèse.

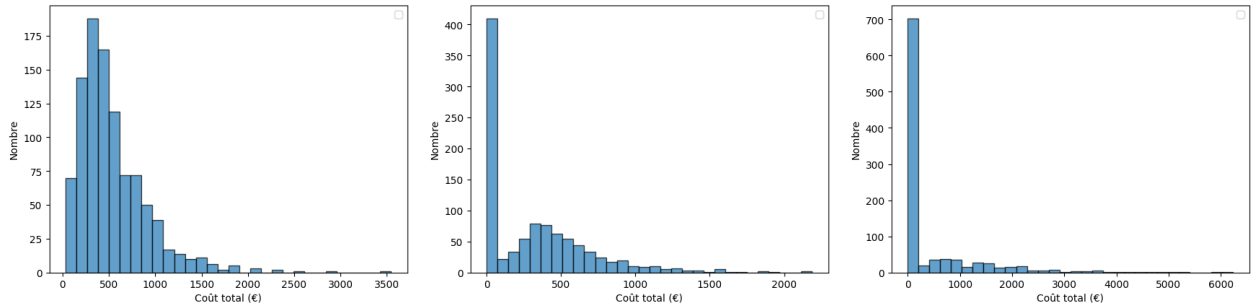
	Coût total	Nombre dans la base	Coût moyen
Portefeuille Total	416 513 €	100	4 165,13 €
Commerce	113 846,73 €	54	2 108,27 €
Industrie	302 666,25 €	46	6 579,70 €

TABLE 1 : Récapitulatif des coûts du portefeuille.

Sensibilité au graphe d'attaque

L'une des forces du modèle repose sur sa capacité à différencier les entreprises selon leur exposition aux risques cyber. Grâce à l'utilisation des graphes d'attaque dans le modèle, il est possible d'observer, dans les résultats, une distinction nette entre les entreprises en fonction de leur niveau de vulnérabilité et de leur capacité à résister aux attaques.

La figure (1) présente le résultat du modèle (distribution des coûts lors d'une attaque) pour trois entreprises du même secteur, similaires en taille et en chiffre d'affaires.



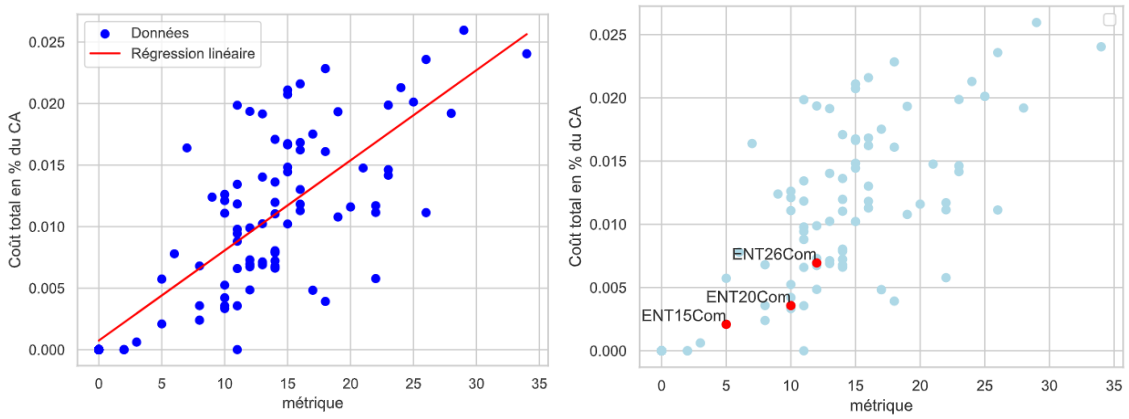
(a) Histogramme des coûts pour l'entreprise *ENT26Com* (b) Histogramme des coûts pour l'entreprise *ENT20Com* (c) Histogramme des coûts pour l'entreprise *ENT15Com*

FIGURE 1 : Comparaison de la répartition des coûts entre trois entreprises.

Ces entreprises, bien que similaires en apparence, présentent des profils de risque très différents. Lorsqu'on compare ce résultat avec les graphes d'attaque de chacune d'entre elles, on observe des structures de vulnérabilités significativement différentes. L'entreprise présentant le graphe d'attaque le plus dense est aussi celle avec la répartition de coûts la plus conséquente, et inversement pour l'entreprise avec le graphe d'attaque le plus réduit.

Ces résultats montrent que la "complexité" du graphe d'attaque semble avoir un impact fort sur les résultats. Pour vérifier cette hypothèse, une étude a été menée sur le rapport entre la complexité du graphe d'attaque et les résultats du coût en pourcentage du chiffre d'affaires. Plusieurs métriques de complexité ont été comparées, et une régression linéaire multiple par sélection exhaustive a permis d'identifier la meilleure composition. Finalement, la longueur et la largeur maximales du graphe sont les deux paramètres retenus dans la régression linéaire multiple avec le meilleur R^2. La figure (2) présente les résultats de cette régression et la position des trois entreprises précédemment étudiées.

La régression linéaire appliquée aux graphes d'attaque met en évidence une relation claire entre la structure du graphe et le niveau de risque, bien que certaines variations subsistent en raison de facteurs spécifiques, tels que les métriques CVE des failles. Ces résultats suggèrent qu'une analyse topologique du graphe d'attaque pourrait permettre d'obtenir une première estimation fiable du risque, ce qui constitue un atout majeur face au manque de données dans le domaine de la cyberassurance.



(a) Régression linéaire sur la métrique choisie et le pourcentage de CA.

(b) Position des entreprises 26, 20 et 15 sur le graphique.

FIGURE 2 : Présentation du coût en pourcentage du chiffre d'affaires en fonction de la métrique choisie.

S'adapter à de nouvelles menaces

L'évolution constante des menaces cyber représente un défi majeur pour l'assurance. Le risque d'une entreprise peut changer radicalement après l'apparition d'une nouvelle faille. Avec une modélisation classique, cette information ne serait pas prise en compte.

L'aspect dynamique du modèle développé dans ce mémoire a été mis en évidence par l'introduction d'une faille critique fictive dans le portefeuille, à partir du portefeuille initial. L'objectif était de simuler l'apparition d'une nouvelle faille à un temps t. Cette faille a été appliquée à deux types d'actifs : un actif central du réseau (pare-feu) et un actif non central. L'évolution des coûts a ensuite été observée.

L'application de la faille sur un actif central entraîne une augmentation du coût total moyen du portefeuille de 7,34 %. Toutes les entreprises ne sont pas égales à cette nouvelle menance, certaines enregistrent une hausse de plus de 100 %, tandis que d'autres ne voient leur coût augmenter que de 1 %. En revanche, l'introduction de la faille sur un actif non central entraîne une augmentation plus modérée, de l'ordre de 1,23 %.

Ces observations montrent que les nœuds centraux du réseau sont nettement plus critiques que les nœuds périphériques, une observation logique car un attaquant doit généralement compromettre un nœud central avant d'atteindre un nœud final.

Conclusion

La gestion du risque cyber est un enjeu majeur pour les assureurs en raison de la complexité croissante des menaces et du manque de données historiques. L'évaluation dynamique du risque s'impose comme une nécessité pour mieux appréhender l'exposition des entreprises, en particulier des PME, souvent les plus vulnérables.

Dans cette étude, nous avons proposé un modèle basé sur les graphes bayésiens d'attaque et les graphes d'impact afin de quantifier le risque cyber de manière plus dynamique. Ce modèle prend en compte la structure interne des entreprises et permet d'évaluer les pertes potentielles en fonction de leur exposition aux cybermenaces. L'approche retenue est en cohérence avec les connaissances en cybersécurité et démontre une capacité d'adaptation face à l'apparition de nouvelles menaces.

Ces outils permettent non seulement d'affiner la quantification du risque, mais également d'accompagner les assurés en proposant des mesures de prévention ciblées. Ainsi, leur intégration dans un

cadre assurantiel réaliste apparaît prometteuse.

Néanmoins, cette étude demeure largement théorique et nécessite des recherches supplémentaires, notamment pour affiner l'estimation de la fréquence des incidents, améliorer la qualité des données utilisées (graphes d'attaque et d'impact) et calibrer les paramètres du modèle.

Synthesis note

Foreword

It should be noted that this thesis is part of an exploratory approach aiming to investigate methodological avenues for improved quantification of cyber risk. Therefore, the approaches proposed here do not claim immediate applicability but rather seek to overcome certain conceptual barriers in a field where much remains to be built.

Context

The widespread digital transformation across all sectors has led to a strong dependence on information systems. While this dependence enhances connectivity and productivity, it also exposes organizations to cyber threats with potentially devastating consequences.

Cyber risk, which impacts the confidentiality, integrity, and availability of data, is growing globally, particularly affecting businesses. Cyber incidents generate various costs (business interruption, remediation, etc.), making cyber insurance attractive. However, this coverage remains largely concentrated on large enterprises, with less than 15% of mid-sized companies insured compared to 98% of large enterprises (AMRAE, 2024). This low adoption is all the more notable as SMEs and mid-sized companies, being more vulnerable, often lack awareness of cybersecurity issues.

The development of cyber insurance is hindered by difficulties in risk quantification. The sector suffers from a lack of reliable data due to the low maturity of market players (emerging risk) and companies' reluctance to disclose information about their cyber incidents, which are often strategic. Modeling is also complicated by interactions between different stakeholders, both at the sectoral level and within information systems, making cyber risk systemic. Furthermore, cyber risk evolves rapidly, at a scale smaller than an annual timeframe. All these factors make it difficult to apply a traditional actuarial methodology (such as the cost-frequency approach).

This study aims to quantify cyber risk dynamically over time and at a granular level, allowing the quantification to be tailored to the specificities of different actors and to capture threats at the information system level. By moving away from standard models, this methodology integrates cybersecurity data, providing an updated view of cyber threat exposure. Thus, this modeling approach paves the way for new opportunities to improve prevention and strengthen support for SMEs.

This paper will begin by presenting the theoretical framework, detailing the tools used and the general philosophy of the modeling approach. Then, an insurance model for business interruption loss will be introduced as the study's application framework. A fictitious dataset, created to observe the results of this model, will then be presented. Finally, the main results will be discussed in the last section.

Theoretical Framework

In this thesis, the modeling focuses on **the attack** rather than the claim. This approach allows for modeling a cyber attack and estimating its frequency instead of adopting the traditional cost-frequency scheme. This choice is based on the fact that cybersecurity data on attacks is more accessible than data on claims, which are often more sensitive. This section will present the modeling framework for an attacker's movements within a company while translating "physical" losses (such as server unavailability) into economic losses for the organization.

Attack Graph - From Descriptive View to Probabilization

The network of a company can be represented as a graph where each edge is a connection between two elements (computers, servers, etc.). The attacker moves through this network by exploiting *human* vulnerabilities (phishing, social engineering, ...) or *technical* vulnerabilities (based on software flaws, misconfigured firewalls, ...). The network can be compared to a house, where the attacker plays the role of a thief trying to reach a treasure. To succeed, they progress by overcoming various security barriers, exploiting each vulnerability along the way, whether it be an unlocked door or the owner's gullibility.

A formalism called the **attack graph** allows for modeling the attacker's movement space. Represented as a directed graph, it maps out the combinations of vulnerabilities that can be exploited to reach a target within a network. Unlike the network, which represents all possible connections, the attack graph retains only the **accessible** paths for the attacker. Returning to the house analogy, not all doors are poorly locked: only those that are represent exploitable entry points and are thus included in the graph. This graph can be constructed at different scales: a physical scale, illustrating the company's assets (computers, servers), or a more abstract scale, describing the privileges acquired by the attacker. This thesis will focus on the first approach, centered on modeling physical assets.

Tatar et al., 2020 propose the use of this type of graph in an insurance context. In addition to the descriptive approach provided by the attack graph, a probabilistic approach is added by quantifying the probability of an attacker exploiting a vulnerability. To assess this probability for a technical vulnerability, the article uses the exploitability metrics from the CVSS score, an index measuring the severity of the flaw. Since all discovered technical vulnerabilities are recorded in the CVE database, this score is quickly calculated and available online.

The probability of exploiting a vulnerability does not directly correspond to the probability of compromising an asset. An asset may have multiple vulnerabilities, increasing its risk of compromise, and some vulnerabilities are only accessible from specific network nodes. By analogy, an unlocked front door cannot be exploited if the outer gate is locked. To evaluate this interdependence and structure, the theory of **Bayesian networks** is used. To compute the conditional probability within the constructed Bayesian network, this thesis employs a method based on the paper by Poolsappasit et al., 2012.

Thus, it is possible to define a probabilistic structure on the attack graph. This structure will establish a methodology for simulating an attack on the company's network.

Impact Graph - From IT Loss to Economic Loss

Once this aspect is considered, it is necessary to translate the observed loss in the IT system into an economic loss for the company. This transformation is carried out using an **Impact Graph**, which models the company's operational dependencies in the form of a directed graph. Introduced by Jakobson, 2011 and adapted to the cyber domain by Tatar et al., 2020, this graph is structured into three layers. Among them, the *asset layer* represents the nodes of the attack graph, while the *business layer* allows for the estimation of economic impacts.

Once the compromised assets are identified during the attack, it becomes possible, thanks to the dependency network, to propagate the loss of operability from the asset layer (where the attack occurred) to the business layer (where economic losses can be quantified).

Stochastic Modeling of Business Interruption Loss

The previous section defined a methodology for modeling and probabilizing an attack, as well as the framework for transforming an IT loss into an economic loss for the company. This section will propose a concrete application of these concepts in the insurance domain within the context of cyber business interruption loss.

The cost model is based on simulating a large number of attacks on the company. It uses a Bayesian attack graph as its probabilistic foundation. At the end of each simulation, the physical loss is converted into an economic loss. This methodology ultimately provides a distribution of the insured's business interruption loss in the event of an attack. The full architecture of the created model is detailed in section (2.3.2).

To focus on business interruption loss, two distinct phases are simulated:

- An **attack phase**, modeling the attacker's movement within the company and the compromise of various assets,
- and a **remediation phase**, during which the company "repairs" the compromised assets, a period during which its productivity will not be optimal.

The company's loss is thus not calculated at a single instant but rather over an entire period, until all assets compromised during the attack have been remediated.

The attack/remediation steps are repeated independently n times to obtain a distribution of business interruption loss costs for a given company, whose attack graph and impact graph are known.

Additionally, an attack frequency estimation model could be incorporated to achieve a complete risk modeling. This model is not covered in this thesis, but the data required for its estimation is more accessible than that for claim frequency, thanks to cybersecurity studies on current threats.

The Fictional Portfolio

The following sections of this thesis aim to observe the results of applying the business interruption loss model. Since the necessary data for using this model does not currently exist at a portfolio scale (particularly the attack graph and impact graph), a fictional portfolio focusing on SMEs was created. This portfolio was constructed using a specific methodology to best interpret the model's applied results.

The Limitation of the Impact Graph on Portfolio Sector Diversity

The choice of sectors is crucial when creating a portfolio, especially in cyber insurance, as each business domain relies on a specific IT and business architecture, leading to distinct sensitivities to cyber threats. The impact graph, which represents the company's architecture, is therefore central to the portfolio's modeling. To construct these graphs for the fictional portfolio, **in-depth interviews with SME executives** were conducted to design representative impact graphs for two key sectors: manufacturing and e-commerce.

Creating an attack graph remains a complex and largely theoretical task. Thus, the decision was made to focus on two sectors to analyze the model's sensitivity to this graph while limiting the number of required assumptions.

Simulating Attack Graphs

The creation of attack graphs for the fictional portfolio is based on a simplified methodology inspired by realistic procedures. This approach includes simulating a company network, assigning vulnerabilities (CVE) to assets, and constructing the attack graph based on these data.

Each company has a randomly generated network based on its size. The assets include computers, servers, firewalls, and routers. The adopted topology follows a multi-star structure, commonly used for centralized management.

A vulnerability library is built in several steps. First, assets are classified by program type, such as operating systems or browsers for a PC. Then, each program type is linked to major existing software and their market shares, for example, Windows, macOS, and Linux for operating systems. Finally, vulnerabilities (CVE) are retrieved via the NVD API by filtering based on software. The software is then distributed across the network according to each competitor's market share. Each asset is assigned a random number of vulnerabilities, the nature of which depends on its technical characteristics.

The attack graph is generated based on the previously constructed network. A link is added between two assets if the target node has exploitable vulnerabilities. The probability of exploitation $\overline{p(e)}$ is calculated by aggregating the associated vulnerabilities.

Although simplified, this methodology allows for the simulation of attack graphs that are consistent with the formulated assumptions. In a real insurance context, tools such as MulVal and TVA would replace these simulations to obtain denser and more precise graphs.

Results and Analysis

The creation of this portfolio enables the application of the model. One hundred companies were simulated while maintaining the relative proportions of the commerce and industry sectors. All the tests presented were conducted on this fictional portfolio.

The summary table (2) illustrates the model's results at the portfolio level.

	Total Cost	**Number in Database**	**Average Cost**
Total Portfolio	€416,513	100	€4,165.13
Commerce	€113,846.73	54	€2,108.27
Industry	€302,666.25	46	€6,579.70

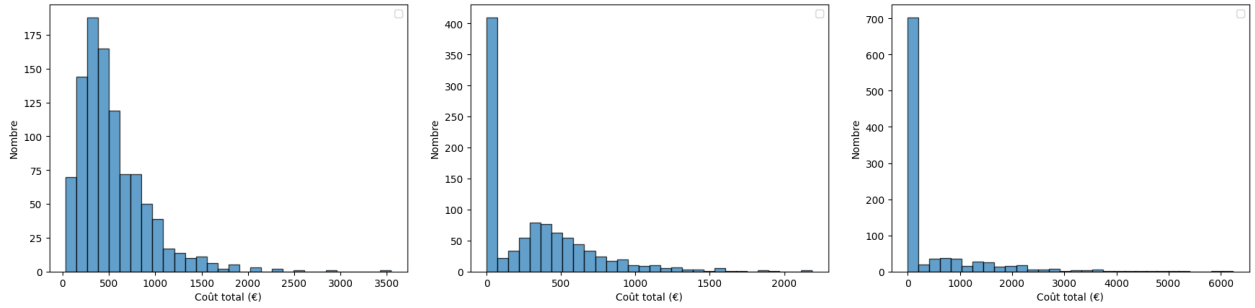
Table 2: Summary of portfolio costs.

These results highlight the model's ability to capture sectoral differences among companies. This variation, linked to the construction of the impact graph specific to each sector, confirms the model's ability to adapt to the specificities of businesses. However, these results are influenced by revenue differences between the two sectors. The thesis proposes a second analysis by expressing costs as a percentage of revenue to validate this hypothesis.

Sensitivity to the Attack Graph

One of the strengths of the model lies in its ability to differentiate companies based on their exposure to cyber risks. By incorporating attack graphs into the model, it is possible to observe a clear distinction in the results between companies depending on their level of vulnerability and their resilience to attacks.

Figure (3) presents the model's results (distribution of costs during an attack) for three companies within the same sector, similar in size and revenue.



(a) Cost histogram for company *ENT26Com*

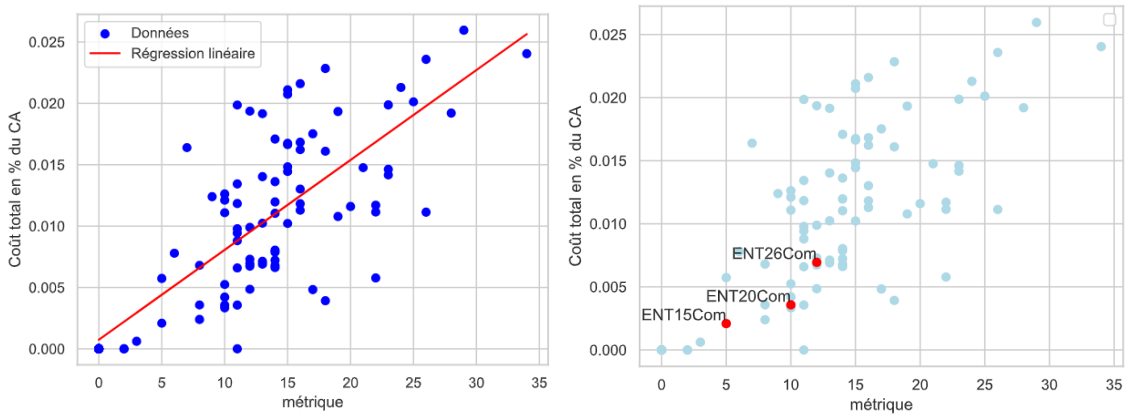
(b) Cost histogram for company *ENT20Com*

(c) Cost histogram for company *ENT15Com*

Figure 3: Comparison of cost distribution across three companies.

Although these companies appear similar, they exhibit significantly different risk profiles. When comparing these results with each company's attack graph, we observe substantially different vulnerability structures. The company with the densest attack graph also has the highest cost distribution, while the company with the least complex attack graph experiences the lowest costs.

These results indicate that the "complexity" of the attack graph seems to have a strong impact on the outcomes. To verify this hypothesis, a study was conducted on the relationship between attack graph complexity and cost results expressed as a percentage of revenue. Several complexity metrics were compared, and a multiple linear regression using exhaustive selection was performed to identify the best combination. Ultimately, the maximum length and width of the graph were selected as the two key parameters in the multiple linear regression with the best R^2. Figure (4) presents the results of this regression and the position of the three previously analyzed companies.



(a) Linear regression on the selected metric and revenue percentage.

(b) Positioning of companies 26, 20, and 15 on the graph.

Figure 4: Representation of cost as a percentage of revenue based on the selected metric.

The linear regression applied to the attack graphs highlights a clear relationship between graph structure and risk level, although some variations remain due to specific factors such as CVE vulnerability metrics. These findings suggest that a topological analysis of the attack graph could provide an initial reliable risk estimate, which is a major advantage given the lack of data in the cyber insurance field.

Adapting to New Threats

The constant evolution of cyber threats represents a major challenge for insurance. A company's risk profile can change drastically following the emergence of a new vulnerability. With a traditional modeling approach, this information would not be taken into account.

The dynamic nature of the model developed in this thesis was demonstrated by introducing a fictional critical vulnerability into the portfolio, based on the initial portfolio. The objective was to simulate the emergence of a new vulnerability at time t. This vulnerability was applied to two types of assets: a central asset in the network (firewall) and a non-central asset. The evolution of costs was then observed.

Applying the vulnerability to a central asset resulted in an average total portfolio cost increase of 7.34%. Not all companies were equally affected by this new threat: some experienced an increase of over 100%, while others saw only a 1% rise in costs. In contrast, introducing the vulnerability on a non-central asset led to a more moderate increase of around 1.23%.

These observations indicate that central nodes in the network are significantly more critical than peripheral nodes—an intuitive finding, as an attacker generally needs to compromise a central node before reaching a final target.

Conclusion

Managing cyber risk is a major challenge for insurers due to the increasing complexity of threats and the lack of historical data. Dynamic risk assessment has become a necessity to better understand companies' exposure, particularly SMEs, which are often the most vulnerable.

In this study, we proposed a model based on Bayesian attack graphs and impact graphs to quantify cyber risk more dynamically. This model takes into account the internal structure of companies and enables the evaluation of potential losses based on their exposure to cyber threats. The chosen approach aligns with cybersecurity knowledge and demonstrates adaptability in response to emerging threats.

These tools not only refine risk quantification but also support insured companies by offering targeted prevention measures. Thus, their integration into a realistic insurance framework appears promising.

However, this study remains largely theoretical and requires further research, particularly to refine incident frequency estimation, improve the quality of the data used (attack and impact graphs), and calibrate the model parameters.

Remerciements

Par la rédaction de ce mémoire se clôturent cinq années d'études, possiblement les plus denses en savoir et en émotions de ma jeune existence. C'est donc tout naturellement que je tiens à remercier l'ensemble des personnes ayant contribué, par leur aide, leur soutien ou leur présence à mes côtés, à l'aboutissement de ce travail et, plus généralement, à la réussite de ces dernières années.

Je souhaite tout d'abord remercier M. Thomas Peyrat, mon tuteur d'entreprise. Les mots me manquent pour exprimer ma gratitude envers son aide, ses conseils et ses remarques, toujours d'une pertinence exceptionnelle. Son encadrement et sa confiance sans faille m'ont permis de surmonter les moments de doute et d'incertitude.

Je tiens aussi à remercier l'équipe EXIOM dans son ensemble pour son accueil chaleureux, sa solidarité et sa bienveillance. En particulier, je tiens à exprimer ma gratitude à l'ensemble du pôle actuariat pour leurs conseils et leur amicalité, qui m'ont accompagné tout au long de ce mémoire.

Je veux également témoigner ma reconnaissance aux personnes m'ayant aidé, conseillé et inspiré tout de la construction de ce travail. Je pense en particulier à M. Unal Tatar, sans qui ce mémoire ne serait sûrement pas le même.

Dans le même esprit, je remercie toute l'équipe pédagogique de l'Université Paris Dauphine, notamment M. Christophe Dutang pour son encadrement lors de ce mémoire et M. Quentin Guibert pour son implication et son dévouement sans faille envers ce master et la réussite de ses étudiants. Je n'oublie pas non plus l'ensemble du corps professoral, dont l'enseignement de qualité, la disponibilité et l'engagement ont été essentiels à la transmission de savoirs précieux.

Je remercie tout aussi chaleureusement l'INSA Rouen, où j'ai eu la chance de passer quatre années des plus enrichissantes. Cet environnement stimulant m'a permis d'approfondir mes passions et de développer tant mes compétences professionnelles que personnelles. Grâce à l'enseignement de qualité et aux nombreuses opportunités offertes, j'ai pu évoluer dans un cadre propice à la réussite et obtenir ce double diplôme, qui représente une étape cruciale dans mon parcours. Je remercie l'ensemble des professeurs qui ont su me transmettre un amour profond pour les mathématiques et une réelle passion pour l'informatique. Mes pensées vont en particulier à M. Nicolas Forcadel, qui m'a donné l'opportunité d'effectuer ce double diplôme, et à M. Antoine Tonnoir, pour nous avoir partagé sa passion des mathématiques de la meilleure des façons.

Je remercie également mes ami.e.s, celles et ceux à qui je voue un amour profond et qui m'ont épaulé, conseillé et avec qui j'ai partagé des moments de vie inoubliables. Votre présence, dans les moments de bonheur comme dans les épreuves, m'a été précieuse, et pour cela, je vous en suis profondément reconnaissant.

Enfin, je souhaite exprimer ma gratitude éternelle à ma famille. Votre amour et votre soutien indéfectible ont été un moteur tout au long de mes études et continueront à l'être dans les années à venir. Je dédie ce mémoire à chacun d'entre eux, et tout spécialement à mon frère, dont je suis immensément fier.

Contents

Introduction

La fin du XXe siècle a marqué l'une des plus grandes révolutions sociétales depuis l'industrialisation : un monde virtuel où les frontières s'effacent, les distances se raccourcissent et les échanges s'accélèrent. Des entreprises aux États, en passant par les particuliers, tous possèdent une identité virtuelle et dépendent aujourd'hui — directement ou indirectement — du numérique, d'Internet et des systèmes d'information en général. Cette dépendance ne fait que s'accroître.

Cependant, ce monde est loin d'être sans failles ni dangers. Ces menaces, connues sous le nom de cybermenaces, sont de nature, d'origine et de conséquences multiples. Leur influence, qui suit l'expansion de la dépendance sociétale au numérique, ne cesse de croître. Leur forte évolutivité et leur hétérogénéité rendent leur étude et leur mitigation complexes.
Selon l'ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information), une cybermenace se définit comme "un risque d'attaque de systèmes informatiques sur les infrastructures d'une compagnie, d'un État, d'une organisation privée ou publique, de son ou de ses systèmes d'information".
Classé premier risque (devant le péril climatique en 2023) par France Assureurs, 2023 depuis leur première cartographie en 2017, ce risque inquiète tant au niveau international qu'en France. Sa maîtrise est désormais un enjeu critique pour la pérennité du monde dans lequel nous vivons.
Que ce soit par la création d'un cadre légal (protection des données personnelles, Data Act), par l'étude et la mise en place de mesures de cybersécurité et de cyberdéfense dans les différentes structures, ou encore par le développement de l'assurance cyber, les différents acteurs sociétaux tentent tant bien que mal de lutter contre ce risque.

Le développement de l'assurance cyber a débuté assez tôt, avec les premiers contrats cyber pour entreprises déployés dans les années 90 aux États-Unis. C'est néanmoins un secteur encore en plein développement et expansion, dont la couverture est hétérogène et parfois limitée. Selon l'étude LUCY ("LUmière sur la CYber-Assurance", AMRAE, 2023), 94 % des grandes entreprises en France bénéficiaient, en 2023, d'une couverture cyber, contre seulement 10 % pour les ETI, et encore moins pour les PME. Cette distinction est d'autant plus problématique que l'impact d'une cyberattaque est bien plus lourd à porter pour les plus petites structures.
Toujours dans le rapport LUCY, on observe que pour les grandes entreprises (entre 50 M€ et 1,5 milliard € de CA), le coût moyen en 2022 s'élève à 900 000 € pour une attaque cyber, tandis qu'il est de 266 000 € pour les ETI (entre 10 M€ et 50 M€ de CA) et de 450 000 € pour les entreprises de taille moyenne (entre 2 M€ et 10 M€ de CA). La non-linéarité de l'impact d'une cyberattaque par rapport à la taille de l'entreprise est alors flagrante.
Cette réalité, combinée à la faible proportion d'entreprises assurées et à une forte proportion d'attaques ciblant les plus petites structures (34 % des attaques par rançongiciels en France visent des TPE/PME/ETI selon le "Panorama de la Cyber Menace" de l'ANSSI, 2023b), incite à la réflexion et à la recherche de solutions pour attirer ces types de structures vers l'assurance tout en réduisant leur exposition aux cybermenaces.

La modélisation du risque cyber dans un cadre assurantiel présente également des défis. Le

manque de données de sinistres rend l'utilisation de modèles classiques complexe et souvent biaisée. Les aspects systémiques et systématiques du risque cyber (développés dans la partie (1.2.3.2)) nécessitent une attention particulière de la part des assureurs, ainsi que des modélisations robustes. La problématique est d'autant plus importante que le risque cyber est fortement **évolutif** : son échelle temporelle est très courte, et différents acteurs du milieu demandent une quantification *dynamique* du risque, non seulement pour mieux assurer, mais aussi à des fins de prévention (Hillairet and Lopez, 2022).

La cybersécurité partage des enjeux communs avec la cyberassurance et présente parfois une maturité supérieure dans la compréhension du risque. Au fil des années, ce secteur a développé un ensemble de connaissances, de méthodes et de modèles qu'il serait intéressant de coupler avec les besoins assurantiels. En particulier, les notions de vulnérabilités techniques, qui sont les failles exploitées par les attaquants pour perpétrer une attaque, peuvent fournir des clés essentielles pour la quantification du risque cyber pour les entreprises.

Ainsi, l'assurance cyber est aujourd'hui en pleine construction. Elle est peu attractive pour les PME et ETI, qui nécessitent un accompagnement plus important. De plus, elle cherche des solutions pour contourner les nombreux défis du risque cyber. Le secteur de la cybersécurité, quant à lui, offre des opportunités pour étudier ce risque de manière plus précise. C'est dans ce contexte que ce mémoire s'intéressera à **l'utilisation de données de cybersécurité pour une quantification dynamique du risque cyber.**

Afin d'étudier cette problématique, le chapitre 1 commencera par présenter le risque cyber en détail. Il abordera des concepts de cybersécurité qui seront repris dans la suite du mémoire. Il présentera également le monde de la cyberassurance, tant par la réglementation que par un panorama de la situation française, ainsi qu'une présentation des défis qui entourent ce secteur.
Le chapitre 2 s'intéressera à l'application de modèles graphiques à la quantification du risque cyber. Une introduction théorique sera proposée et des modèles seront présentés. Une application à la perte d'exploitation sera ensuite réalisée par la modification des modèles étudiés et la présentation d'un modèle stochastique.
Enfin, le chapitre 3 se concentrera sur le contexte des PME. Un portefeuille fictif sera construit et le modèle sera appliqué. Cela permettra d'étudier les réponses apportées par ce type de modèle aux différentes problématiques établies au fil de ce mémoire.

Chapter 1

Présentation du monde de la cybersécurité et de la cyber-assurance

Peu peuvent se vanter aujourd'hui de n'avoir jamais entendu parler de *cyber*. Que ce soit à la télévision, lors d'un discours présidentiel sur la *cyberdéfense*, dans des articles de presse en raison d'un vol de données massif lié à une *cyberattaque*, ou encore lorsque, par inadvertance, vous avez cliqué sur un lien téléchargeant un *virus* rendant votre ordinateur inutilisable.
Néanmoins, cette proximité n'est pas synonyme de connaissance. Un vocabulaire ainsi que des outils, dont les sous-jacents sont souvent étrangers, sont alors manipulés sans trop s'en rendre compte. Définir et comprendre ce monde inconnu pour beaucoup peut permettre de mieux en saisir les enjeux et ainsi de mieux s'y adapter dans le monde assurantiel.

Ce chapitre commencera par définir les contours et les aspects essentiels du risque cyber. Seront ensuite étudiées les méthodes et connaissances liées à la gestion de ce risque au sein des entreprises. Enfin, une assurance cyber en pleine évolution sera présentée, avec une discussion sur ses limites et sur les nouvelles perspectives liant deux mondes : celui de l'actuariat et celui de la cybersécurité.

1.1 Le Risque Cyber

Cette section vise à apporter des éléments de cybersécurité afin de mieux comprendre ce domaine. Avant de s'intéresser au secteur assurantiel, le risque cyber sera donc introduit ici, en le catégorisant et en observant ses spécificités techniques.

1.1.1 Introduction au risque cyber

Selon Ministère de l'économie des Finances et de la Souveraineté Industrielle et Numérique, 2022, le risque cyber se définit comme *"l'ensemble des risques liés à l'usage des technologies numériques et peut être défini comme un risque opérationnel portant sur la confidentialité, l'intégrité ou la disponibilité des données et systèmes d'information.[...]"*.
Dans cette définition, trois mots-clés majeurs ressortent : **Confidentialité**, **Intégrité** et **Disponibilité**. Ce sont en effet les trois piliers de la *CIA Triad* (pour *Confidentiality*, *Integrity* et *Availability*).

- La **Confidentialité** correspond à la capacité d'une donnée à n'être consultable et modifiable que par un public défini. Un vol de données privées, par exemple, est une attaque qui remet en cause ce pilier.

- L'**Intégrité** désigne la capacité d'une donnée à être délivrée conformément aux décisions établies. Si la donnée est uniquement consultable par un public défini, toute modification non autorisée constitue une atteinte à son intégrité. Une modification malicieuse d'un site web, changeant les informations affichées, est une attaque affectant son intégrité.
- Enfin, la **Disponibilité** correspond à la capacité d'un service à être accessible aux utilisateurs autorisés lorsque cela est requis par son détenteur. Une attaque de type *ransomware*, bloquant l'ensemble des fichiers d'une entreprise en attendant une rançon, est une atteinte à la disponibilité.

Ainsi, un attaquant aura pour objectif de diminuer ou de supprimer l'un (ou plusieurs) des trois piliers pour un ensemble de systèmes donné. Ces actions définissent les prémices du risque cyber.
La figure 1.1 représente schématiquement les trois piliers ainsi que les résultats des actions de l'attaquant (en rouge).

(a) Schéma présentant l'atteinte à la Confidentialité.

(b) Schéma présentant l'atteinte à l'Intégrité.

(c) Schéma présentant l'atteinte à la Disponibilité.

Figure 1.1: Présentation schématique des trois piliers.

Le risque cyber couvre un domaine extrêmement vaste, englobant un large éventail de sources et d'objectifs. Ce risque émerge de multiples causes et se manifeste sous différentes formes d'attaques, affectant particulièrement les trois piliers essentiels que sont la confidentialité, l'intégrité et la disponibilité. Afin d'appréhender efficacement ce domaine complexe, il est nécessaire de le subdiviser. Ce découpage permet de transformer une entité hétérogène et opaque en un ensemble structuré et compréhensible.

Ainsi, dans ce chapitre, cette approche sera appliquée en divisant les différents aspects du risque cyber afin de mieux les analyser et les maîtriser.
Dans un premier temps, cette section s'intéressera à la présentation de l'évolution temporelle de ce risque. Elle montrera qu'il s'agit d'un risque croissant et donnera des exemples d'attaques ayant marqué le monde de la cybersécurité. Dans un second temps, le risque cyber sera présenté comme étant divisé en deux sous-risques : le risque interne et le risque externe. Ensuite, une variété de motivations et d'acteurs seront étudiées, avant d'examiner en détail les dommages causés par ce risque

et les typologies de cibles.
Cette présentation générale offrira un panorama plus clair du risque cyber et permettra d'approfondir les détails techniques sur la procédure d'attaque dans la section suivante.

1.1.1.1 Le risque cyber, un risque en constante croissance

Depuis la création de l'informatique moderne, la question de l'imperfection des systèmes d'information et de la faillibilité des opérateurs humains est un enjeu majeur. De l'ingénierie humaine dont a fait preuve Alan Turing et son équipe pour décoder les codes allemands à la fin de la Seconde Guerre mondiale, à la théorisation de la possible création d'un virus informatique par John von Neumann en 1949, les prémices des attaques sur les systèmes d'information ne sont pas nouvelles.

Néanmoins, l'évolution technologique du XXIe siècle et la dépendance sociétale aux outils numériques font du cyber un enjeu pour toutes les couches de la société.
L'intérêt des criminels pour ce domaine est alors en constante augmentation. En effet, le monde devient de plus en plus connecté, comme l'illustre l'essor de l'Internet des objets (*Internet of Things, IoT*) dont la croissance exponentielle est visible sur la figure 1.2. De plus en plus d'objets du quotidien (et par extension, d'objets utilisés dans un cadre professionnel) sont interconnectés : montres, portes, caméras de sécurité et, bien évidemment, serveurs, ordinateurs, routeurs... Cela augmente d'autant plus les cibles potentielles d'attaques, élargissant ainsi la surface d'attaque.

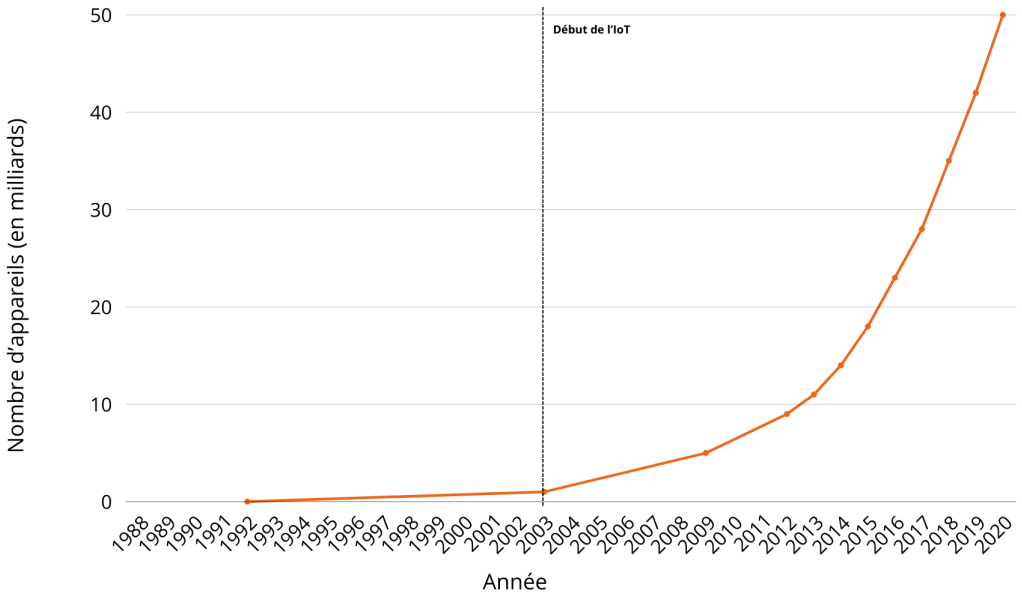


Figure 1.2: Croissance exponentielle du nombre d'objets connectés – Suarez and Salcedo, 2017

Le monde devient également bien plus dépendant de ces nouvelles technologies. Pour les attaquants, cette situation équivaut à un marché en plein essor. Pour illustrer cette tendance, il est possible d'observer la somme annuelle rapportée à l'IC3 (Internet Crime Complaint Center) depuis les années 2000 aux États-Unis (figure 1.3). Ce graphique, tiré de Statista, 2024a, met en évidence l'évolution significative de la cybercriminalité sur les vingt dernières années. Présentant une croissance exponentielle, il témoigne de l'impact croissant de ce risque sur la société.

Notons que si la tendance reflète fidèlement la situation réelle (croissance exponentielle), les chiffres sont en revanche bien inférieurs à la réalité, notamment en raison de la réticence des entreprises à

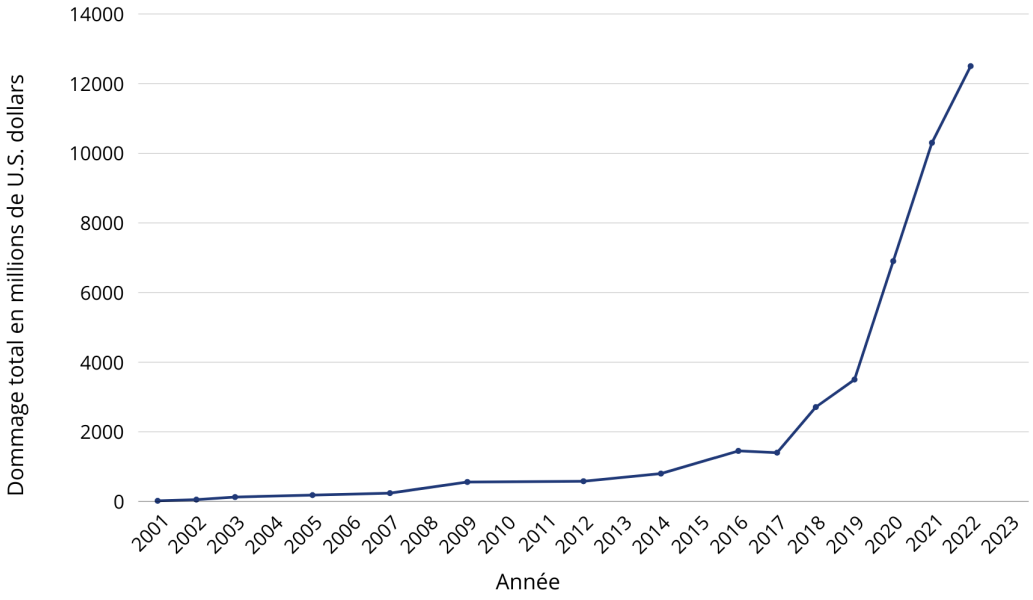


Figure 1.3: Somme annuelle des cyberattaques rapportées à l'IC3 aux États-Unis – Statista, 2024a

divulguer des informations sur le sujet. En examinant plus en détail ces chiffres, les 12,5 milliards de dollars rapportés paraissent fortement sous-estimés. Selon le World Economic Forum (World Economic Forum, 2024a), la perte mondiale liée à la cybercriminalité en 2022 s'élèverait plutôt à environ 8 400 milliards de dollars, ce qui relativise les 10,3 milliards enregistrés aux États-Unis, première puissance économique mondiale. Toujours selon cette source, cette perte devrait continuer à croître pour atteindre 23 000 milliards de dollars en 2027, représentant près de 12% du PIB mondial de 2023 (estimé à 104 476 milliards de dollars selon Wikipédia).

Bien que ces chiffres ne soient que des estimations et que le coût réel de ce type de risque reste très difficilement quantifiable à grande échelle (voir Dreyer et al., 2018 pour une analyse des difficultés d'estimation du coût du risque cyber à l'échelle mondiale et des différentes méthodes associées), le constat demeure alarmant quant à l'importance de ce risque pour l'économie mondiale.

Exemples d'attaques - Une composante systémique Pour conclure ce panorama sur l'évolution du risque, examinons quelques exemples d'attaques de masse ayant eu lieu au cours des dix dernières années.

Depuis les années 2010, les attaques se sont multipliées et ont gagné en gravité, certaines faisant la une des médias : *NotPetya, Log4Shell, Uber Data Breach...* Ces attaques illustrent à elles seules la dangerosité potentielle du risque cyber et la diversité des moyens employés par les attaquants pour rentabiliser leurs actions. Certaines d'entre elles présentent même un caractère **systémique**, ce qui préoccupe particulièrement les assureurs et sera abordé dans la section 1.2.3.2.

La liste suivante n'a pas vocation à être exhaustive, mais vise à présenter au lecteur quelques exemples marquants du monde de la cybersécurité de la dernière décennie :

WannaCry et NotPetya - 2017 Une des paires d'attaques les plus connues et les plus référencées dans les domaines de la cybersécurité et de l'assurance. WannaCry, un *ransomware* (1.1.2.1)

exploitant une faille Windows (Eternal Blue), a touché des centaines de milliers d'ordinateurs à travers 150 pays. Se présentant sous forme de ver (1.1.2.1), il s'est propagé à une échelle extrêmement large (Boydron, 2018).
NotPetya, quant à lui, est un *wiper* (1.1.2.1) supprimant toutes les données du poste infecté en utilisant la même faille que WannaCry. Initialement ciblée sur l'Ukraine, cette attaque, issue d'une opération russe, s'est propagée à l'échelle mondiale, affectant des entreprises majeures comme FedEx, Merck et Maersk. Considérée comme l'une des attaques les plus coûteuses de l'histoire (environ 10 milliards de dollars de pertes), elle souligne l'importance cruciale de la cyber-résilience. La faille exploitée par ce *wiper* avait en effet déjà été corrigée par Windows avant l'attaque, ce qui met en lumière le risque lié à l'absence de mises à jour régulières (Greenberg, 2018).

Colonial Pipeline Ransomware Attack - 2021 Cette attaque a forcé l'arrêt d'une des plus grandes infrastructures pétrolières des États-Unis, provoquant des pénuries de carburant à l'échelle nationale. Cet incident illustre la vulnérabilité des infrastructures critiques face au risque cyber (wikipedia, n.d.).

Log4j Vulnerability - 2021 Une faille critique "*zero-day*", exploitée par les attaquants avant d'avoir été détectée par les experts en cybersécurité, permettant l'exécution de code arbitraire sur des systèmes utilisant la bibliothèque Log4j. Cet exemple montre que, malgré une hygiène numérique rigoureuse, le risque cyber ne peut jamais être totalement éliminé.

UK Electoral Commission Attack - 2023 Des attaquants ont réussi à voler des informations personnelles d'électeurs entre 2014 et 2022 au Royaume-Uni, soulignant la vulnérabilité des processus démocratiques face au cyber-risque (Giannotti, 2023).

Ce paragraphe introduit divers types d'attaques externes, mais d'autres menaces existent. Le paragraphe suivant présentera un panorama global des risques cyber dans le contexte des entreprises.

1.1.1.2 Une division interne/externe

Lorsque le risque cyber est défini, il est possible de distinguer deux types de risques distincts au sein d'une entreprise :

Risques internes Cette catégorie regroupe les risques dont l'origine se situe à l'intérieur de l'entreprise. Il peut s'agir d'actes malveillants internes ou d'erreurs commises par une ou plusieurs personnes en interne, entraînant une perte économique. La distinction entre l'acte malveillant et l'erreur réside dans l'intention de l'acteur : dans le premier cas, il s'agit d'un acte délibéré, tandis que dans le second, l'erreur est involontaire. Par exemple, un employé mécontent pourrait décider de divulguer publiquement des informations sensibles de l'entreprise (*fuite de données*). À l'inverse, une mauvaise manipulation par un utilisateur disposant de droits élevés sur le serveur pourrait entraîner son arrêt et provoquer une interruption de la production (*perte de production*).

Risques externes Ces risques proviennent de menaces extérieures à l'entreprise. Les types d'acteurs sont multiples et leurs motivations variées. Néanmoins, et cela sera examiné plus en détail ultérieurement, la procédure demeure similaire : un acteur externe met en œuvre diverses méthodes pour s'introduire dans l'entreprise et atteindre ses objectifs. La section 1.1.1.3 présentera plus en détail les motivations des attaquants. Pour s'infiltrer, les acteurs malveillants peuvent exploiter leurs connaissances techniques (*exploitation de failles techniques*), utiliser des scripts ou des logiciels malveillants (par exemple, les *malwares*, ou *maliciels* en français), ainsi qu'exploiter les failles humaines (*ingénierie sociale*).

Qu'il soit intentionnel ou non, interne ou externe, le risque est omniprésent, tant à l'intérieur qu'à l'extérieur de l'entreprise. La suite de cette section s'attardera plus en détail sur les différentes motivations et les acteurs de cette menace cyber intentionnelle.

1.1.1.3 Une variété de motivations et d'acteurs

Après cette première segmentation du risque, une seconde approche sera présentée en répondant aux questions : *Pourquoi attaquer ?* et *Qui sont les attaquants ?* Un document de l'ANSSI, 2022b (Agence Nationale de la Sécurité des Systèmes d'Information) offre une bonne vision de la question.

Ainsi, plusieurs raisons peuvent inciter un attaquant à passer à l'action.

L'appât du gain L'objectif pour l'attaquant est ici de tirer un bénéfice financier, que ce soit de manière directe ou indirecte. Cela peut se concrétiser par divers moyens, allant d'une attaque ciblée à des attaques de masse, les voies pour parvenir à ses fins étant multiples.

L'espionnage L'attaquant cherche à obtenir des renseignements stratégiques présents dans le système d'information de la structure ciblée. Pour maximiser son efficacité, il tentera de rester le plus longtemps possible sur le réseau sans être détecté, afin d'extraire un maximum d'informations. L'espionnage peut concerner aussi bien des entreprises (espionnage industriel dans des secteurs stratégiques comme l'armement, le spatial ou l'énergie, par exemple) que des États (organisations ou gouvernements cherchant à obtenir un avantage stratégique).

La déstabilisation L'objectif de l'attaquant est ici de « modifier les perceptions d'une population ou de déstabiliser un acteur donné ou un processus démocratique » (ANSSI, 2022b).
L'attaque peut être le fait d'un État ou d'un groupe cherchant à perturber un processus électoral, ou encore celui de *hacktivistes* désireux de transmettre une idéologie ou un message à travers leur action.

Les exemples fournis illustrent la diversité des attaquants. Des **États** et autres **agences de renseignement**, aux **organisations criminelles** et **hacktivistes**, en passant par de simples **amateurs**, le spectre des acteurs est large.

Ces acteurs communiquent de plus en plus via des forums et des marchés noirs où ils monnayent conseils et informations. On observe ainsi l'émergence de "sociétés" spécialisées dans la cyberattaque, fournissant des outils clés en main permettant à des organisations criminelles, ou à des individus mal intentionnés, d'attaquer des cibles sans disposer de connaissances techniques avancées. Nous entrons dans l'ère du *"Cybercrime-as-a-Service (CaaS)"*. Loin de l'image du génie isolé derrière son ordinateur cherchant des failles dans un système, la réalité actuelle révèle l'existence d'une économie parallèle, organisée et extrêmement prolifique.

Dans cette section, les différentes motivations incitant aux cyberattaques ont été abordées. La section suivante adoptera le point de vue de la victime (l'entreprise) et examinera les divers types de dommages pouvant être subis en raison d'un sinistre cyber.

1.1.1.4 Les dommages du risque cyber

Pour une entreprise, subir une cyberattaque peut avoir diverses conséquences. Celles-ci dépendent, d'une part, des intentions de l'attaquant, du cadre légal, mais aussi de la réaction des clients et du public si l'information devient médiatisée. En fonction du type d'entreprise et de l'attaque subie, une

ou plusieurs des catégories présentées peuvent entraîner des pertes significatives.

D'après RiskLens, 2021, un acteur majeur de la cybersécurité, les dommages subis par une entreprise peuvent être classés en six catégories distinctes :

La perte d'exploitation (ou de productivité) Lorsqu'une attaque bloque le fonctionnement de l'entreprise, par exemple en empêchant l'utilisation de certaines fonctionnalités internes ou en rendant les services inaccessibles aux clients (site internet, logiciel, etc.), cela entraîne une **perte directe** de productivité. Cette situation peut se traduire par une diminution du chiffre d'affaires. Certains employés se retrouvent sans mission et certains services deviennent temporairement indisponibles, ce qui génère des pertes pour l'entreprise. Par exemple, pour une entreprise de commerce en ligne, la fermeture temporaire de son site web pendant une semaine engendrerait nécessairement une perte de ventes.

Les pertes liées à la remédiation Outre les pertes causées par l'attaque elle-même, la gestion de celle-ci et l'élimination de la menace représentent un surcoût important pour l'entreprise. Si elle ne dispose pas d'un service de cybersécurité (comme c'est souvent le cas pour les ETI/PME ou les entreprises peu « cyber-résilientes »), elle devra, par exemple, faire appel en urgence à des spécialistes pour traiter la menace. Le coût du traitement, de l'expertise et de la mobilisation des ressources humaines (réunions de crise, etc.) est élevé, notamment en proportion du chiffre d'affaires. Il s'agit là aussi d'un coût **direct** de l'attaque.

Le coût légal Il s'agit d'un coût **indirect**, qui s'ajoute aux conséquences de l'attaque et résulte des pénalités légales imposées à l'entreprise (pouvant inclure des sanctions financières ou des obligations de divulgation). Ces pénalités peuvent découler du non-respect d'obligations contractuelles ou des cadres légaux entourant les données personnelles. Par exemple, en cas de fuite de données (*Data Breach*), de nombreuses lois à travers le monde, comme le RGPD en Europe, imposent des sanctions sévères. Uber en a fait l'expérience en 2016, en payant une amende de 148 millions de dollars après le vol de 600 000 profils de conducteurs et de 57 millions de clients. L'entreprise avait tenté de dissimuler cet incident en versant 100 000 $ aux attaquants pour garder l'événement confidentiel, ce qui a conduit à sa condamnation (AP news, 2018).

Le coût de réputation Il correspond à la perte de confiance envers l'entreprise de la part des parties prenantes (clients, fournisseurs, banques, etc.). Par exemple, en tant que consommateur, on ferait moins confiance à une banque ayant perdu des données sensibles lors d'une cyberattaque. Il s'agit d'une perte **indirecte** pour l'entreprise.

Le coût en avantage compétitif Ce coût est lié à la perte de compétitivité après une cyberattaque ou à la perte d'un avantage stratégique sur le marché (données confidentielles, processus internes, etc.). Ces pertes sont particulièrement préoccupantes en cas de fuite de données stratégiques. Il s'agit d'une perte **indirecte**, souvent difficile à quantifier.

Dans une étude menée par le World Economic Forum, 2024b sur la cybersécurité des entreprises, un graphique met en évidence les préoccupations des dirigeants concernant ces différentes pertes (figure 1.4).

On remarque que les pertes directes constituent la principale source d'inquiétude, en particulier la perte d'exploitation, qui préoccupe fortement les responsables de la cybersécurité. Le coût de remédiation arrive en deuxième position. L'impact sur la réputation est également une préoccupation majeure, bien que les autres coûts indirects, comme le cadre légal renforcé, soient jugés moins prioritaires.

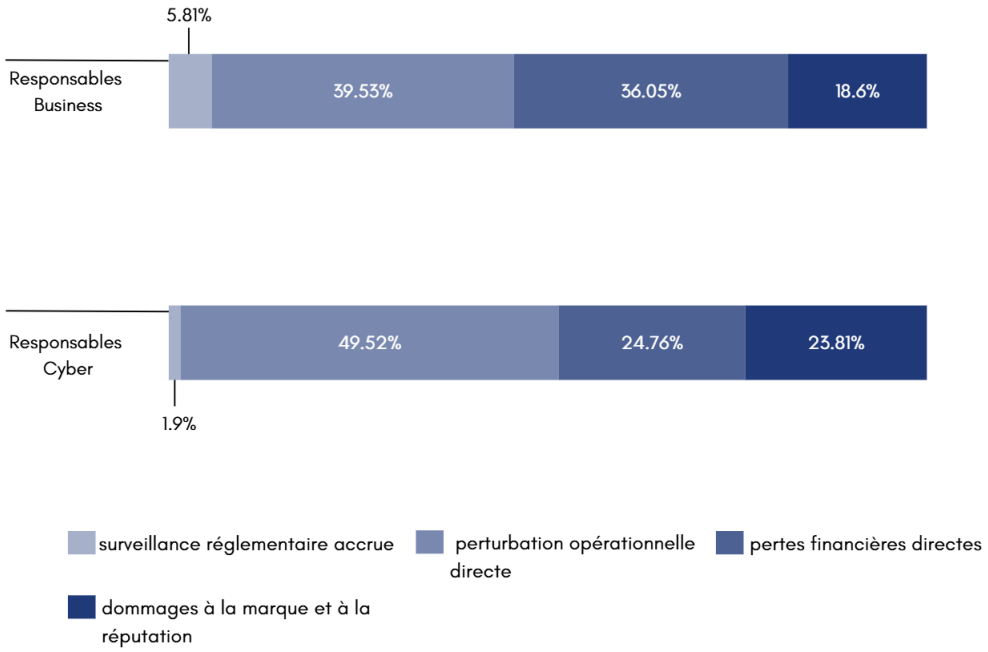


Figure 1.4: Préoccupations principales des dirigeants dans le monde après une cyberattaque – World Economic Forum, 2024b

Il est à noter que l'étude met en avant certaines spécificités régionales : en Asie, en Afrique et en Amérique latine, l'accent est davantage mis sur les coûts financiers, tandis que les dirigeants européens et nord-américains se préoccupent davantage des coûts "opérationnels".

Cette section a présenté les différents types de coûts pouvant être subis par une entreprise à la suite d'un sinistre cyber. Toutefois, toutes les entreprises ne sont pas exposées au même niveau de risque ; certaines sont plus ciblées que d'autres. La section suivante se concentrera sur l'analyse des cibles potentielles et les différentes typologies d'entreprises les plus exposées.

1.1.1.5 Typologies des cibles

Les cibles d'attaques sont de multiples natures. Dans le cadre d'un panorama de la cybermenace, une étude de ANSSI, 2023b sur les attaques à but lucratif en France fournit la répartition des victimes de rançongiciels en 2023. Cette répartition est visible sur la figure 1.5. Il est à noter la présence importante des TPE/ETI/PME, qui représentent 58% des attaques recensées, ce qui montre que les petites structures sont particulièrement vulnérables (et cette dynamique ne se limite pas aux rançongiciels). Néanmoins, la grande diversité des victimes est également remarquable. Des entreprises stratégiques, des associations ou encore des collectivités territoriales ont toutes été touchées par des cyberattaques en 2023, et ce, dans des proportions significatives.

Comme l'ont souligné différentes études, les coûts associés aux cyberattaques ainsi que leur fréquence varient en fonction du secteur d'activité de l'entreprise. En effet, selon une présentation de l'institut FAIR (organisation dédiée à la recherche et au partage de connaissances en management du risque cyber), il existe de fortes disparités dans la fréquence des vols de données selon les secteurs d'activité, comme illustré dans la figure 1.6.

Cependant, cette étude étant basée sur la base de données VERIS (principalement alimentée par des incidents aux États-Unis), ses conclusions ne peuvent être appliquées directement au contexte

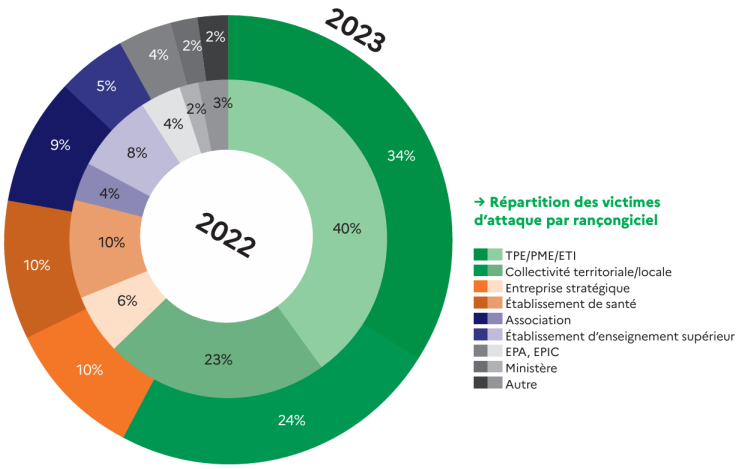


Figure 1.5: Répartition des victimes d'attaques par rançongiciel en 2023 - ANSSI, 2023b

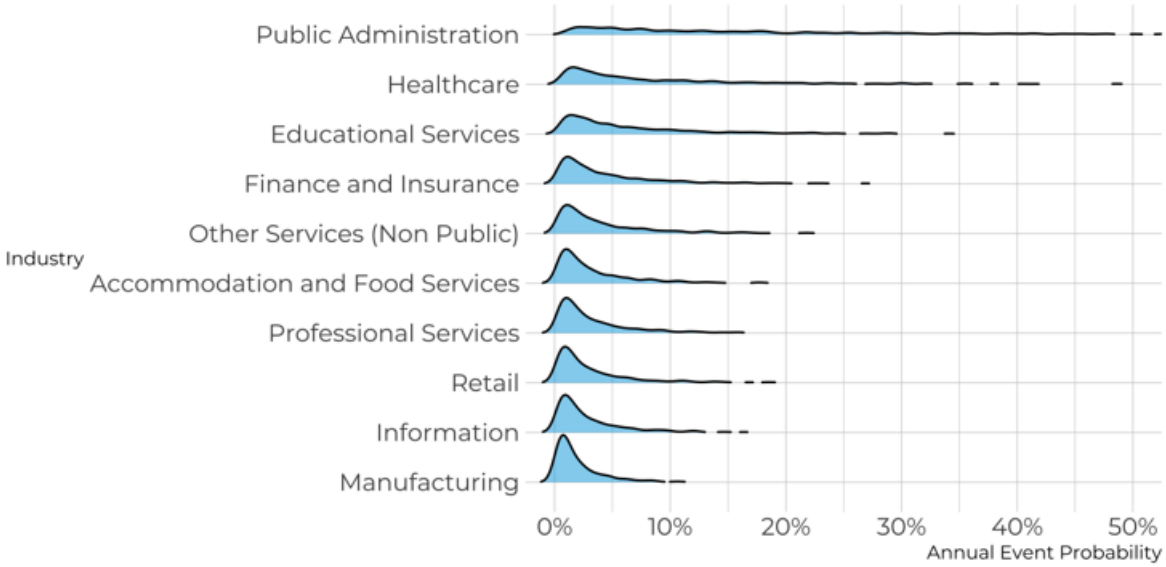


Figure 1.6: Probabilité d'attaque annuelle selon les secteurs - FAIR Institute and EY, 2024

français. Néanmoins, elle met en évidence l'impact du secteur d'activité sur l'exposition au risque cyber.

Au-delà du secteur, le risque d'une attaque dépend également du type de victime. En effet, bien que les petites entreprises subissent des coûts plus faibles en valeur absolue, elles sont aussi les plus exposées, et le coût d'un sinistre cyber représente une part proportionnellement plus importante de leur chiffre d'affaires. Plusieurs études illustrent ce phénomène. À titre d'exemple, dans le contexte français, les chiffres de LUCY (AMRAE, 2023) indiquent qu'en 2022, le coût moyen d'une cyberattaque pour une grande entreprise (entre 50 M€ et 1,5 milliard € de CA) était de 900 000 €, tandis qu'il s'élevait à 266 000 € pour les ETI (entre 10 M€ et 50 M€ de CA) et à 450 000 € pour les entreprises

de taille moyenne (entre 2 M€ et 10 M€ de CA). Cela montre qu'une cyberattaque représente un fardeau bien plus lourd pour les petites structures.

Néanmoins, la survenue de sinistres graves, causant d'importants dégâts (comme illustré précédemment dans les exemples d'attaques, section 1.1.1.1), peut rendre une mauvaise année particulièrement difficile à gérer, même pour une grande entreprise. Ces situations peuvent même être qualifiées de **catastrophes cyber**, et constituent l'une des raisons pour lesquelles certaines compagnies d'assurance hésitent à proposer des produits spécifiques à ce secteur.

D'autres facteurs **internes** influent sur le niveau de risque auquel une entreprise est exposée, tels que son niveau d'hygiène cyber ou son importance stratégique. Ces aspects seront abordés plus en détail ultérieurement dans ce mémoire.

1.1.2 Dans la tête d'un attaquant

Après cette introduction au risque cyber, qui permet d'observer son impact global ainsi que son hétérogénéité tant sur les cibles que sur les coûts et les volontés des attaquants, l'accent sera maintenant mis plus particulièrement sur les méthodes et stratégies d'attaque. Cela permettra de mieux comprendre comment ce risque se décompose en détail et ainsi commencer à poser les bases de l'étude future.

1.1.2.1 Différentes méthodes d'attaques

Lorsqu'il s'agit d'attaques cyber, certains mots comme *DDoS*, *Phishing*, *Malware* reviennent souvent. En effet, les différentes techniques présentées donnent naissance à des attaques de diverses natures, chacune avec des terminologies propres. Dans cette section liée à la cybersécurité, un glossaire des méthodes d'attaque les plus utilisées par les cybermalfaiteurs sera d'abord proposé. Ensuite, des catégories et classes d'attaque seront dénotées, car en cybersécurité, il faut diviser pour mieux régner.

Glossaire Pour établir ce glossaire, il sera fait référence à l'article de Baker, 2024.

Malware Le *Malware* (*"Malicious Software"*) est une méta-catégorie, il correspond à tout code ou programme dont l'objectif est d'endommager un système d'information. Cette méthode peut utiliser des **vulnérabilités** techniques ainsi que des erreurs de configuration pour s'appliquer. Il enveloppe néanmoins un ensemble de méthodes n'ayant pas le même fonctionnement. Parmi elles, les plus connues sont :

- **Le Virus** est un bout de code malveillant qui se greffe sur une application ou un fichier bénin. Par exemple, un fichier Excel infecté par ce type de code. Notons que certaines catégories ci-dessous peuvent prendre la forme d'un virus comme d'un programme à part entière.
- **Le Ransomware** est un logiciel malicieux cryptant les données d'une machine, d'un serveur ou d'une entreprise entière, dont la clé de décryptage est fournie (théoriquement) après le paiement d'une rançon. Une variante est la *Wiper Attack* qui, au lieu de crypter les données, les détruit définitivement.
- **Le *Rootkit*** est un ensemble de logiciels permettant à un attaquant de prendre le contrôle (l'accès *"root"*) d'une machine ou d'un réseau entier.
- **Les Vers (*Worms*)** sont des programmes qui se dupliquent et se propagent sur un réseau.

Phishing Le *Phishing* est une cyberattaque utilisant les moyens de communication courants pour inciter la cible à partager des informations sensibles ou à télécharger un logiciel malveillant.

Ingénierie Sociale L'attaquant manipule la cible pour obtenir des documents, de l'argent ou des informations en se faisant passer pour une entité de confiance.

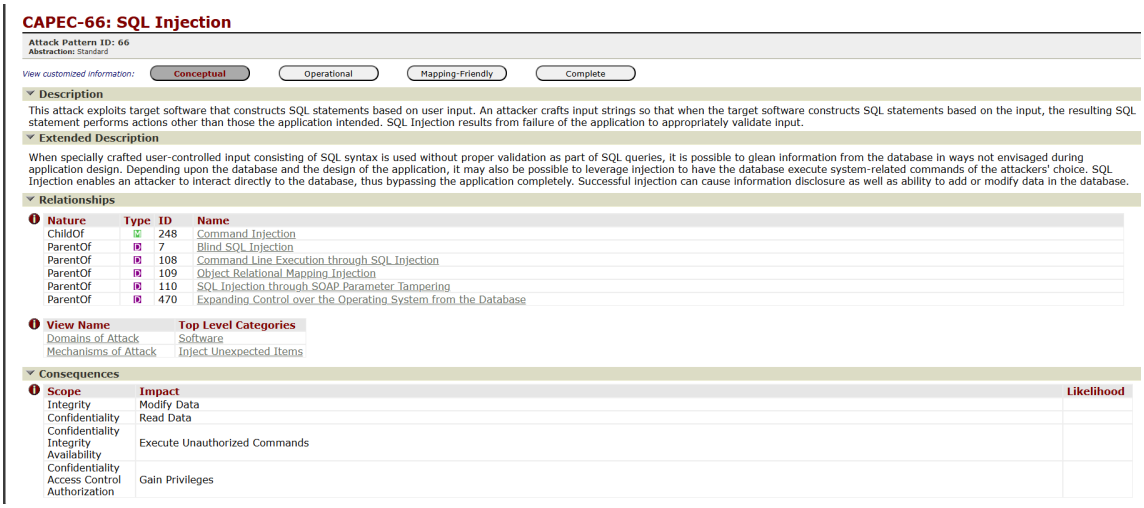
Injection de code L'objectif est d'utiliser des failles du réseau pour injecter du code malveillant.

Le Man in the Middle L'attaquant intercepte les communications entre la victime et la ressource ciblée.

Les attaques DoS ou DDoS L'objectif est de rendre indisponible un service en surchargeant son réseau de requêtes.

Supply Chain Attack L'attaquant infecte un fournisseur de services pour propager un logiciel malveillant aux utilisateurs.

Ce glossaire présente un ensemble de méthodes souvent évoquées lorsqu'il est question d'attaques informatiques. Une classification plus détaillée est disponible via MITRE, qui propose la liste CAPEC (extitCommon Attack Pattern Enumeration and Classification) (MITRE, 2019).



CAPEC-66: SQL Injection

Attack Pattern ID: 66
Abstraction: Standard

View customized information: Conceptual | Operational | Mapping-Friendly | Complete

Description

This attack exploits target software that constructs SQL statements based on user input. An attacker crafts input strings so that when the target software constructs SQL statements based on the input, the resulting SQL statement performs actions other than those the application intended. SQL Injection results from failure of the application to appropriately validate input.

Extended Description

When specially crafted user-controlled input consisting of SQL syntax is used without proper validation as part of SQL queries, it is possible to glean information from the database in ways not envisaged during application design. Depending upon the database and the design of the application, it may also be possible to leverage injection to have the database execute system-related commands of the attackers' choice. SQL Injection enables an attacker to interact directly to the database, thus bypassing the application completely. Successful injection can cause information disclosure as well as ability to add or modify data in the database.

Relationships

Nature	Type	ID	Name
ChildOf	V	248	Command Injection
ParentOf	B	7	Blind SQL Injection
ParentOf	B	108	Command Line Execution through SQL Injection
ParentOf	B	109	Object Relational Mapping Injection
ParentOf	B	110	SQL Injection through SOAP Parameter Tampering
ParentOf	B	470	Expanding Control over the Operating System from the Database

View Name	Top Level Categories
Domains of Attack	Software
Mechanisms of Attack	Inject Unexpected Items

Consequences

Scope	Impact	Likelihood
Integrity	Modify Data	
Confidentiality	Read Data	
Confidentiality Integrity Availability	Execute Unauthorized Commands	
Confidentiality Access Control Authorization	Gain Privileges	

Figure 1.7: CAPEC-66 Injection SQL - MITRE CAPEC

Une description détaillée du mode opératoire est fournie ainsi qu'un arbre de relation avec des méthodes de plus en plus précises. Cette ressource permet d'avoir une arborescence précise des types de menaces et d'associer à chaque attaque une conséquence en termes d'intégrité, de confidentialité et de conformité.

1.1.2.2 Menaces Humaines et Techniques

Parmi la multitude de types d'attaques introduits ci-dessus, deux grandes catégories peuvent être distinguées. Celles utilisant des **failles techniques**, se basant sur des vulnérabilités logicielles, des mauvaises configurations de pare-feu, de site web, de ports, etc., comme les *injections de code*, certains *malwares* ou le *DDoS*, et celles exploitant les **failles humaines**, telles que le *phishing* et l'*ingénierie sociale*.

La composante humaine est un facteur clé à ne pas sous-estimer, car elle représente une part significative des incidents de sécurité en entreprise. D'après CyberMalveillance, 2022, 13 % des recherches d'assistance pour les entreprises sont dues au *phishing*. De plus, selon Verizon, 2023, environ 82 % des violations de données impliquent un facteur humain, tel que le phishing ou des erreurs de configuration. Cela souligne l'importance de la formation et de la vigilance des employés dans la sécurité informatique des entreprises.

Ainsi, pour une entreprise, se concentrer sur un type de menace sans s'intéresser à l'autre ne peut pas fonctionner. L'enjeu est à la fois d'avoir une équipe consciente du risque et proactive, par des entraînements de phishing par exemple, mais aussi un processus de cyberdéfense bien fonctionnel.

1.1.2.3 Une méthodologie d'attaque

Un attaquant n'utilise rarement qu'une seule des méthodes et des outils proposés pour pénétrer une cible. En effet, dans la plupart des cas, l'exploitation d'une seule faille, qu'elle soit technique ou humaine, ne suffit pas à mettre à genoux l'ensemble de la défense de l'entreprise.

L'attaquant suit un plan d'attaque précis, utilisant des failles comme éléments essentiels à chaque étape de sa progression. À chaque phase, il applique des méthodes spécifiques qui lui permettent de mener à bien son attaque. L'organisation MITRE a proposé un modèle d'attaque largement adopté par de nombreux acteurs du domaine. Ce modèle, illustré à la figure (1.8), décrit les différentes étapes d'une attaque. Cette figure souligne que ce processus est itératif et requiert de l'attaquant l'utilisation d'une variété de méthodes pour atteindre ses objectifs.

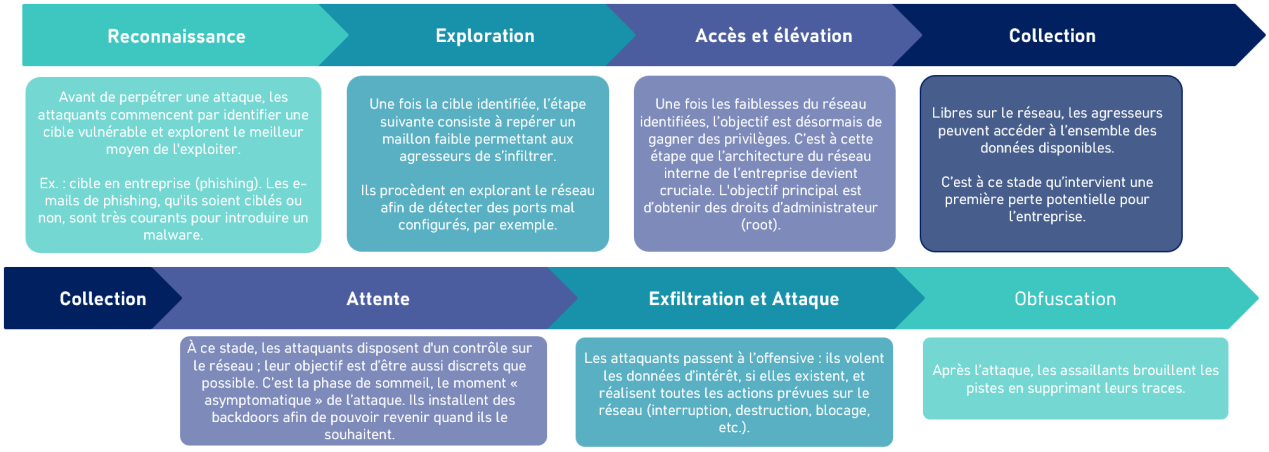


Figure 1.8: Étapes d'un plan d'attaque

Dans l'objectif de rendre la cybersécurité encore plus normée et exhaustive face aux menaces, MITRE a créé la matrice ATT&CK. Cette matrice permet d'étudier en détail l'ensemble des méthodes et des sous-méthodes associées à une étape d'un processus d'attaque, visibles sur la figure (1.8). Très complète, elle permet d'avoir une vision des possibles méthodes et, combinée avec une connaissance des faiblesses de l'entreprise ainsi que des menaces et vulnérabilités existantes, elle permet également de déduire les chemins potentiels qu'un attaquant pourrait emprunter. Certains sites, souvent de fournisseurs de solutions de cyberdéfense, permettent d'ailleurs de visualiser en temps réel les menaces et vulnérabilités présentes dans le monde sur une matrice ATT&CK. Ceci est visible, par exemple, sur le site de Fortiguard.

Cette matrice permet également de distinguer plusieurs phases dans l'attaque. Comme le montre

Figure 1.9: MITRE ATT&CK et position de l'attaquant

la figure (1.9), l'attaque se déroule en deux étapes, indiquées en jaune.

Dans un premier temps, l'attaquant est en dehors de l'entreprise et va tenter de s'y infiltrer par l'une des différentes méthodes présentées en détail dans la matrice et de manière générale dans notre glossaire. Pour minimiser les risques d'intrusion, l'entreprise cible doit avoir une surface d'attaque bien protégée — des remparts solides — en particulier dans les deux catégories mentionnées dans la section (1.1.2.2).

Dans un second temps, l'attaquant est à l'intérieur du réseau ou du système et entame une phase de prise de pouvoir et de mouvement latéral pour identifier des cibles d'intérêt. Ici, la capacité de réponse de l'entreprise est cruciale, notamment en ce qui concerne la détection de l'attaque et la rapidité de sa réaction.

Enfin, lorsque l'attaquant a acquis les privilèges nécessaires, il passe à l'attaque. C'est alors la préparation en amont de l'entreprise qui influe sur les coûts qu'elle subira à la suite de l'attaque, notamment en termes de sauvegardes et d'infrastructure résiliente.

L'examen de la distinction entre le risque humain et technique, ainsi que la division temporelle d'une attaque cyber, a permis de mettre en lumière les différents enjeux de la cybersécurité et de la cyberdéfense. En intégrant l'ensemble de ces modules dans un modèle actuariel, il serait possible d'améliorer la segmentation du risque.

Pour conclure cette partie sur l'étude détaillée de la menace cyber, une présentation précise des vulnérabilités logicielles sera proposée. En effet, la majorité des méthodes techniques reposent sur l'exploitation de ces vulnérabilités pour s'infiltrer et endommager les systèmes cibles. À cette occasion, des données précieuses seront également présentées pour la quantification évolutive du risque cyber.

1.1.3 Présentation des données de vulnérabilité

Aucun système n'est totalement infaillible, et c'est sur cette réalité que reposent de nombreuses cyberattaques. Les attaquants cherchent ainsi à exploiter ces vulnérabilités pour en tirer un avantage. Comme exposé dans la section précédente, ces vulnérabilités peuvent se manifester à travers des **failles humaines**, telles que celles décrites dans (1.1.2.2), ou des **failles techniques**. En effet, chaque système est conçu, développé, mis à jour et utilisé par des individus, ce qui introduit des risques inhérents.

1.1.3.1 Définition des vulnérabilités dans les systèmes d'information

Il existe, dans l'ensemble des systèmes, applications et services utilisés par les entreprises (et plus généralement dans l'ensemble de la société), des défauts dans la conception, l'implémentation ou la configuration de ceux-ci. Appelés **vulnérabilités**, ces défauts constituent une porte d'entrée potentielle pour l'attaquant. Elles n'ont néanmoins pas toutes la même *exploitabilité*. Si nous prenons l'allégorie du vol dans une maison :

- Une porte ouverte est une vulnérabilité facilement exploitable, et ce, par n'importe qui.
- Une fenêtre ouverte au premier étage demande plus d'efforts, mais lorsqu'on a une échelle, c'est une tâche réalisable.
- Enfin, un défaut de conception dans la serrure de la porte, permettant à un expert de l'ouvrir sans clé, est une vulnérabilité bien plus difficile à exploiter.

Il en va de même pour les vulnérabilités informatiques. Remplaçons la porte par une vulnérabilité facilement exploitable, comme un défaut de configuration de ports sur un serveur ouvert à l'internet, permettant d'entrer directement dans celui-ci. L'échelle représente un logiciel malveillant déjà existant, qu'il est facile de se procurer sur les marchés noirs (comme mentionné précédemment avec le *"Cybercrime-as-a-service (CAAS)"* (1.1.1.3)). Le défaut de conception technique, quant à lui, peut être vu comme une vulnérabilité dont aucune exploitation n'a encore été réalisée et qui demande un grand niveau d'expertise pour être exploitée.

De plus, comme évoqué dans la section (1.1.2.3), le processus d'attaque comporte différentes étapes. L'existence d'une faille n'implique pas son exploitabilité directe par l'attaquant. Sa facilité d'exploitation n'est donc pas suffisante pour expliquer la vulnérabilité globale de l'entreprise. En reprenant notre exemple de la maison, le fait que la porte soit ouverte n'est peut-être pas directement exploitable si une barrière entoure la maison. Il lui faudra ainsi d'abord trouver un moyen de passer à travers pour atteindre la porte.

Ainsi, si une faille se trouve à l'intérieur, mais que la surface d'attaque est parfaitement protégée, rien ne pourra arriver... ceci n'est malheureusement jamais le cas. Ce concept sera repris lors de la modélisation effectuée dans le chapitre 2, en particulier avec les idées développées en section (2.2.2.1).

L'exploitation de la faille ne donne pas forcément non plus le même résultat. Une brèche dans le mur de la maison permet de voir ce qu'il y a à l'intérieur, mais pas d'y rentrer.

Un attaquant prendra donc en compte la difficulté de l'**exploitabilité** de la vulnérabilité, mais aussi son **impact** pour le choix de son utilisation.

Ainsi, une vulnérabilité dans un logiciel (ou toute autre pièce du puzzle informatique) grandement utilisé peut rapidement devenir problématique. Du point de vue assurantiel, ces informations sont particulièrement intéressantes, car la connaissance des vulnérabilités permettra non seulement de mieux quantifier le risque, mais aussi d'accompagner l'assuré.

Pour suivre au mieux la découverte et la quantification de la dangerosité d'une vulnérabilité, différentes méthodes et systèmes de notation ont été mis en place. Leur compréhension est un support essentiel pour la thématique de ce mémoire.

1.1.3.2 Les CVE, piliers de la sécurité informatique moderne

Les CVE (*Common Vulnerabilities and Exposures*) sont des identifiants uniques attribués à des vulnérabilités spécifiques dans des logiciels ou des systèmes informatiques. Ceux-ci permettent aux

différents acteurs de disposer d'un moyen normé de partage de l'information. Ce programme, et la nomenclature associée, sont gérés par MITRE (dont il a déjà été question en (1.1.2.3) pour la matrice ATT&CK) et sont subventionnés par la CISA (*Cybersecurity and Infrastructure Security Agency*).

Comme présenté dans Red Hat, 2021, certaines sociétés sont déléguées à l'attribution d'un code CVE à une vulnérabilité. Ces organismes font partie du CNA (*CVE Numbering Authority*) et seuls eux peuvent déclarer directement une CVE. Il existe une centaine de membres, parmi lesquels figurent des acteurs majeurs du monde informatique (Microsoft, IBM, etc.), ainsi que des entreprises de sécurité et des laboratoires de recherche.

L'ensemble des règles d'écriture et de gestion du programme CVE peut être consulté sur CVE Program, 2024.

En particulier, MITRE alloue chaque année des blocs d'identifiants à chaque CNA. Lorsqu'un de ces organismes souhaite déclarer une vulnérabilité sur son produit ou sur un produit tiers, il la déclare avec l'un des identifiants attribués (dans l'ordre croissant).

CVE-2017-0144

Figure 1.10: Exemple de nom de CVE

Comme présenté sur la figure (1.10), la nomenclature d'une vulnérabilité CVE est fixe. Elle se compose du terme "CVE", suivi de l'année de la découverte ou de l'attribution initiale de la vulnérabilité, et enfin d'un numéro unique pour chaque vulnérabilité au sein de l'année. L'exemple donné correspond à la 144e vulnérabilité répertoriée pendant l'année 2017. Elle fait partie des vulnérabilités exploitées lors de l'attaque WannaCry (voir le paragraphe (1.1.1.1) pour un rappel de l'attaque WannaCry).

1.1.3.3 Une évolution de la base CVE

Les CVE sont aujourd'hui utilisées par l'ensemble de la communauté cyber et, à l'image de l'utilisation croissante des outils informatiques, le nombre de vulnérabilités divulguées croît lui aussi d'année en année. La figure (1.11) illustre cette tendance, montrant une augmentation continue du nombre de vulnérabilités CVE depuis 1999.

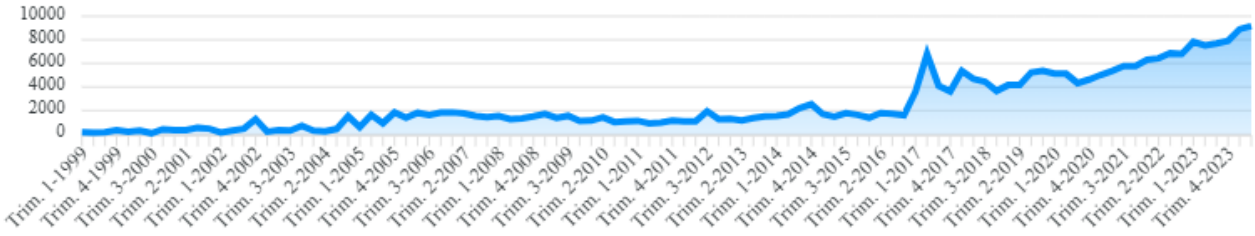


Figure 1.11: Nombre de vulnérabilités par trimestre de 1999 à 2023

Cette tendance à la hausse des vulnérabilités divulguées n'est peut-être pas aussi négative qu'elle en a l'air. Certes, elle s'inscrit dans une augmentation générale du nombre de failles connues, en raison de l'accroissement des services et logiciels utilisés. Cependant, elle reflète également une nouvelle volonté de la communauté professionnelle et de cybersécurité de favoriser le partage d'informations.

En effet, il n'existe aucune obligation légale pour les fournisseurs de services qui ne sont pas en *Open Source* (code accessible à tous) de partager les vulnérabilités qu'ils auraient repérées sur leurs

systèmes. Cela repose donc sur la bonne volonté des différents acteurs.

Cette évolution des pratiques est analysée dans une excellente étude de Cyentia Institute, 2023, que le lecteur intéressé est invité à consulter.

Comme expliqué par RedHat, 2023, en raison de l'augmentation du nombre de CVE publiées chaque mois, les professionnels de la cybersécurité ne peuvent pas s'intéresser à toutes les vulnérabilités. Ils doivent donc prioriser. Ainsi, il existe une volonté de quantifier le risque lié à une vulnérabilité afin de mieux allouer les **ressources et le temps disponibles**.

1.1.3.4 Les différentes composantes de la base CVE

Ce programme et cette nomenclature permettent ainsi d'obtenir une liste de vulnérabilités à laquelle peut s'ajouter un grand nombre d'informations, facilitant une meilleure analyse du risque lié à chacune des vulnérabilités. Ces données sont parfois disponibles directement dans la déclaration sur GitHub du programme CVE, où se trouve la liste des différentes CVE pour toutes les années. Elles peuvent également être complétées par certaines sources externes, offrant ainsi davantage d'informations.

En particulier, la NVD (*National Vulnerability Database*), un répertoire du gouvernement américain, fournit une base de données enrichie de certaines informations. Cette base sera utilisée dans le chapitre 3 lors de la création d'un portefeuille fictif (voir section 3.1.3.3). D'autres ressources en *Open Source* seront citées dans la suite de ce mémoire pour apporter une valeur ajoutée à la base.

Par la suite, les différentes informations associées au numéro CVE seront détaillées. Celles-ci apportent une plus-value pour la gestion moderne du risque cyber.

La CISA KEV En plus de la base CVE, la CISA (*Cybersecurity and Infrastructure Security Agency*) recense dans la base KEV (*Known Exploited Vulnerabilities*) l'ensemble des vulnérabilités CVE connues pour être ou avoir été exploitées, et renvoie vers les conseils du fournisseur pour mitiger la vulnérabilité dans les systèmes de l'entreprise, par exemple en mettant à jour le logiciel concerné. L'ensemble des vulnérabilités de cette base ont donc été exploitées par des attaquants lors d'un processus d'attaque. Elle est donc composée de vulnérabilités qui doivent être étudiées avec attention.

Les CPE, une normalisation des noms de programmes Si la volonté est de fournir une liste utile dans le cadre des CVE, la première information importante est de savoir quels sont les services ou programmes touchés.

Or, un simple nom de programme ne suffit pas. Les vulnérabilités se trouvent souvent sur des versions précises et sont généralement rapidement corrigées ("*patchées*") dans une nouvelle version. Il est donc nécessaire de disposer d'une codification détaillée du service utilisé. Comme présenté dans Sanguino and Uetz, 2017, le standard CPE (pour *Common Platform Enumeration*), proposé par la NIST, permet en théorie d'adopter une méthode unique pour nommer un programme et ainsi faciliter la transmission d'informations entre différents acteurs. Parmi les informations fournies dans le CPE figurent : le fournisseur, le produit, la version, l'édition, la langue ainsi qu'un grand nombre d'autres éléments permettant d'identifier de manière unique un produit.

En prenant l'exemple de la version 122.0 de Firefox, illustré dans la figure 1.12, une écriture normalisée au format XML est observée. Par exemple, la première ligne présente le nom de la CPE sous la forme *cpe:/part:fournisseur:produit:version*.

Cet exemple provient du dictionnaire mis à jour par le NIST, qui offre une liste quasi exhaustive de l'ensemble des produits informatiques disponibles sur le marché.

```
<cpe-item name="cpe:/a:mozilla:firefox:122.0">
    <title xml:lang="en-US">Mozilla Firefox 122.0</title>
    <references>
      <reference href="https://www.mozilla.org/en-US/security/
        advisories/mfsa2024-01/">Advisory</reference>
    </references>
    <cpe-23:cpe23-item name="cpe:2.3:a:mozilla:firefox:122.0:*:*:*
      :*:*:*:*"/>
  </cpe-item>
```

Figure 1.12: Exemple de code CPE pour Firefox version 122.0.

Ainsi, dans la description d'une CVE, se trouve une liste de tous les programmes affectés au format CPE, comme illustré par l'exemple donné dans la figure 1.13.

Figure 1.13: CPE de la CVE-2024-24919 sur la base NVD.

En théorie, toute CVE devrait avoir une liste de CPE associée et chaque identifiant CPE devrait être unique. Malheureusement, comme expliqué dans Sanguino and Uetz, 2017, ce n'est pas le cas. Certaines CVE existent sans CPE, bien qu'un CPE pour ce produit soit présent dans le dictionnaire, et certains identifiants ne sont pas uniques. Il reste donc un travail de normalisation à poursuivre dans le domaine de la cybersécurité afin d'assurer une automatisation robuste des processus.

La description, forte valeur ajoutée, faible automatisation En plus des identifiants des produits, les vulnérabilités CVE présentes sur la base NVD disposent d'une description détaillée. En reprenant l'exemple de la figure 1.13, la description illustrée dans la figure 1.14 peut être consultée. Elle révèle, par exemple, la possible exploitation de la faille par un attaquant, ainsi que l'existence d'une remédiation sous la forme d'une mise à jour corrective.

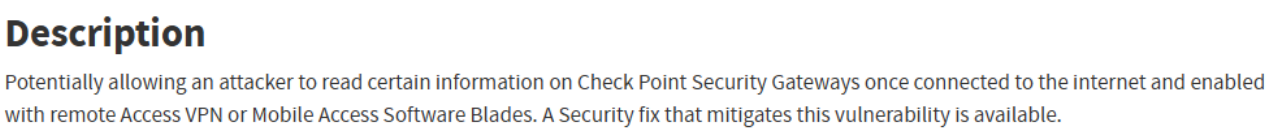


Figure 1.14: Description de la CVE-2024-24919 sur la base NVD.

Cependant, cette description est rédigée en langage naturel, elle n'est pas normée, rendant ainsi son utilisation complexe pour une automatisation de la quantification du risque, ce qui serait pourtant utile dans un contexte assurantiel. Les professionnels de la cybersécurité sont également conscients de ce problème. En effet, une normalisation est nécessaire pour une mitigation optimale du risque, en s'appuyant sur des méthodes où la donnée fiable et normalisée est essentielle.

La description représente une forte valeur ajoutée mais nécessite un traitement détaillé. Certains

travaux de recherche tentent néanmoins d'en tirer parti pour mieux quantifier le risque. Par exemple, l'étude menée dans Kanakogi et al., 2021 a pour objectif d'associer, à partir de la description, un identifiant CAPEC (présenté dans la section 1.1.2.1) en utilisant des modèles de langage.

Les scores, un outil de qualification et peut-être de quantification L'analyse des vulnérabilités techniques sera désormais abordée à travers différents indicateurs. Comme précisé en conclusion de la section (1.1.3.3), dans le monde de la cybersécurité, les ressources et le temps sont limités. Disposer d'indicateurs évolutifs et précis d'une menace, ici provenant d'une vulnérabilité, est donc essentiel pour allouer le budget "défense" aux problématiques réellement importantes.

Cette partie présentera en détail l'indicateur **CVSS** (pour *Common Vulnerability Scoring System*), ainsi que d'autres indicateurs utilisés dans le domaine de la cybersécurité de manière plus succincte.

Le Score CVSS est un système standardisé d'évaluation de la gravité des vulnérabilités CVE. Plusieurs versions existent, la dernière en date étant CVSS 4.0, publiée en novembre 2023. Comme expliqué par Wikipedia, 2024, l'objectif de ce score est de permettre une évaluation objective et mesurable du risque porté par une vulnérabilité afin de prioriser et allouer les ressources à la mitigation des problématiques les plus urgentes. Le jugement des différentes métriques qui composent ce score est effectué par des professionnels de la cybersécurité et est publié quelque temps après la publication de la CVE.

Existant depuis 2005, l'évolution de ce score est désormais gérée par FIRST, un organisme international qui développe et promeut des standards pour la gestion des vulnérabilités et des incidents de sécurité. L'objectif de ce score est de fournir une note de 0 à 10 pour chaque CVE en fonction de sa dangerosité. Cette note comprend une composante fixe et une composante évolutive, permettant ainsi de suivre l'évolution de la dangerosité de la faille au fil du temps.

L'utilisation de ce type de score dans un contexte de tarification assurantielle sera abordée dans ce mémoire au chapitre 2.

Comme mentionné, la nouvelle version de CVSS date de novembre 2023. Elle sera examinée en détail, mais il est important de noter qu'elle n'est pas encore parfaitement implémentée (dans NVD par exemple), et qu'il faudra attendre pour pouvoir bénéficier de l'ensemble des informations qu'elle apporte. L'explication détaillée du score CVSS 3.1 est disponible sur le site de FIRST.

Notre explication se basera sur le document de spécification du score CVSS 4.0 (FIRST, 2023).

Comme le montre la figure 1.15, le score CVSS est divisé en plusieurs catégories.

Chacune de ces catégories est divisée en métriques, et pour chaque métrique, il existe un ensemble de **valeurs possibles**. Par exemple, pour la métrique *User Interaction* dans le groupe des métriques de base, trois valeurs sont possibles : (N) pour *None*, signifiant que l'utilisateur n'a pas besoin d'agir pour que la faille soit exploitée, (P) pour *Passive*, signifiant que l'utilisateur (victime) a une faible implication dans l'action, et (A) pour *Active*, signifiant que la faille nécessite une interaction active de l'utilisateur cible pour être exploitée. Une vulnérabilité avec (N) dans cette catégorie est donc plus dangereuse qu'une faille (P) ou (A), toutes choses égales par ailleurs.

Toutes les métriques et leurs valeurs ne seront pas explicitement détaillées ici. Cependant, l'ensemble des informations peut être consulté dans FIRST, 2023. La suite de ce paragraphe présentera néanmoins la signification des différentes catégories.

- **Le groupe de métriques de Base** correspond aux métriques intrinsèques à la vulnérabilité. Ces métriques sont invariantes dans le temps et indépendantes du contexte de l'entreprise. Ce groupe est divisé en deux sous-groupes :

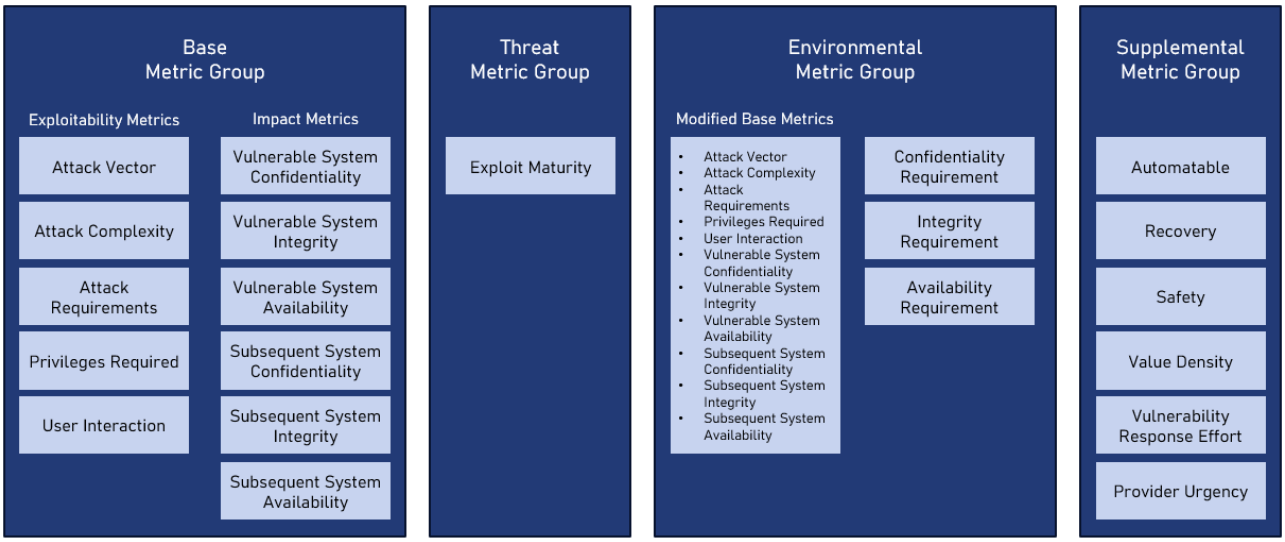


Figure 1.15: Score CVSS 4.0.

  - **Exploitabilité** : métriques évaluant les propriétés du système vulnérable. Par exemple, la métrique *User Interaction* analyse la facilité d'exploitation de la vulnérabilité.
  - **Impact** : métriques évaluant les conséquences de la vulnérabilité sur le système concerné (*Vulnerable System*) et sur les systèmes connexes (*Subsequent System*).

  Ce groupe rappelle le constat effectué en (1.1.3.1) sur la double importance de l'exploitabilité et de l'impact pour l'attaquant lors du choix de l'utilisation d'une vulnérabilité. Les deux sous-groupes seront d'ailleurs exploités pour la création du modèle lors du chapitre 2.
- **Le groupe des métriques des Menaces** correspond aux métriques évolutives dans le temps mais indépendantes des caractéristiques du système. Ce groupe ne contient qu'une seule métrique, *Exploit Maturity*, qui évalue la maturité des outils disponibles pour exploiter la vulnérabilité.
- **Le groupe des métriques Environnementales** inclut des métriques dépendantes du contexte de l'entreprise. Ce groupe modifie les métriques de base et introduit trois nouvelles métriques relatives à la Confidentialité, l'Intégrité et la Disponibilité, qui sont les piliers dont il était question lors de la partie (1.1.1)).
- **Le groupe de métriques Supplémentaires** fournit un contexte et mesure des attributs extrinsèques d'une vulnérabilité. Ces métriques, définies par l'utilisateur, n'affectent pas le score final du CVSS, mais permettent une analyse plus adaptée au contexte spécifique.

Comme mentionné précédemment, ces métriques sont catégorielles et non numériques. L'ensemble des informations est alors compilé dans un vecteur associant à chaque métrique une valeur. Par exemple, la figure 1.16 illustre un vecteur CVSS où "AV:N" correspond à la métrique *Attack Vector* définie à N.

CVSS:4.0/AV:N/AC:L/AT:N/PR:H/UI:N/VC:L/VI:L/VA:N/SC:N/SI:N/SA:N

Figure 1.16: Exemple de vecteur CVSS.

Pour transformer cela en score, FIRST a mis en place des "macros vecteurs", qui permettent d'approximer l'ensemble des possibilités dans un univers de choix plus restreint, afin de générer une note entre 0.1 et 10. Cette approche a été introduite avec CVSS 4.0, alors que CVSS 3.1 utilisait une formule directe de calcul du score.

Le Score EPSS D'autres scores existent avec le même objectif de prioriser les vulnérabilités à étudier. Parmi eux, le score EPSS (pour *Exploit Prediction Scoring System*), créé en 2019, est né du constat que la mitigation d'une vulnérabilité grave pouvant être exploitée de manière automatique augmente la sécurité des systèmes de 62,81 %, tandis que la mitigation d'une vulnérabilité grave sans kits n'améliore la sécurité que de 3,2 % (Alay-Eddine, 2022). L'idée derrière ce score est de répondre à la question "La vulnérabilité sera-t-elle réellement exploitée dans les 12 prochains mois ?" (Alay-Eddine, 2022). Pour ce faire, douze caractéristiques ont été sélectionnées et un modèle de régression a été entraîné sur la base CVE de MITRE afin d'obtenir les poids de chaque variable explicative. La méthodologie ainsi que l'ensemble des variables d'intérêt du score EPSS sont disponibles dans Jacobs et al., 2023. Dans l'univers de la cybersécurité, ce score ne fait pas encore l'unanimité, notamment en raison de son absence de neutralité dans le choix des variables (Alay-Eddine, 2022). Cette méthode permet néanmoins d'avoir un aperçu rapide et automatique du niveau d'exploitabilité d'une vulnérabilité sans attendre l'avis d'un expert.

D'autres scores D'autres entités ont développé leurs propres scores de vulnérabilité pour mieux répondre à leurs besoins spécifiques. Par exemple, l'assureur cyber Coalition a créé son propre score, le CESS (*Coalition Exploit Scoring System*), adapté aux besoins de ses clients (Coalition, 2023).

Pour conclure cette partie sur l'évaluation du risque lié à une vulnérabilité, il est clair qu'un effort constant est déployé dans le domaine de la cybersécurité pour affiner la qualification des risques, tant par la mise à jour des méthodes existantes, comme le passage de CVSS 3.1 à 4.0, que par la création de nouvelles approches, telles que le score EPSS. Cette dynamique reflète une volonté d'extraire davantage de valeur ajoutée des déclarations de vulnérabilités. Ces informations constituent un véritable atout pour le secteur de l'assurance, car elles permettent une meilleure évaluation de l'environnement dans lequel évolue l'assuré, offrant ainsi une vision plus précise de son exposition au risque. Toutefois, ces données restent encore sous-exploitées à ce jour.

1.1.4 La gestion du risque cyber en entreprise

Pour conclure ce parcours dans le monde du risque cyber, l'attention se portera sur les enjeux de sa gestion en entreprise. Pour ce faire, certains concepts discutés précédemment seront approfondis afin de donner une vue d'ensemble et permettre une transition vers le monde de l'assurance ainsi que vers la thématique principale de ce mémoire, qui est l'utilisation des données cyber pour quantifier dynamiquement le risque d'une entreprise.

1.1.4.1 Des entreprises peu préparées

Le risque cyber est complexe, évolutif et particulièrement difficile à contrer. Avec l'émergence de l'intelligence artificielle, les entreprises redoutent une intensification de ce danger (World Economic Forum, 2024b), notamment en raison des techniques de *DeepFakes* et de génération de texte, qui rendent l'exploitation des failles humaines encore plus redoutable et efficace. Le sondage de World Economic Forum, 2024b dépeint un monde de l'entreprise qui se sent en faible confiance face à l'avenir du cyber. Un manque de préparation est particulièrement ressenti par les entreprises à faible et moyen revenu, parmi lesquelles seulement 49,2% estiment avoir les connaissances nécessaires pour atteindre leur objectif de cybersécurité en 2024. Cette prise de conscience de la part des petits acteurs est récente. En 2022, selon le même sondage, 94,7% d'entre eux pensaient avoir ces connaissances.

Cette information s'inscrit en parallèle de la remarque faite sur les PME-ETI dans la partie (1.1.1.5), où il a été observé que ces entreprises sont également les plus touchées en termes de ratio de chiffre d'affaires. Une précarité cyber de certains acteurs économiques de la société se dénote alors. Ce sentiment est détaillé dans une étude de CrowdStrike, 2021. La figure (1.17) présente les chiffres clés de cette étude.

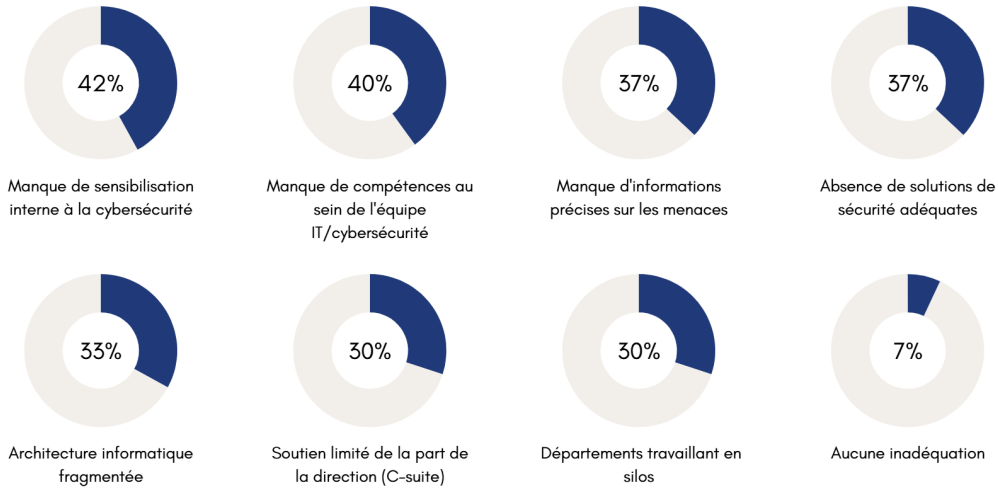


Figure 1.17: Sondage sur l'inadéquation des entreprises au risque cyber selon CrowdStrike, 2021.

Le constat est ici double. Tout d'abord, 42% des entreprises souffrent d'un manque de sensibilisation interne à la cybersécurité, avec des formations insuffisamment développées pour les employés. Par ailleurs, une impression de déficit de connaissances techniques et d'"intelligence de la menace" est également présente chez les participants. Précédemment, la distinction entre les failles **humaines** et **techniques** a été évoquée. Il apparaît ici que les entreprises semblent déficientes dans la maîtrise de ces deux types de vulnérabilités.

Il convient donc de conclure sur la nécessité cruciale d'accompagner les entreprises, en particulier les petites et moyennes entreprises, qui disposent de ressources encore plus limitées pour développer leur résilience et mettre en place une véritable *Threat Intelligence*. Dans le cadre assurantiel, cela suggère une évolution vers une assurance plus polyvalente, qui ne se contente pas uniquement de couvrir les risques mais intègre également des actions de prévention et un soutien continu dans la gestion des cybermenaces.

1.1.4.2 Une existence de la gestion du risque

Néanmoins, une réelle intelligence de la menace existe du côté cybersécurité, avec des outils tels que la matrice ATT&CK, les bases CVE, CAPEC, etc.

De la connaissance du risque via la *Threat Intelligence*, à la prévention et à la mise en place de pratiques renforçant la résilience, en passant par la gestion de crise, un large éventail de connaissances est déjà disponible et pourrait contribuer à mieux préparer les entreprises face aux menaces cyber. Des gestes simples et des pratiques faciles à mettre en œuvre existent pour réduire significativement le risque, notamment pour les petites et moyennes entreprises, où les solutions sont souvent peu coûteuses et rapidement bénéfiques.

1.1.4.3 Un besoin d'assurance

Le risque cyber peut causer des dommages parfois conséquents aux entreprises (voir section 1.1.1.5), et il peut être difficile pour elles de s'en relever sans aide. La gestion de leur risque ne suffit pas à combler les pertes. Il est donc nécessaire d'assurer un partage du risque dans ce domaine et de favoriser une mutualisation entre acteurs afin d'évoluer dans un environnement résilient face à la menace cyber.

L'assurance constitue ainsi un levier majeur et essentiel pour une bonne évolution de la gestion du risque cyber. Néanmoins, comme cela sera examiné, cette assurance demeure hétérogène et encore en cours de développement.

1.2 Assurance et Réglementation

Dans la partie précédente, l'intérêt s'est porté sur le risque cyber dans son cadre général. L'évolution de ce risque, certaines de ses spécificités, ainsi qu'une analyse de la méthodologie d'une attaque du point de vue de l'attaquant ont été présentées. Une analyse des vulnérabilités techniques a ensuite été menée. Il en a été conclu que certaines entreprises sont extrêmement vulnérables et que le besoin d'assurance se fait sentir.

Dans cette partie, nous étudierons plus en détail l'aspect assurantiel du risque. Une rapide présentation des différentes réglementations entourant ce domaine sera effectuée en premier lieu. Les diverses offres assurantielles du marché seront ensuite présentées. Les limites de l'assurance cyber actuelle, tant en termes de modélisation que de manque de données et de non-assurabilité, seront ensuite détaillées. Enfin, le principe d'assurance "dynamique" sera examiné, ainsi que la volonté d'obtenir une quantification continue du risque porté par un assuré.

1.2.1 Une réglementation en évolution

En raison de l'impact croissant des systèmes d'information et des données sur notre société, un cadre réglementaire structurant ce domaine se renforce progressivement. Que ce soit dans la protection des données personnelles, la réglementation des infrastructures critiques ou les mesures de résilience en général, ce cadre est aujourd'hui en pleine émulation. Quelques lois et autres directives seront présentées ici afin d'obtenir une vision plus précise du cadre légal entourant le monde de la cybersécurité en France et dans le monde.

1.2.1.1 La protection des données

Dans le domaine de l'assurance, la donnée est le nerf de la guerre. La protection des données personnelles est donc devenue une priorité pour les régulateurs. En France et en Europe, plusieurs textes législatifs ont été adoptés pour protéger les informations personnelles des individus et garantir leur confidentialité.

L'une des lois les plus significatives dans ce domaine est le Règlement Général sur la Protection des Données (RGPD) ou (EU) 2016/679, entré en vigueur en mai 2018 (CNIL, 2016). Ce règlement impose des obligations strictes aux entreprises concernant la collecte, le traitement et le stockage des données personnelles. Les principales exigences du RGPD incluent le consentement explicite des individus pour la collecte de leurs données, ainsi que des droits d'**accès**, de **rectification** et d'**effacement**, entre autres. Il impose également aux entreprises de **notifier les autorités compétentes en cas de violation de données** (Article 33), incluant bien évidemment les cyberattaques. Les entreprises

se portent garantes des données qu'elles détiennent. Ainsi, en cas de vol ou de compromission des données, elles peuvent être légalement responsables en vertu du RGPD, s'exposant ainsi à des sanctions de la Commission Nationale de l'Informatique et des Libertés (CNIL), ainsi qu'à d'éventuels dommages et intérêts.

Plus récemment, et toujours au niveau européen, le *Data Governance Act* et le *Data Act* s'inscrivent dans la stratégie européenne pour les données de 2020 (CNIL, 2022). Le premier, un règlement sur la gouvernance des données adopté en 2022, vise à structurer le partage des données personnelles au sein de l'UE, notamment en mettant en place des certifications pour les fournisseurs de services d'intermédiation de données. Le second, adopté en 2023 et centré sur les données non personnelles, a pour objectif de favoriser une meilleure accessibilité aux données (notamment dans le cadre de l'*Internet of Things* (IoT)), de renforcer les droits des utilisateurs sur les données qu'ils génèrent, et de permettre à terme la création d'un marché unique européen des données (Ministère de l'économie, 2023).

Outre le RGPD et le *Data Act*, la France a également mis en place des lois nationales comme la Loi Informatique et Libertés, adaptée pour être conforme au RGPD (CNIL, 2019). Cette loi encadre l'utilisation des données personnelles et veille à ce que les droits des citoyens soient protégés.

Le cadre réglementaire poursuit un double objectif : d'une part, favoriser le partage des données à l'échelle européenne en normalisant les transactions, et d'autre part, protéger les citoyens contre les abus. Cet ensemble législatif s'inscrit dans le risque de coût légal évoqué dans la partie 1.1.1.4.

1.2.1.2 Cadre des infrastructures critiques

Les infrastructures critiques, telles que les réseaux de communication, les systèmes financiers et les services publics, sont essentielles au bon fonctionnement de la société. Leur protection contre les cybermenaces est donc cruciale.

En France, la Loi de Programmation Militaire (LPM) de 2013, complétée par la LPM 2019-2025 et celle de 2024-2030, prévoit des mesures spécifiques pour protéger les opérateurs d'importance vitale (OIV) (CyberInstitut, n.d.). Elle impose, entre autres, l'obligation pour les OIV de mettre en place des plans de sécurité, de réaliser régulièrement des audits de sécurité et de notifier les incidents de sécurité à l'Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI).

Au niveau européen, la directive NIS (*Network and Information Systems Directive*) vise également à renforcer la cybersécurité des infrastructures critiques. Adoptée en 2016 et transposée en 2018, la directive NIS 1 visait à améliorer la gouvernance de la cybersécurité dans les États membres, en instaurant des autorités nationales de cybersécurité (comme l'ANSSI en France), en encourageant la coopération entre États, et en imposant des exigences de sécurité aux acteurs de dix secteurs stratégiques (ANSSI, 2022a).

L'adoption de NIS 2 en 2022 (directive (UE) 2022/2555) marque une volonté d'étendre ce cadre de sécurité avec un "changement de paradigme". Parmi les propositions figurent la régulation des petites structures, qui, comme évoqué précédemment, sont particulièrement sensibles et mal préparées à ce risque. Le nombre d'entités concernées passe ainsi d'une centaine sous NIS 1 à plusieurs milliers à l'échelle nationale. Les critères d'intégration reposent principalement sur l'activité de l'entreprise et son chiffre d'affaires (ANSSI, 2023a).

1.2.1.3 L'encadrement de la résilience des entreprises

La volonté d'encadrer les infrastructures critiques existe depuis longtemps, mais l'extension du cadre législatif pour inclure la cyberrésilience des entreprises en général est relativement récente. Cette volonté est néanmoins en action au niveau européen, comme l'a démontré NIS2. D'autres projets de lois sont en cours à l'échelle européenne, à l'image du règlement DORA (*Digital Operational Resilience Act*), qui s'applique à un large panel d'entités financières, y compris les assurances. Son objectif est d'améliorer significativement la résilience cyber de ces structures. Entrant en application en 2025, ce règlement impose des exigences en matière de gestion des risques, de tests de résilience et de communication des incidents (Pwc, 2023).

En France, des initiatives telles que le Plan de Continuité d'Activité (PCA) et le Plan de Reprise d'Activité (PRA) sont encouragées afin d'assurer la continuité des services en cas de cyberattaque. Ces plans incluent des mesures visant à identifier les risques, mettre en place des procédures de réponse et garantir une reprise rapide des opérations.

1.2.1.4 Une évolution également dans le monde assurantiel

Ce cadre légal, en constante évolution en Europe et en France ces dernières années, a également des répercussions sur le secteur de l'assurance, où la gestion des risques cybernétiques devient de plus en plus structurée. Nous pouvons citer, par exemple, la création de deux nouvelles catégories relatives au risque cyber dans le code des assurances. Il s'agit du compte 32, relatif aux "Dommages aux biens consécutifs aux atteintes aux systèmes d'information et de communication", et du compte 33, relatif aux "Pertes pécuniaires consécutives aux atteintes aux systèmes d'information et de communication" (Article A344-2 du Code des assurances, suite à la publication de Legifrance, 2022).

La réglementation en matière de cybersécurité et de protection des données est en constante évolution pour répondre aux menaces croissantes. Les lois et directives mentionnées ci-dessus illustrent l'effort concerté des autorités pour protéger les données personnelles, sécuriser les infrastructures critiques et assurer la résilience des systèmes d'information. L'assurance évolue elle aussi, afin de prendre en compte le risque cyber de manière plus fiable et sereine. Tout cela s'inscrit dans le développement de solutions pour un risque de plus en plus préoccupant.

1.2.2 L'assurance cyber - un panorama de la situation française

Cette sous-section présentera des chiffres clés de l'assurance cyber. Cette introduction contribuera à mieux comprendre le contexte et les enjeux majeurs de l'assurance cyber.

1.2.2.1 Une variété de garanties

Comme mentionné dans la partie (1.1.1.4), les dommages subis par les entreprises lors d'un incident cyber sont multiples. Les garanties proposées par l'assurance cyber sont à l'image de cette multitude de pertes. Comme présenté par Hillairet and Lopez, 2022, les garanties d'un contrat d'assurance cyber incluent typiquement :

- Une garantie de remboursement des dommages et réparation des systèmes, pour combler les coûts de remédiation.
- Une garantie de perte d'exploitation, pour combler le manque à gagner.

- Une garantie responsabilité civile pour combler le coût légal (en particulier avec les pertes de données et l'application de la RGPD évoqué lors de la partie (1.2.1.1))
- une garantie sur les coûts de communication de crise pour combler les coûts de réputation

En comparant avec les dommages liés au risque cyber décrits dans la section (1.1.1.4), il apparaît que l'assurance cyber offre une couverture étendue pour la plupart des types de dommages auxquels les entreprises peuvent être confrontées.

1.2.2.2 Un volume de primes relativement faible et en pleine évolution

En examinant le panorama de l'assurance cyber en France, le marché apparaît encore en développement. Au niveau des primes, le volume total était de 183 M€ en 2021 et de 316 M€ en 2022 (AMRAE, 2023), représentant une augmentation de 72% en un an. L'augmentation entre 2022 et 2023 est, quant à elle, plus faible, avec un volume total de 328 M€ pour 2023, soit une croissance de seulement 4% (AMRAE, 2024). Cette faible augmentation est en grande partie due à la stagnation du volume de primes des grandes entreprises qui, comme nous le verrons, constituent le nœud central de la cyberassurance en France. Pour les autres tailles d'entreprises, la croissance reste forte : 27% entre 2022 et 2023 pour les ETI et 73% pour les entreprises de taille moyenne, selon l'étude LUCY 2024.

Néanmoins, ces chiffres doivent être remis en perspective. L'assurance cyber ne représente que 3% des primes (et 0,35% du chiffre d'affaires) des assurances de biens et responsabilité (Ministère de l'économie, 2022). Pour un risque classé en tête de liste par France Assureurs, 2023 depuis plusieurs années, ce pourcentage semble assez faible et témoigne d'une assurance encore imparfaitement implantée. Comme l'explique Philippe Cotelle dans LUCY AMRAE, 2024, "Il suffit de quelques sinistres d'importance pour venir rogner considérablement sur cette prime".

À l'international, ce développement est également marqué. Selon Munich Re, 2023, le marché cyber était évalué à 11,9 milliards de dollars en 2022 et devrait atteindre 33,3 milliards en 2027, soit une croissance de plus de 22% par an.

1.2.2.3 Une assurance encore concentrée sur les grands acteurs

Avec 280 grandes entreprises assurées en 2023, 94% des grandes entreprises françaises sont couvertes par une assurance cyber (AMRAE, 2024). Ce chiffre est resté stable entre 2022 et 2023, selon l'étude LUCY, indiquant une constance dans la couverture de ce type d'entreprises.

Ce même constat ne peut être fait pour les ETI et les entreprises de taille moyenne. Comme observé précédemment, la forte croissance des primes entre 2022 et 2023 témoigne d'un marché en plein développement. En termes d'entreprises assurées, le nombre a également fortement augmenté, avec une croissance de 47% pour les ETI et de 194% pour les entreprises de taille moyenne (AMRAE, 2024).

Deux points sont néanmoins à relever :

- **La proportion d'ETI et d'entreprises de taille moyenne assurées reste très faible**. En 2023, seulement 15% des ETI étaient couvertes par une assurance cyber, et cette proportion est encore plus faible pour les PME. Ce constat est problématique, car il a été souligné à plusieurs reprises que ces entreprises sont les plus vulnérables aux cybermenaces.
- **La concentration des primes reste largement dominée par les grandes entreprises**.

Toujours selon LUCY, bien qu'elles ne représentent que 2% des polices analysées par l'étude, elles pèsent à elles seules 80% des primes collectées.

Le marché français est donc principalement dominé par les grandes entreprises, tant en termes de primes que de pourcentage de souscription. Toutefois, la croissance rapide de la proportion d'ETI et de PME assurées révèle un intérêt croissant pour ce type d'assurance parmi les plus petits acteurs.

1.2.2.4 Un S/P encore très instable en France

Un S/P stable au niveau macro est une preuve d'une bonne connaissance du risque et d'une bonne santé de l'assurance en général. Le ratio sinistres/primes (S/P) de l'assurance cyber, quant à lui, est particulièrement volatile. Après une période de hausse ayant atteint le niveau préoccupant de 167 % en 2020, il semble suivre une tendance à la baisse depuis. La figure 1.18 illustre cette évolution ainsi que les montants des primes et des sinistres au niveau français (AMRAE, 2024).

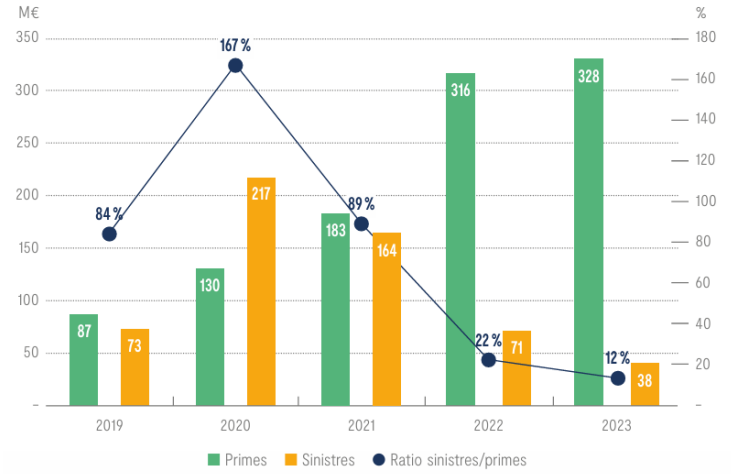


Figure 1.18: S/P français dans l'assurance cyber de 2019 à 2023 - AMRAE, 2024.

Cette baisse permet au S/P d'atteindre 12% en 2023. Néanmoins, comme le présente le baromètre Cesin Opinion Way Cesin, 2024, le nombre d'attaques réussies ne diminue pas, il se stabilise ou augmente. En effet, 66% des membres interrogés (sur les 456 répondants) considèrent que le nombre de cyberattaques est resté stable, tandis que 23% estiment qu'il a augmenté.

Plusieurs raisons expliquent cette baisse du S/P. La première est une politique de "*hard market*" menée par les assureurs à la suite du S/P inquiétant de 2020, caractérisée par une augmentation des primes, la mise en place de franchises et de limites plus strictes, ainsi qu'un tri plus rigoureux des assurés. Ce *hard market* semble s'atténuer progressivement pour favoriser le développement de l'assurance et la rendre plus attractive, les assureurs étant rassurés par la stabilisation du S/P. La deuxième raison pourrait être une meilleure hygiène numérique des entreprises. Toutefois, comme observé, le nombre de cyberattaques réussies ne semble pas diminuer. Toujours d'après LUCY et Cesin Cesin, 2024, trois quarts des entreprises ayant subi une cyberattaque ne font pas appel à leur assurance cyber, ce qui laisse "supposer que l'impact de ces incidents est inférieur aux montants des franchises" (AMRAE, 2024).

Enfin, il est important de noter que le poids des grandes entreprises influence fortement le S/P global. En 2023, toutes les classes d'entreprises (ETI, PME et grandes entreprises) affichent un S/P

très positif. Cependant, en 2022, les moyennes entreprises affichaient un S/P de 100%, tandis que le S/P global n'était que de 22%. Cela s'explique par le fait que les grandes entreprises n'avaient qu'un S/P de 16%. Une mauvaise année pour ces grandes structures pourrait donc remettre en cause cette vision globale.

Il est également pertinent de considérer que le bon S/P des PME/ETI en 2023 résulte peut-être d'un tri sélectif, laissant les entreprises moins résilientes sans assurance.

La section précédente met en lumière un marché en pleine évolution. Il est encore peu développé pour les ETI/PME, qui sont des acteurs sensibles au risque cyber. La quantité de primes est en augmentation, mais reste relativement faible, et le secteur de l'assurance cyber pourrait rencontrer des difficultés en cas d'année marquée par de grands sinistres. Il s'agit d'une assurance en construction, comme le dépeint AMRAE, 2024, une construction qui semble suivre une trajectoire prometteuse, mais avec une maturité des assurés très disparate. Comme relevé par Ministère de l'économie, 2022, "L'assurance du risque cyber constitue pourtant un levier essentiel du renforcement de la résilience de notre tissu productif", soulignant ainsi la nécessité de continuer à étudier le risque sous-jacent afin de mieux assurer et protéger les différents niveaux de la société.

1.2.3 Les défis du risque cyber en assurance

Dans cette section, les caractéristiques qui rendent le risque cyber particulièrement complexe à modéliser et à assurer seront examinées en détail. Au-delà de son instabilité réglementaire et de son évolution rapide, abordées dans les sections précédentes, d'autres facteurs contribuent à cette difficulté. Cette partie explorera ces différents aspects qui rendent la gestion du risque cyber si délicate.

1.2.3.1 Un manque de données

Le manque de données fiables représente un véritable défi dans le domaine de l'assurance cyber. Contrairement à d'autres risques pour lesquels les données publiques sont facilement accessibles, ou où les assureurs disposent de bases de données robustes et représentatives grâce à une longue expérience, le risque cyber se distingue par une carence significative dans ce domaine.

Les assureurs sont souvent insuffisamment matures pour disposer de bases de données solides, et il existe trop peu de données publiques disponibles. Cette situation s'explique en grande partie par la réticence des entreprises victimes d'attaques à partager des informations, par crainte de nuire à leur image ou de subir un désavantage concurrentiel. Ainsi, l'ensemble des acteurs du secteur de l'assurance réclament des données plus transparentes et accessibles afin de mieux évaluer et gérer les risques cyber.

Lorsque Hillairet and Lopez, 2022 évoquent l'idée de tendre vers une application des méthodes utilisées pour les maladies (réseau Sentinelles) pour le partage d'informations et une méthodologie plus claire pour la menace cyber, ou lorsque le Ministère de l'économie, 2022 incite à "faciliter la transmission d'informations entre assureurs" et à "harmoniser" les méthodologies, c'est pour permettre d'obtenir un "degré d'informations qui manque aujourd'hui au suivi de la menace cyber et à l'anticipation de ses évolutions" (Hillairet and Lopez, 2022).

Des bases de données de sinistres publiques existent, comme la base PRC (*Privacy Rights Clearinghouse*), qui fournit une liste des fuites de données aux États-Unis. En effet, dans ce pays, certaines failles sont rendues publiques par le gouvernement. Cependant, les conclusions qui peuvent être tirées de ces données ne reflètent pas nécessairement la situation en France. Ces données sont principalement axées sur les pertes de données et peuvent être biaisées (Hillairet and Lopez, 2022). Il est d'ailleurs

mentionné sur le site de PRC : *"[PRC] should not be considered a complete and accurate representation of every data breach in the United States"*.

D'autres bases existent, comme la base de Veris, qui propose un framework pour normaliser la déclaration de sinistres. Toutefois, cette base repose sur le volontariat, ce qui la rend extrêmement biaisée, car les entreprises sont généralement très réticentes à partager ce type d'informations, comme cela a été mentionné dans la partie 1.1.1.1.

Un autre exemple de base publique est **Hackmageddon**, créée et régulièrement mise à jour par le professeur *Paolo Passeri* à partir de diverses sources publiques. Cette base permet de suivre et d'analyser le risque cyber. Bien qu'elle soit loin d'être parfaite (car elle repose uniquement sur des données publiques et est tenue à jour par un seul individu), elle offre néanmoins la possibilité d'étudier certains aspects du risque cyber. Elle est, par ailleurs, utilisée pour l'application d'un modèle de Hawkes par Boumezoued et al., 2023.

1.2.3.2 Les défis de la fréquence

Cette section se penche sur les défis liés à la fréquence des incidents cyber. En s'appuyant largement sur l'analyse de Awiszus et al., 2022, les trois composantes principales qui rendent le risque cyber particulièrement complexe à modéliser du point de vue de la fréquence seront examinées.

Le Risque Idiosyncratique Il s'agit de la partie de risque propre à chaque assuré. Par exemple, le risque de commettre une erreur en interne est a priori indépendant de celui du voisin. Cette composante est relativement simple à modéliser, car elle suit les hypothèses classiques établies précédemment.

Le Risque Systématique Cette composante de risque découle des vulnérabilités communes à différentes entités (plusieurs entreprises utilisant un même outil souffrant d'une faille). En raison de cette composante, l'hypothèse d'indépendance entre les incidents (ou entre les individus, selon la modélisation utilisée) devient invalide. Plus précisément, il est constaté qu'une faille exploitable et exploitée facilite le processus d'attaque pour les cybercriminels, rendant celle-ci plus attrayante. L'auto-corrélation du nombre d'attaques a été vérifiée sur les bases publiques évoquées précédemment (voir Hillairet and Lopez, 2022, p.22, pour un exemple sur la base PRC). L'excitation est ici **externe**, car elle ne dépend pas du nombre d'assurés déjà infectés, mais de facteurs externes à l'infection (failles existantes chez l'assuré, etc.).

Le Risque Systémique Ce risque fait référence à la contagion pouvant se produire entre différents acteurs. Qu'il s'agisse du risque lié à la *Supply Chain*, abordé en section 1.1.2.1, ou de la propagation de virus à travers les réseaux interconnectés entre entreprises, ce type de risque est particulièrement préoccupant pour les assureurs. En raison de l'interconnexion croissante entre les entités, les perturbations peuvent se propager rapidement d'une organisation à l'autre, entraînant des conséquences financières et opérationnelles majeures pour l'ensemble du portefeuille d'assurés. Ce **phénomène d'accumulation** remet en question non seulement l'indépendance des intertemps, mais aussi le caractère mutualisable du risque dans son ensemble.

Il est à noter que, dans d'autres littératures, la distinction entre les risques systématique et systémique n'est pas toujours faite, les deux typologies étant souvent regroupées sous l'appellation "systémique". Il est néanmoins intéressant d'analyser le risque sous ce prisme. La distinction entre ces deux concepts peut être difficile à établir. Par exemple, l'attaque *NotPetya* exploitait une faille Windows (risque systématique), mais s'est également propagée via la chaîne d'approvisionnement (*Supply Chain*), ce qui relève du risque systémique.

Cependant, certains outils existent pour prendre en compte ces spécificités. Il est possible de mentionner les processus de Hawkes à deux facteurs, permettant d'étudier l'auto-excitation et l'excitation externe. Un article consacré à ces processus dans le cadre du risque cyber fournit des résultats prometteurs (Boumezoued et al., 2023). Un mémoire sur ce sujet a également été réalisé par Bessy-Roland, 2019.

Concernant les risques d'infection et de propagation en particulier (composante systémique), la recherche s'appuie également sur les modèles épidémiologiques issus du domaine médical. Un excellent mémoire traite de cette problématique dans le cadre du **cyber silencieux** (Peyrat, 2022). Le cyber silencieux correspond aux polices d'assurance ne prenant pas explicitement en compte le risque cyber dans la tarification, mais étant activées par celui-ci (par exemple, une police de perte d'exploitation mal formulée n'excluant pas le risque cyber). D'après Opinions & Débats, il s'agit également d'un des grands enjeux du monde de la cyberassurance (Hillairet and Lopez, 2022).

Ces théories restent néanmoins très difficiles à appliquer en pratique en raison du manque crucial de données et de la complexité de la calibration de ces modèles sophistiqués.

1.2.3.3 Les défis du coût

Le coût des sinistres présente également des défis importants. Comme le souligne Hillairet and Lopez, 2022, le caractère mutualisable de ce risque nécessite l'existence d'une espérance μ. Cependant, la forte disparité des coûts de sinistres observée dans certains cas et la présence de risques extrêmes conduisent souvent à l'utilisation de lois à queue lourde dans les modélisations actuarielles classiques. Cela soulève des questions sur l'assurabilité de certains acteurs (voir Hillairet and Lopez, 2022 pour une présentation détaillée de cette problématique). La variabilité des coûts de sinistres est illustrée dans la figure 1.19.

La présence de sinistres XXL, de faible occurrence mais de fort impact, confirme cette forte variabilité. Bien qu'une baisse encourageante de ce type de sinistres soit visible sur le graphique, la réalité du risque cyber n'est pas amoindrie, et le marché français n'est pas à l'abri d'un sinistre de très grande ampleur touchant l'un de ses plus grands acteurs (Hillairet and Lopez, 2022).

Il est, en outre, difficile de tirer des conclusions sur une amélioration durable en se basant uniquement sur les observations de quelques années.

La modélisation actuarielle du risque cyber est encore en pleine évolution. Comme le montre le panorama de l'assurance cyber présenté tout au long de la partie 1.2, le secteur est en transformation et les primes ne sont pas encore entièrement stabilisées. Les défis de modélisation, associés à un manque crucial de données, compliquent la construction de modèles basés sur les sinistres passés. L'analyse des entreprises présentée dans la section 1.1.4.1 révèle la nécessité impérative d'accompagner les PME et les ETI dans le renforcement de leur résilience cyber et l'acquisition d'une assurance appropriée. De plus, notre exploration du domaine de la cybersécurité au cours de la première partie de ce chapitre a mis en évidence l'existence d'outils permettant une meilleure quantification du risque cyber.

Dans cette dernière partie, la nécessité d'une quantification continue du risque individuel sera examinée, et la problématique qui servira de base à la modélisation sera définie.

1.2.4 Un besoin de quantification en continu

Tout au long de cette première partie, la quantification du risque cyber est apparue comme un enjeu majeur. Que ce soit pour les professionnels de la cybersécurité, cherchant à mieux appréhender l'évolution du risque et les méthodes utilisées par les attaquants, ou à un niveau plus global pour les

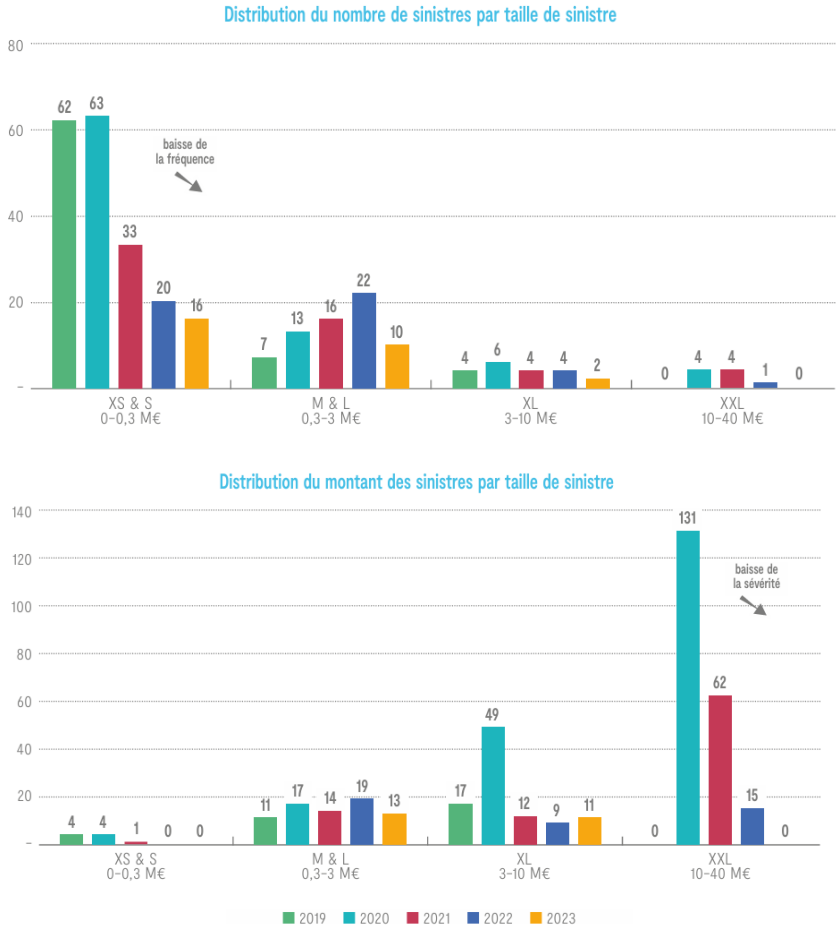


Figure 1.19: Nombre et montant par taille de sinistre sur les années 2019 à 2024 - AMRAE, 2024.

assureurs afin de concevoir des contrats attractifs et robustes. Les modèles classiques peinent à être pleinement utilisables en cyberassurance.

Néanmoins, les connaissances en cybersécurité établies et les données disponibles dans ce domaine permettent d'envisager une **estimation "sans sinistre" de la prime pure**, dépendante du profil de risque de l'assuré et de l'environnement extérieur. Cette estimation pourrait être réalisée automatiquement et en continu, permettant ainsi d'observer avec précision l'évolution du portefeuille de l'assureur.

L'appel à une quantification dynamique et continue se fait de plus en plus pressant au sein des différents acteurs. Opinion & Débats Hillairet and Lopez, 2022 mentionne que "la prévention, qui est un élément important de la réduction de la menace, bénéficierait fortement d'un baromètre en temps réel, d'autant que l'échelle temporelle du cyber-risque est assez courte".

Ce besoin de suivi est non seulement crucial pour développer une méthode de quantification plus adaptative, mais également, comme mentionné précédemment, pour permettre aux assureurs de mener des actions de prévention efficaces. Ce point est particulièrement important pour les petites et moyennes structures qui ne disposent pas nécessairement d'un environnement de cybersécurité très développé. L'investissement de l'assureur en tant qu'intermédiaire dans la communication sur le risque en temps réel peut permettre de réduire considérablement l'impact des menaces.

De plus, il permettrait de diminuer le risque humain et technique intrinsèque à l'entreprise en

proposant des conseils et des audits, tels que des formations aux risques liés au *phishing* ou la vérification de la fermeture de ports sur le site web. Ces différents aspects peuvent significativement réduire le risque d'une structure, apporter une forte valeur ajoutée à l'assurance (et donc en améliorer l'attractivité) sans pour autant accroître de manière excessive le coût pour l'assureur. Comme précisé dans Wang, 2019, cela pourrait permettre de "cueillir les fruits bas", ceux qui ne demandent pas beaucoup d'efforts, mais produisent des résultats significatifs.

Certaines entreprises en France et à l'étranger fournissent déjà ce type de services. Coalition, une entreprise offrant un service d'assurance "active", propose un produit qui intègre un score de risque évolutif dans le temps. Ce produit fournit également un service d'évaluation des risques pour l'assuré, comme illustré dans la figure 1.20.

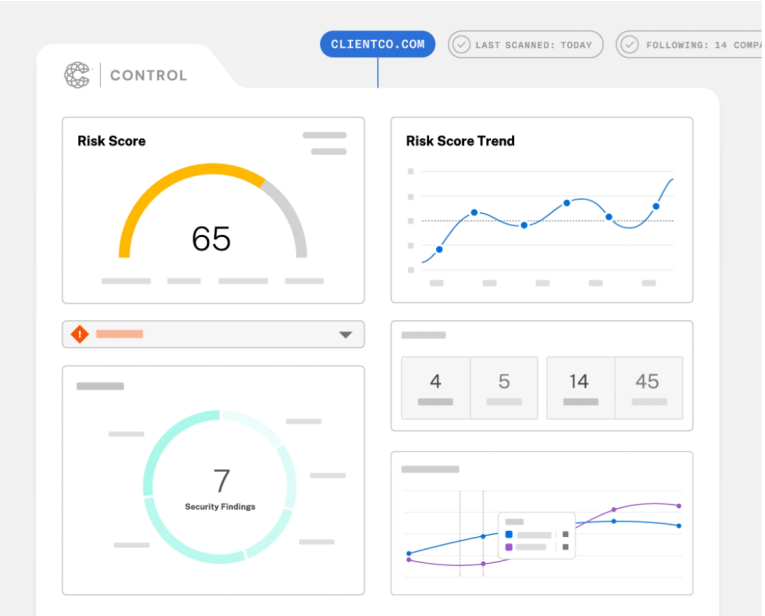


Figure 1.20: Image commerciale des services de Coalition - Une connaissance du risque en continu.

En France, certaines entreprises d'*insurtech*, telles que Dattak ou Stoik, proposent également ce type de services.

Proposer des méthodes permettant d'exploiter les connaissances sur les typologies d'entreprises, les méthodes d'attaque et les pratiques de cybersécurité pour évaluer le risque peut contribuer à faire évoluer l'assurance cyber. Cette approche est d'autant plus cruciale dans un contexte de manque de données.

L'objectif, dans la suite de ce mémoire, est alors de rechercher et de construire les bases d'une modélisation pour la quantification du risque utilisant les données de cybersécurité (CVE, CVSS, etc.) (présentées en partie 1.1.3.2), les méthodologies d'attaques (ATT&CK, etc.) (discutées en partie 1.1.2.3), ainsi que des méthodes alternatives à celles utilisées classiquement par l'actuariat.

Nous visons ainsi à démontrer **l'existence de pistes dans la modélisation actuarielle du risque cyber** capables de répondre aux défis actuels :

- à la problématique du manque de données de sinistres fiables (sous-section 1.2.3.1), par un modèle dépendant le moins possible des sinistres (la quantification "sans sinistre" évoquée dans la partie 1.2.4);
- aux limites de la modélisation classique se heurtant à l'évolutivité de ce risque (sous-section

1.2.3.2), par un modèle dynamique, dont la nécessité pour une évolution pérenne de l'assurance cyber a été traitée dans la section 1.2.4;

- mais aussi à un marché français encore trop peu centré sur les PME/ETI (sous-section 1.2.2.3).

Cela se fera avec une connaissance plus précise de l'assuré et de ses systèmes cyber, et sera bénéfique :

- pour l'assuré, qui, notamment pour les PME/ETI, bénéficiera d'une expertise dont il a besoin et à laquelle il a peu accès;
- pour l'assureur, qui disposera d'un moyen plus précis d'évaluer le risque individuel de son portefeuille.

Chapter 2

Modélisation dynamique du risque cyber — théories et recherches

Le chapitre 1 introduit deux aspects critiques de l'étude du risque cyber. Dans l'observation de la situation du monde cyber-assurantiel, le lecteur découvre un ensemble d'enjeux et de besoins (quantification continue, sans sinistre, plus précise, etc.) liés à la difficulté d'adaptation des connaissances et des méthodologies classiques au monde du cyber (voir partie 1.2). Dans l'étude des connaissances en cybersécurité, il découvre des méthodes d'observation, de gestion et d'analyse du risque qui ne sont néanmoins pas encore totalement appliquées à l'assurance (voir partie 1.1).

Ce chapitre 2 a pour objectif de construire les bases d'une modélisation pour la quantification du risque cyber, faisant **le pont entre les mondes** de l'assurance et de la cybersécurité, et ainsi de montrer qu'il existe des méthodes viables dans ce secteur. Les recherches se sont orientées vers une modélisation répondant aux enjeux assurantiels. L'objectif est, d'une part, de pallier le manque de données grâce à une quantification *"sans sinistre"*, et d'autre part, de proposer une approche *dynamique* permettant de suivre l'évolution du risque. Ce modèle doit également être applicable aux petites structures, souvent les plus vulnérables.

Pour ce faire, des modèles graphiques ont été étudiés. Ce type de modèle peut être adapté au contexte de la problématique établie, du fait de sa versatilité et de son adaptabilité, tant dans la prise en compte des connaissances en cybersécurité que des besoins en assurance. Une première partie abordera les bases théoriques : d'abord par une introduction à la théorie des graphes, qui constitue la base des modèles graphiques, puis par une introduction aux réseaux bayésiens, un modèle graphique probabiliste qui sera une des clefs de la modélisation construite. Une seconde partie traitera des applications des modèles graphiques à la quantification du risque cyber. Un article présentant un *framework* reliant cybersécurité et quantification assurantielle à l'échelle **micro** sera détaillé et servira de base pour notre modélisation. Une troisième partie présentera les bases d'un modèle **applicable à l'assurance perte d'exploitation cyber**. La modélisation reprendra les concepts présentés en seconde partie, tout en les adaptant aux besoins de l'assurance. Cette partie sera l'aboutissement des recherches effectuées.

2.1 Graphes et Modèle Graphique Bayésien — Cadre Théorique

Les modèles graphiques sont encore peu communs en actuariat classique, même dans un cadre cyber-assurantiel. Ceux présentés dans la suite de ce mémoire se basent sur des graphes et une approche stochastique particulière : les réseaux bayésiens.

Dans un objectif de clarification, cette première partie présentera les fondements théoriques qui seront utilisés dans la suite de ce chapitre et permettra ainsi de poser un socle solide pour l'application de modèles graphiques au monde du cyber. Cette partie se décompose ainsi : dans un premier temps, la théorie des graphes sera introduite au lecteur. Dans un second temps, les réseaux bayésiens seront étudiés dans l'objectif de présenter l'approche probabiliste utilisée par les modélisations explorées dans la suite de ce mémoire.

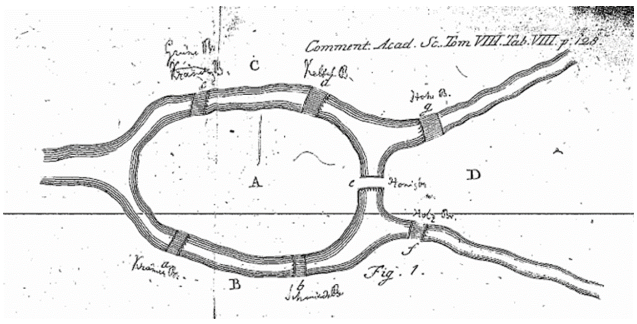
2.1.1 Théorie des Graphes

Avant d'étudier la théorie derrière les réseaux bayésiens, cette section posera quelques bases en théorie des graphes. D'une part, car les réseaux bayésiens sont des modèles graphiques et ont donc un lien théorique avec cette dernière, mais aussi parce que les graphes en tant que tels jouent un rôle primordial dans notre modélisation, notamment pour les graphes d'impacts dans la section 2.2.2.2. Cette section complétera les notions introduites dans le mémoire de Peyrat, 2022.

2.1.1.1 Un peu d'histoire

Aujourd'hui un champ central des mathématiques, la théorie des graphes est étudiée en tant que telle depuis près de trois siècles. En effet, ses débuts sont attribués aux travaux du mathématicien Leonhard Euler en 1736 sur le problème des sept ponts de Königsberg (Euler, 1736).

(a) Plan de la ville de Königsberg à l'époque d'Euler



(b) Modélisation faite par Euler. Avec A, B, C, D les quartiers (sommets) et a, b, c, d, e, f les ponts (arêtes)

Figure 2.1: Modélisation du problème des sept ponts de Königsberg par Euler, 1736

Le problème posé à Euler était le suivant : dans la ville de Königsberg (alors en Prusse, aujourd'hui Kaliningrad en Russie), il existait à l'époque sept ponts reliant chaque quartier de la ville (comme nous le voyons sur la sous-figure 2.1a). Était-il alors possible de traverser l'ensemble des ponts en ne les empruntant qu'une seule fois ?

D'abord réticent à l'idée d'écrire sur ce problème, qu'il considérait davantage lié à la logique qu'aux mathématiques (*Correspondances d'Euler avec Ehler, avril 1736*, Sachs et al., 1988), Euler finit par se laisser intriguer par une question qui, sous des apparences simples, dissimulait une difficulté inattendue.

Il écrira dans une lettre au mathématicien Giovanni Jacobo Marinoni que ce problème méritait une attention particulière, car ni la géométrie, ni l'algèbre, ni même le dénombrement ne semblaient des outils suffisants pour le résoudre. Euler démontrera alors, dans un papier de 1736, que cela n'était pas possible dans le cas de la ville de Königsberg. Par la suite, un chemin passant par toute arête

exactement une fois fut nommé chemin eulérien. Si, de plus, le chemin se termine là où il a commencé, c'est alors un cycle eulérien et le graphe associé est appelé graphe eulérien.

Euler donna par ailleurs des conditions nécessaires et suffisantes pour qu'un graphe connexe (non orienté et où tout point est connecté à tout autre point) soit eulérien, sans pour autant les prouver. Elles le seront 130 ans plus tard par le mathématicien Carl Hierholzer. Euler avait formulé qu'un graphe n'était eulérien que si chaque sommet avait un nombre pair d'arêtes (connu aujourd'hui comme le théorème d'Euler). Hierholzer ira plus loin en présentant et en prouvant une deuxième caractérisation : un graphe connexe admet un chemin eulérien si, et seulement si, tous ses sommets ont un degré pair, à l'exception d'au plus deux sommets qui peuvent avoir un degré impair (*Ponts de Königsberg et cycle eulérien* n.d.).

En utilisant ce théorème d'Euler (aussi appelé théorème d'Euler-Hierholzer), nous pouvons facilement résoudre le problème des sept ponts. Si nous prenons le sommet A, celui-ci disposant de 5 arêtes, il nie alors la première condition d'Euler. Les points B et D ayant, eux aussi, un nombre d'arêtes impair, cela contredit également la seconde condition. Ainsi, ni cycle eulérien, ni même chemin eulérien ne sont possibles.

C'est de cette histoire que naît la première propriété des graphes clairement explicitée comme telle. Aujourd'hui, le domaine a grandement évolué et est fortement utilisé dans la modélisation d'interactions et de réseaux (Internet, épidémiologie, etc.).

Dans la suite de cette partie, une base dans ce domaine sera établie en s'appuyant sur le cours de Michel Rigo de l'Université de Liège (Rigo, 2009).

2.1.1.2 Graphes orientés

De manière formelle, un graphe **orienté** se définit comme une paire de la forme $G = (V, E)$ où V est un ensemble (fini ou infini) et E est une partie de $V \times V$ appelée une *relation* sur V. Les éléments de V sont appelés les **sommets** du graphe, et les éléments de E, les **arêtes**.
Si les n éléments de V sont notés $(v_i)_{1 \leq i \leq n}$, alors les éléments de E seront de la forme $(v_j, v_k)_{1 \leq j,k \leq n}$, avec j le sommet de départ de l'arête et k le sommet d'arrivée.
Il est alors remarquable que, pour un graphe orienté fini (pour lequel $\text{card}(V) < +\infty$), le nombre maximal d'arêtes (sans duplication d'arêtes) est de $(\#V)^2$, où $\#$ représente le cardinal.

Comme mentionné précédemment, un graphe a pour usage de représenter des relations entre différents éléments ; il est donc souvent représenté sous une forme graphique (représentation sagittale), comme illustré dans la figure 2.2.

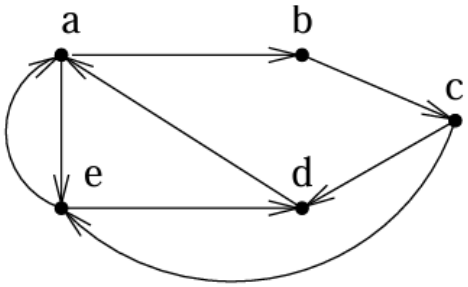


Figure 2.2: Représentation sagittale d'un graphe orienté - (Rigo, 2009)

Degrés Concernant une arête $a_r = (v_i, v_j)$, on dira qu'elle est *sortante* de v_i et *entrante* à v_j. Pour un sommet v, l'ensemble de ses arêtes sortantes sera noté $w^+(v)$ et l'ensemble de ses arêtes entrantes $w^-(v)$. Enfin, l'ensemble de ses arêtes *incidentes* sera noté $w(v) = w^-(v) \cup w^+(v)$.
De plus, le *demi-degré* sortant (resp. entrant) d'un sommet v est défini par

$$d^+(v) = \#(w^+(v)) \text{ (resp. } d^-(v) = \#(w^-(v)).$$

Et, de même, le *degré* de v est défini comme

$$d(v) = d^+(v) + d^-(v).$$

On peut alors démontrer (lemme des poignées de main) que, pour un graphe fini,

$$\sum_v d^+(v) = \sum_v d^-(v).$$

Cela s'obtient simplement en notant que

$$\sum_v d^+(v) = \sum_{v \in V} \sum_{1 \leq i \leq n} \mathbb{1}_{v_i = v} = \sum_{1 \leq i \leq n} \sum_{v \in V} \mathbb{1}_{v_i = v} = \sum_{1 \leq i \leq n} 1 = \#E.$$

En posant $a_r = (v_i, v_j)$, la même procédure peut être appliquée pour $d^-(v)$. Ainsi, le lemme des poignées de main stipule finalement que

$$\sum_v d(v) = \sum_v d^+(v) + \sum_v d^-(v) = 2 \times \#(E).$$

Voisins Enfin, il peut être utile de définir la notion de *prédécesseurs* et de *successeurs*. On notera $succ(v) = \{s_1, ..., s_k\}$ (resp. $pred(v) = \{p_1, ..., p_l\}$) l'ensemble des sommets tels que $(v, s_i) \in w^+(v)$ (resp. $(p_i, v) \in w^-(v)$).
Enfin, la liste des voisins de v s'exprime simplement comme $\nu(v) = pred(v) \cup succ(v)$.

Graphe simple ou multigraphe Un *graphe simple* (qu'il soit orienté ou non) est un graphe pour lequel il ne peut exister au plus qu'une seule arête entre deux sommets et aucune boucle (d'arête de la forme (a, a) par exemple). Un *multigraphe* autorise, lui, ces deux catégories de liens.

2.1.1.3 Graphes non orientés

Il est aussi possible de définir la notion de graphe non orienté. Fondamentalement, soit $G = (V, E)$ un graphe, il sera dit non orienté lorsque E est une *relation symétrique* sur V. En d'autres termes, cela signifie que

$$\forall a, b \in V \times V, (a, b) \in E \implies (b, a) \in E.$$

Il est souvent convenu de représenter la double arête (a, b), (b, a) par un unique élément $\{a, b\}$ (ensemble ne prenant pas en compte le sommet d'arrivée ou de départ, juste la liaison), E devenant alors un ensemble d'ensembles. Dans la représentation sagittale, cela est illustré par un trait unique sans flèche, comme présenté dans la figure 2.3.

Il est possible, à la manière des graphes orientés, de définir l'ensemble des arêtes incidentes $w(v)$, ainsi que le degré de v, défini par $d(v) = \#w(v)$. En notant que, par la définition de E en ensemble d'ensembles, le cardinal de E se trouve mathématiquement divisé par deux. Le lemme de la poignée de main s'applique toujours sous la forme

$$\sum_v d(v) = 2 \times \#(E).$$

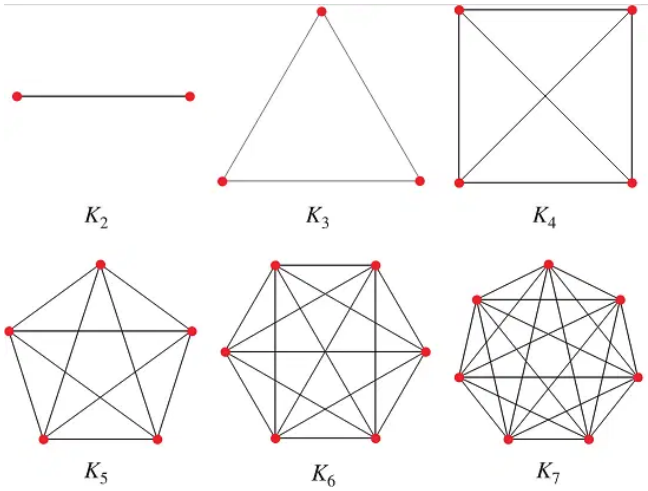


Figure 2.3: Représentation sagittale de graphes non orientés

Régularité et complétude Un graphe simple est dit *k-régulier* lorsque $\forall v \in V, d(v) = k$. De plus, si $k = \#V - 1$, on dit que le graphe est *complet*. En d'autres termes, un graphe complet est un graphe dans lequel tous les sommets sont adjacents deux à deux.
Un graphe complet à n sommets est, par convention, nommé K_n. La figure 2.3 présente les sept premiers graphes complets. Il est par ailleurs simple de démontrer que, pour un graphe complet, le nombre d'arêtes est donné par

$$\sum_{i=1}^{n}(n-i) = \frac{n(n-1)}{2}.$$

Arbres Un graphe simple non orienté est nommé *arbre* s'il est connexe et sans cycles. Par exemple, l'ensemble $(\{a,b\}, \{b,c\}, \{c,a\})$ forme un cycle de longueur 3. Un graphe dont chaque composante connexe est un arbre est appelé une *forêt*.

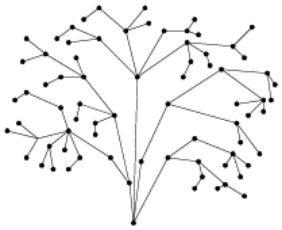


Figure 2.4: Représentation sagittale d'un arbre - (Rigo, 2009)

2.1.1.4 Graphes pondérés et représentation matricielle

Il est aussi possible d'étendre la notion de graphe (orienté ou non) en attribuant des *poids* aux différentes arêtes. Ce concept, nommé *graphe pondéré*, est particulièrement utile pour modéliser et optimiser des systèmes où les relations entre nœuds ont des caractéristiques quantitatives, comme les réseaux de transport, de communication ou de logistique.
D'un point de vue formel, on dit que $G = (E, V, f)$ est *étiqueté* par f lorsque

$$f : E \to \Sigma.$$

Si $\Sigma \subset \mathbb{R}$, alors G est *pondéré*.

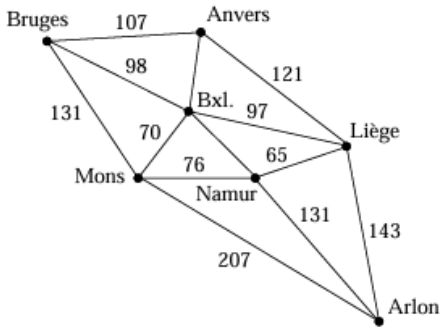


Figure 2.5: Représentation sagittale d'un graphe pondéré - (Rigo, 2009)

Représentation matricielle Depuis le début de cette section, la représentation sagittale d'un graphe a été utilisée. Bien que cette approche permette de facilement modéliser et visualiser un graphe, elle peut s'avérer limitée pour l'analyse, en particulier dans un contexte informatique. Une autre possibilité est de représenter le graphe par sa **matrice d'adjacence**.
La matrice d'adjacence A est une matrice $n \times n$ (avec n le nombre de sommets). En prenant $G = (V, E)$ et $V = (v_i)_{i \leq n}$, la matrice $A = (a_{i,j})_{1 \leq i,j \leq n}$ est définie comme

$$a_{i,j} = \begin{cases} 1 \text{ si } (v_i, v_j) \in E, \\ 0 \text{ sinon.} \end{cases}$$

Pour un graphe pondéré, le 1 peut être remplacé par le poids de l'arête. La matrice d'un graphe non orienté sera donc symétrique. La représentation matricielle d'adjacence de K_3 (figure 2.3) est donc

$$A_{K_3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

De même, pour le premier exemple (figure 2.2), dans le cas d'un graphe orienté, la matrice serait

$$A_{or} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

On peut remarquer la symétrie de la première matrice et la non-symétrie de la deuxième. Il est à noter que cette représentation est **unique**, et représente donc un et un seul graphe.

Au-delà de permettre une représentation des graphes adaptée au monde informatique (et utilisée lors du codage du modèle de ce mémoire), elle permet aussi, dans un cadre plus théorique, de mener une analyse spectrale des graphes. Celle-ci consiste à étudier les propriétés des graphes en examinant les valeurs propres et les vecteurs propres de leurs matrices d'adjacence (ou de Laplace). Cette approche permet de révéler des informations importantes sur la structure du graphe, telles que la *présence de communautés*, la *robustesse* et la *dynamique des réseaux*. Elle est particulièrement puissante pour résoudre des problèmes complexes en théorie des graphes et en informatique théorique.

Dans cette section, les concepts fondamentaux de la théorie des graphes ont été définis. La notion de parenté, qui sera utile tout au long de la modélisation FDNA, a notamment été introduite. Ensuite, les graphes orientés, base mathématique des graphes d'attaque, ainsi que les pondérations, essentielles pour la section suivante et pour le reste du mémoire, ont été abordés. Enfin, la représentation matricielle des graphes, qui sera utilisée dans le chapitre 3, a également été présentée.

Dans la section suivante, les réseaux bayésiens, qui font le lien entre la théorie des graphes et les probabilités, seront introduits.

2.1.2 Introduction générale aux réseaux bayésiens

Comme précisé dans *Secure IT Systems* (Chockalingam et al., 2017), le manque de données historiques, en particulier dans les intrusions de sécurité, constitue un frein au développement de modèles réalistes en cybersécurité. Cette problématique est également rencontrée dans le domaine de la cyberassurance.
Néanmoins, les Réseaux Bayésiens (RBs) ont pour atout de répondre aux différents enjeux posés par cette conjoncture. En particulier, leur capacité d'adaptabilité et leur aptitude à combiner différentes sources de connaissances peuvent permettre de rendre le contexte de **manque de données** moins insurmontable.

Cette partie s'intéressera à la présentation théorique des réseaux bayésiens, cadre stochastique sur lequel se baseront les différents travaux étudiés dans la suite de ce mémoire.

2.1.2.1 Introduction

Comme expliqué dans *Causality* (Pearl, 2013), les réseaux bayésiens remplissent principalement trois rôles :

1. fournir un moyen efficace d'exprimer des hypothèses, en particulier celles relatives à la dépendance entre différentes variables ;

2. donner une représentation économique des fonctions de probabilité jointe ;

3. faciliter l'*inférence* à partir d'observations.

L'objectif de ce type de réseau est fondamentalement de structurer la dépendance entre différentes variables aléatoires afin de mieux inférer et comprendre l'évolution des probabilités. Ils reposent sur un formalisme à la croisée des chemins entre la théorie des graphes et celle des probabilités (Naïm et al., 2011).

2.1.2.2 Formule de Bayes

Le domaine est appelé réseau **bayésien**, car la théorie de Bayes joue un rôle fondamental dans la construction dudit réseau. Un rappel sur cette théorie probabiliste sera effectué en se référant au travail de Jeamaon and Khemapatapan, 2020. L'ensemble des concepts mathématiques liés aux probabilités ne sera cependant pas défini ici. Le lecteur intéressé par un développement plus poussé pourra se référer à Dauxois, 2014.

Cas des événements Soit un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et un ensemble d'événements disjoints A_i formant une partition de l'univers Ω, ainsi que deux événements E et F quelconques, avec $\mathbb{P}(F) \neq 0$ et $\mathbb{P}(E) \neq 0$ (un exemple est illustré sur la figure 2.6). Dans un premier temps, la formule de Bayes nous indique que

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(F|E) \times \mathbb{P}(E)}{\mathbb{P}(F)}.$$

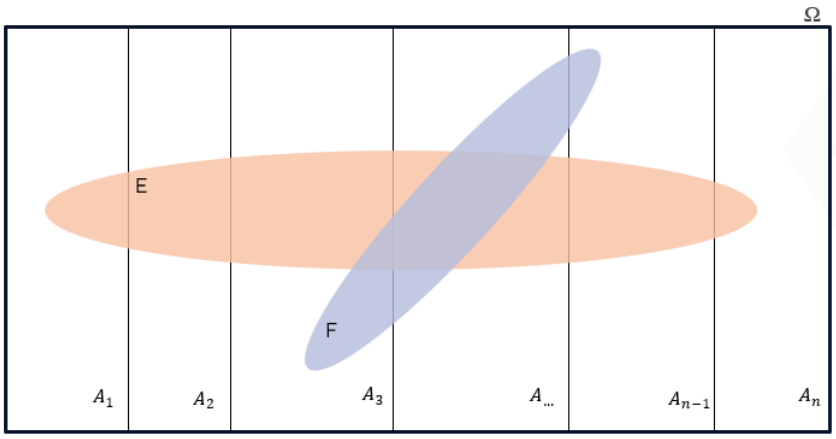


Figure 2.6: Représentation schématique de la division des événements

De manière purement informelle, ce théorème permet d'inverser la structure de la connaissance. Il est possible d'obtenir la connaissance de E sachant F en connaissant F sachant E (sous condition de connaître E et F). Si la connaissance de E et de F dépend uniquement des A_i, il est aussi possible de réécrire le théorème sous la forme :

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(F|E) \times \sum_i \mathbb{P}(E|A_i)\mathbb{P}(A_i)}{\sum_i \mathbb{P}(F|A_i)\mathbb{P}(A_i)}.$$

Pour des variables aléatoires discrètes, E et F représentent des événements distincts, et les probabilités sont calculées en sommant les valeurs individuelles.

Cas des variables continues La formule de Bayes s'applique également aux variables continues. Soient X et Y deux variables aléatoires continues, alors la densité conditionnelle de X sachant $Y = y$ s'exprime comme

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y) \times f_X(x)}{f_Y(y)}.$$

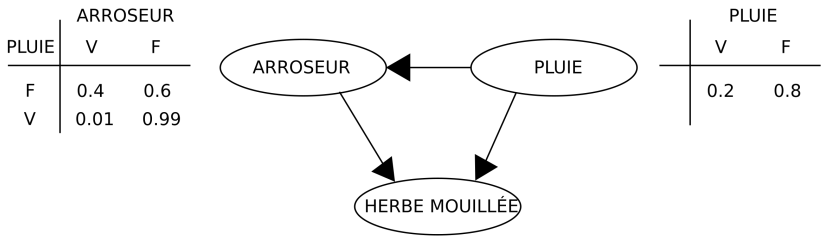
où f_X (resp. f_Y) désigne la fonction de densité de X (resp. Y), appelée densité *à priori*, et $f_{X|Y=y}/f_{Y|X=x}$ sont les densités conditionnelles respectives, appelées densités *à posteriori*.

2.1.2.3 Réseau bayésien

Le réseau bayésien a pour objectif de créer un réseau de dépendance sur lequel il est possible de construire une structure de probabilité. Sa force réside non seulement dans son explicabilité, mais aussi dans sa capacité d'applicabilité, aussi bien dans des domaines avec peu (voire sans) données d'apprentissage que dans des domaines où le volume de données est plus conséquent.

Formellement, comme présenté dans Naïm et al., 2011, un réseau bayésien se définit par :

1. un graphe acyclique $G = (V, E)$, correspondant à la structure du réseau ;
2. un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$;
3. un ensemble de variables aléatoires V_i correspondant à l'ensemble des nœuds du graphe G ($(V_i)_{i \leq n} = V$).



PLUIE	ARROSEUR V	ARROSEUR F
F	0.4	0.6
V	0.01	0.99

PLUIE V	PLUIE F
0.2	0.8

ARROSEUR	PLUIE	HERBE MOUILLÉE V	HERBE MOUILLÉE F
F	F	0.0	1.0
F	V	0.8	0.2
V	F	0.9	0.1
V	V	0.99	0.01

Figure 2.7: Premier exemple d'un réseau bayésien - Source : Wikipédia

Un premier exemple Pour mieux comprendre l'intérêt du réseau bayésien, un premier exemple est présenté avec la figure 2.7. Trois variables aléatoires sont visibles ($\{arroseur, pluie, herbe\_mouillee\}$), chacune avec un résultat binaire (soit il pleut, soit il ne pleut pas).
Associées à celles-ci, les probabilités conditionnelles des événements sont définies en fonction des parents ($w^-(a)$ du sommet a).

Ainsi, la probabilité que l'arroseur soit allumé dépend du fait qu'il pleuve ou non, et la probabilité que l'herbe soit mouillée dépend à la fois de la pluie et du fait que l'arroseur soit allumé. Cette structure de dépendance est facilement visualisable grâce au graphe, permettant ainsi une meilleure compréhension pour un humain (l'interprétabilité étant un atout majeur de ce type de modélisation).

Parents L'ensemble des parents de V_i dans le graphe G est noté $C(V_i)$. En reprenant les notations introduites plus tôt dans ce mémoire, cela correspond à $w^-(V_i)$. Comme expliqué dans Pearl, 1985, l'ensemble des parents de V_i regroupe les variables jugées *directement* reliées.

Dans l'exemple de la figure 2.7, les parents directs de "herbe mouillée" sont "pluie" et "arroseur". L'état de l'herbe dépend donc de ces deux variables. Il est également observable que "arroseur" dépend lui-même de "pluie", créant une structure déjà plus complexe, où le fait qu'il pleuve influence à la fois "arroseur" et "herbe mouillée". Ce type de relation est un exemple du défi de l'**inférence**, qui sera abordé plus tard.

Loi jointe La connaissance de V_i (en tant que variable aléatoire) n'est pas totale, mais conditionnée à ses parents directs. En d'autres termes, la loi de V_i n'est pas connue directement, mais celle de

$V_i|C(V_i)$ l'est. Cela est également visible sur la figure 2.7 : la loi de *herbe mouillée* n'est pas connue directement, mais sa loi conditionnelle aux deux autres variables l'est.

Soit $\mathbb{P}(V) = \mathbb{P}(\cap_{i=1}^{n} V_i) = \mathbb{P}(V_1, ..., V_n)$, la distribution jointe des probabilités. Alors, par définition (Pearl, 1985), et dans une volonté de *complétude* et *consistance*, avec la topologie du graphe, celle-ci s'exprime comme

$$\mathbb{P}(V) = \prod_{i=1}^{n} \mathbb{P}(V_i|C(V_i)). \tag{2.1}$$

Cette écriture apporte l'une des propriétés fondamentales d'un réseau bayésien : chacun de ses sommets est conditionnellement indépendant de ses non-descendants, sachant ses parents directs (Naïm et al., 2011). En partant de cette propriété, l'écriture (2.1) peut être retrouvée, et inversement.

Figure 2.8: Exemple d'une causalité en chaîne

En effet, en prenant l'exemple de la figure 2.8, si cette propriété n'avait pas été instaurée, $C|B$ aurait très bien pu être dépendant de A. Cela aurait soulevé la question de savoir pourquoi ne pas avoir ajouté une flèche directement de A à C, en plus de celle allant vers B (comme dans l'exemple de la figure 2.7). C'est donc pour assurer la cohérence entre la structure de dépendance et sa représentation graphique que cette propriété est essentielle. Si l'on suppose que $C|B$ est indépendant de A, alors $\mathbb{P}(A, B, C)$ s'écrit :

$$\mathbb{P}(A, B, C) = \mathbb{P}(C|A, B)\mathbb{P}(A, B) = \mathbb{P}(C|B)\mathbb{P}(A, B) = \mathbb{P}(C|B)\mathbb{P}(B|A)\mathbb{P}(A).$$

En adoptant une approche plus "probabiliste", on appelle **parents markoviens** (notés VA_i) de la variable aléatoire V_i le plus petit ensemble de variables aléatoires tel que $V_i|VA_i$ soit indépendant du reste de ses non-descendants (Pearl, 2013), p.14.

Il peut être prouvé que si $\mathbb{P}$ est strictement positive, alors il existe un unique réseau bayésien pouvant représenter cette structure de probabilité et $C(V_i) = VA_i$ (Pearl, 1988).
De plus, s'il existe, pour un graphe G donné, une probabilité $\mathbb{P}$ vérifiant (2.1), alors G et $\mathbb{P}$ sont dits **Markov-compatibles** (Pearl, 2013), p.16. Ces notions permettent de justifier l'équivalence entre graphe et probabilité dans la structure de causalité illustrée par la figure 2.8.

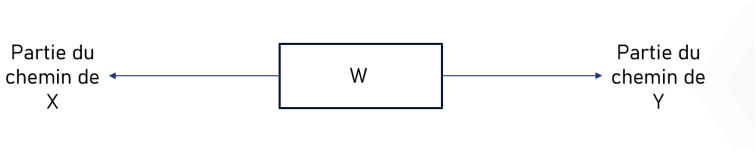
D-séparation Définissons un concept permettant de lier la topologie du graphe à la structure de dépendance dans l'objectif de pouvoir *inférer* des informations au sein de cette structure de connaissance.

On dit que deux sommets (variables aléatoires) X et Y sont *d-séparés* par un ensemble de sommets $\mathbf{Z} = (Z_1, ..., Z_k)$ si, pour tous les chemins (dans le sens non orienté) entre X et Y, au moins l'une des deux conditions suivantes est vérifiée :

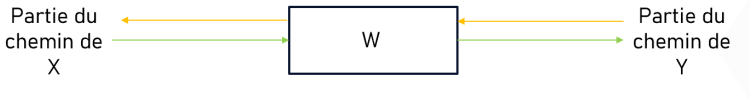
1. Le chemin **converge** (voir figure 2.9) vers un sommet W tel que $W \notin \mathbf{Z}$, et W n'est pas un parent direct d'un élément de $\mathbf{Z}$.

2. Le chemin passe par un sommet $W \in \mathbf{Z}$, et est soit **divergent**, soit **en série** en ce sommet (voir figure 2.10).

Figure 2.9: Visualisation de la notion de convergence d'un chemin



(a) Visualisation d'un chemin divergeant en un sommet



(b) Visualisation d'un chemin en série en un sommet

Figure 2.10: Visualisation des notions de divergence et de série

Cette définition peut être étendue à deux ensembles de nœuds $\mathbf{X} = (X_1, ..., X_{k_1})$ et $\mathbf{Y} = (Y_1, ..., Y_{k_2})$, en disant que $\mathbf{X}$ est d-séparé de $\mathbf{Y}$ par $\mathbf{Z}$ si chaque élément de l'un est d-séparé de chaque élément de l'autre par $\mathbf{Z}$.

Indépendance et D-séparation La d-séparation est une notion "graphique". Néanmoins, par la construction probabiliste d'un réseau bayésien, celle-ci peut être liée à des propriétés probabilistes des variables aléatoires qui constituent G. En effet, il peut être prouvé que si X est d-séparé de Y par $\mathbf{Z}$, alors X est indépendant de Y pour toute probabilité Markov-compatible avec le graphe. De plus, si X et Y ne sont pas d-séparés, alors ils ne sont pas indépendants (Pearl, 2013).

Enfin, il peut aussi être prouvé qu'une variable aléatoire V_i est d-séparée du reste du graphe par l'ensemble de ses parents, de ses enfants et des autres parents de ses enfants (Naïm et al., 2011).
Ces notions permettent de rendre locaux tous les calculs dans un graphe causal, c'est-à-dire que deux parties d-séparées peuvent être calculées séparément, optimisant ainsi les calculs dans le graphe.

Inférence Comme vu avec l'exemple de la pluie et de l'herbe, il ne suffit pas de suivre les flèches pour comprendre comment l'*information* se déplace dans le graphe ; l'ordre est important (Naïm et al., 2011). L'objectif dans un graphe est souvent de calculer la probabilité conditionnelle d'une variable aléatoire par rapport à une autre (ou un ensemble d'autres). Ce calcul s'appelle l'*inférence*.

D'un point de vue théorique, ce calcul est parfaitement trivial au vu des propriétés de Bayes. En effet, en prenant X et Y dans V (des variables aléatoires (sommets) ou des sous-ensembles de V (ensemble de sommets)) et S l'ensemble des variables aléatoires de G qui ne sont ni X ni Y, nous avons

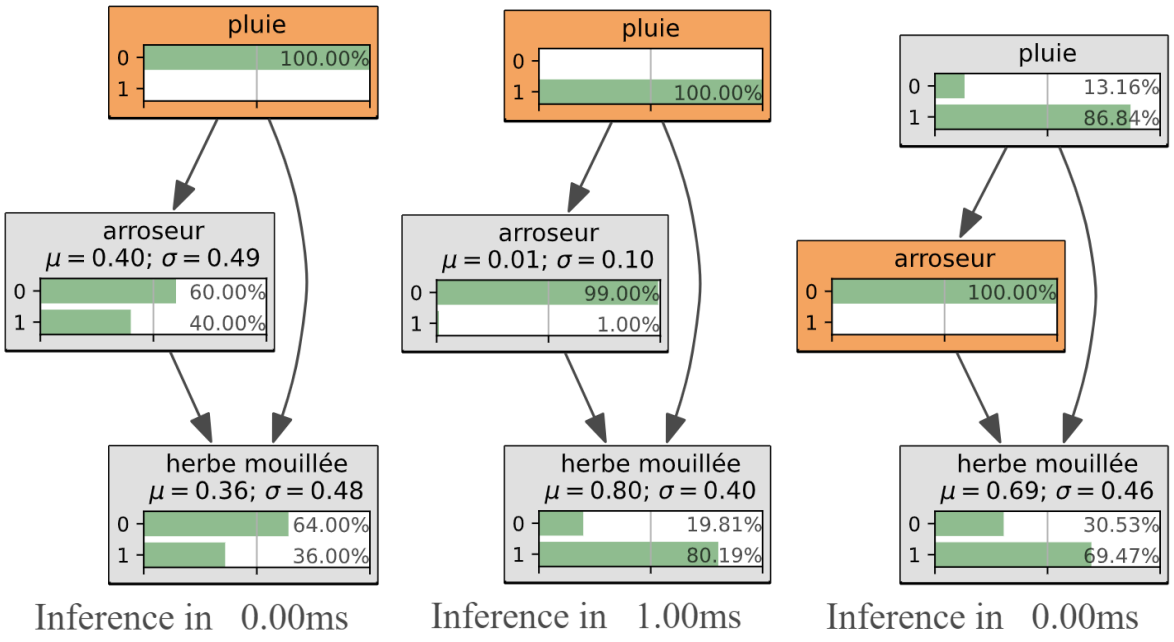
$$\mathbb{P}(Y = y|X = x) = \frac{\sum_s \mathbb{P}(Y = y, X = x, S = s)}{\sum_{y,s} \mathbb{P}(Y = y, X = x, S = s)}. \tag{2.2}$$

Il est à noter qu'ici les variables aléatoires sont supposées discrètes. La généralisation au cas continu sera discutée dans le paragraphe sur les Réseaux Bayésiens Hybrides (2.1.2.4).

Cependant, toute la complexité de ce problème réside dans le fait de le réaliser de manière optimale

en considérant uniquement les connaissances données par la topologie du graphe G. Ce problème, à première vue simple, a été démontré comme NP-difficile et particulièrement dépendant de la structure du graphe par G.F. Cooper en 1990 (Cooper, 1990).

En reprenant l'exemple de la figure 2.7, il serait possible de vouloir calculer $\mathbb{P}$(Herbe Mouillée|Pluie = Vrai). Dans ce cas, il faudrait d'abord calculer la probabilité que l'arroseur soit allumé (qui est de 0.01) pour ensuite pouvoir inférer la probabilité que l'herbe soit mouillée.



(a) Inférence avec la connaissance qu'il ne pleut pas

(b) Inférence avec la connaissance qu'il pleut

(c) Inférence avec la connaissance que l'arroseur n'est pas allumé

Figure 2.11: Exemples d'inférences sur le cas de la figure 2.7

La figure 2.11 présente plusieurs exemples d'inférences. Dans le cas (a), il est inféré qu'il ne pleut pas. Un changement est alors observé dans la probabilité de l'arroseur, mais aussi dans celle de l'herbe mouillée. De même, en (b), lorsque qu'il pleut, les résultats changent en faveur d'une herbe mouillée et d'un arroseur éteint. Enfin, comme observé en (c), il est possible d'inférer à partir d'une connaissance qui ne se trouve pas à la racine du graphe. Par exemple, en sachant que l'arroseur est éteint, des changements sont observés sur la probabilité que l'herbe soit mouillée.
Dans cet exemple, la logique peut paraître simple, mais avec une structure de dépendance plus complexe, les calculs deviennent rapidement plus lourds.

Plusieurs bibliothèques sont disponibles pour l'étude et le calcul de ce type de graphe. En ce qui concerne Python, l'une d'elles est *PyAgrum*, qui est utilisée pour l'exemple de la figure 2.11. Cette bibliothèque repose sur *aGrUM*, une bibliothèque en C++, et propose un ensemble complet d'outils pour la visualisation et la modélisation de réseaux bayésiens.
Notons par ailleurs que *PyAgrum* affiche des valeurs μ et σ sur ses variables inférées (comme sur la figure 2.11). Celles-ci correspondent à la moyenne et à l'écart-type de la variable aléatoire. Ici, ces paramètres n'ont pas un intérêt majeur puisque toutes nos variables n'ont que deux états, ce qui signifie que $\mu = P(X = 1)$ et $\sigma = \sqrt{\mu(1-\mu)}$. Toutefois, si les variables avaient plus d'états (et donc une distribution plus complexe), ces mesures seraient plus pertinentes pour comprendre la répartition des probabilités.

La bibliothèque utilise des algorithmes efficaces pour le calcul d'inférences, tels que la *Lazy Propagation* et la *Shafer-Shenoy Inference* pour une inférence exacte, ainsi que la *Loopy Belief Propagation* pour une inférence approximative lorsque l'inférence exacte devient trop coûteuse (PyAgrum, n.d.).

2.1.2.4 Au-delà du réseau bayésien

Dans la partie précédente, les réseaux bayésiens ont été introduits dans leur ensemble, selon leur définition la plus fondamentale. Cependant, certaines remarques complémentaires concernant l'évolution de ces réseaux sont particulièrement intéressantes et ouvrent de nouvelles perspectives de modélisation.

Les réseaux bayésiens hybrides Dans l'ensemble de nos exemples, les variables utilisées sont discrètes. Néanmoins, le concept de réseau bayésien peut s'étendre à un mélange de variables continues et discrètes sans modifier la définition de base ni aucune des propriétés énoncées. Ce type de réseau bayésien est appelé **réseau bayésien hybride** (*HBN* en anglais). Ils sont de plus en plus utilisés dans divers domaines tels que l'étude des réseaux, la psychologie, la gestion des risques et la cybersécurité. Le papier de Xing et al., 2017 présente, par exemple, une application des réseaux bayésiens pour reconstruire les réseaux de régulation génique (GRN), illustrant une application de ces réseaux dans le domaine génétique.

Apprentissage de la structure Comme précisé précédemment, la force des réseaux bayésiens réside dans leur flexibilité. Dans le cadre de l'apprentissage automatique, il est possible d'apprendre la structure d'un réseau bayésien à partir de données, permettant ainsi de créer un réseau causal ou de corrélation de manière automatique. Par exemple, Zhu and Nguyen, 2022 propose des fondements théoriques pour l'apprentissage de la structure des *HBN*.
L'objectif est, connaissant V et en notant $\mathbb{P}(V|G)$ ce qui avait été noté par simplicité $\mathbb{P}(V)$ dans l'équation (2.1), d'étudier la structure G du graphe en remarquant que

$$\mathbb{P}(V|G) \propto \mathbb{P}(G|V)\mathbb{P}(G),$$

où $\propto$ signifie "proportionnel à".

Apprentissage des paramètres Il est également possible d'apprendre les "paramètres" d'un réseau bayésien, c'est-à-dire de déterminer les probabilités conditionnelles associées aux relations entre les nœuds du réseau. Ces paramètres quantifient la force et la nature des dépendances entre les variables. De manière usuelle, les nœuds dans un réseau bayésien sont représentés par des variables discrètes ou gaussiennes. Il est alors possible d'apprendre les paramètres de la loi conditionnelle en fonction des données.

L'apprentissage des paramètres peut se faire à partir de données observées en utilisant des méthodes telles que l'estimation du maximum de vraisemblance ou l'estimation bayésienne. L'estimation du maximum de vraisemblance ajuste les probabilités conditionnelles pour maximiser l'accord entre le modèle et les données observées. L'estimation bayésienne, quant à elle, incorpore des distributions a priori et les met à jour à mesure que de nouvelles données sont observées, offrant ainsi une approche plus flexible et robuste dans des contextes incertains ou avec des données limitées.

Réseau bayésien dynamique Il est également possible d'étendre la définition des réseaux bayésiens en y ajoutant une dimension temporelle, permettant de modéliser l'évolution des variables composant

G à des temps discrets ($t \in \mathbb{N}$). Il est important de noter que, dans ce type de réseau, ce sont les variables qui évoluent dans le temps et non la structure du graphe (E).

Comme présenté dans Murphy, 2002, un réseau bayésien dynamique peut être défini comme une paire $(G_1, G_\rightarrow)$, où G_1 est un réseau bayésien définissant la probabilité initiale $\mathbb{P}(V^1) = \mathbb{P}(V_1^1, ..., V_n^1)$ (le réseau au temps initial) et où $G_\rightarrow$ est un réseau bayésien définissant $\mathbb{P}(V^t|V^{t-1})$ comme

$$\mathbb{P}(V^t|V^{t-1}) = \prod_{i=1}^{n} \mathbb{P}(V_i^t|C_\rightarrow(V_i^t)). \tag{2.3}$$

où V_i^t est le i-ème nœud au temps t et $C_\rightarrow(V_i^t)$ représente les parents de ce nœud dans le graphe $G_\rightarrow$. G_1 est utilisé uniquement pour l'initialisation, tandis que l'évolution est décrite par $G_\rightarrow$, permettant aux parents d'être issus du même temps ou du temps $t-1$.

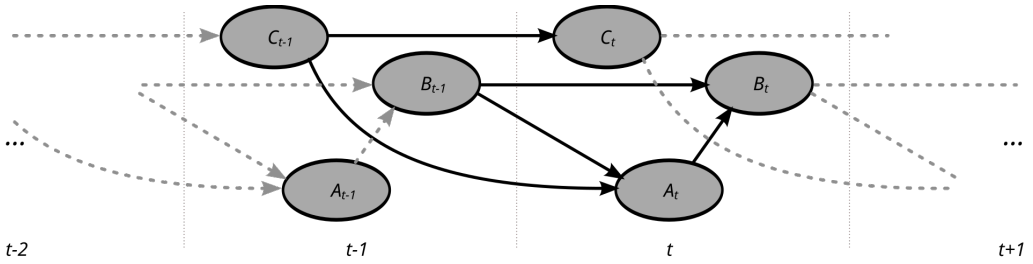


Figure 2.12: Exemple de réseau bayésien dynamique, présentation d'un $G_\rightarrow$ (source : Wikipédia)

Dans la figure 2.12, on observe qu'un nœud peut avoir pour parent des variables du même temps ou d'un temps $t-1$. Par exemple, pour B_t, on constate que $C_\rightarrow(B_t) = \{A_t, B_{t-1}\}$. Si B_t représente la température d'une pièce au temps t et A_t l'activité du chauffage, alors le graphe montre que la température dépend à la fois de celle de $t-1$ et de l'activité du chauffage à t. Cela permet de prendre en compte à la fois des influences temporelles et causales. La structure du graphe reste inchangée entre chaque instant, ce qui est illustré par les flèches en pointillés. Il est aussi possible de définir des variables statiques influençant les variables évoluant dans le temps.

Ce type de modèle se rapproche des modèles de Markov cachés (*HMMs*) et des filtres de Kalman, comme discuté dans Murphy, 2002. Les RBDs sont appliqués dans des domaines tels que la finance pour la prévision des tendances de marché, la robotique pour le suivi d'état et la prise de décision en temps réel, la bio-informatique pour l'analyse de séries temporelles, et l'intelligence artificielle pour le traitement du langage naturel et la reconnaissance vocale.

Dans cette partie, nous avons introduit un ensemble de concepts théoriques. Tout d'abord, la théorie des graphes, base des modèles graphiques, a été présentée. Ensuite, les réseaux bayésiens, à l'intersection entre modèle graphique, structure probabiliste et causalité, ont été explorés. Nous avons abordé leur évolution récente à travers les notions d'apprentissage automatique et de réseaux bayésiens dynamiques.

2.2 Une revue des approches graphiques dans le monde du cyber

Cette section s'intéressera aux approches graphiques et aux applications des concepts théoriques vus dans la section précédente. Dans un premier temps, une vision propre à la cybersécurité sera présentée, mettant en place une application des réseaux bayésiens pour quantifier le risque cyber au sein d'une entreprise. Les apports et limites de cette démarche seront discutés, en expliquant pourquoi

une vision plus **micro** est nécessaire. Dans un second temps, une vision cyber-assurantielle basée sur des modèles graphiques sera introduite. Les différents aspects de ce modèle ainsi que les apports et les limites à combler dans la prochaine section seront détaillés et discutés.

L'objectif principal de cette section est de démontrer qu'une recherche est en cours pour mettre en commun des connaissances de cybersécurité et assurantielles, dans l'objectif de quantifier ce risque. L'enjeu de la quantification du risque cyber ne se limite pas au monde assurantiel, mais s'étend également aux entreprises, elles-mêmes dans la nécessité de mieux connaître leur risque.

2.2.1 Méthode FAIR bayésienne - Une recherche de quantification pour les entreprises

Cette partie présentera l'adaptation des réseaux bayésiens à une méthode visant à faciliter la communication du risque cyber entre les décisionnaires et les pôles techniques.

L'objectif est de montrer que les réseaux bayésiens et les approches graphiques en général sont au cœur de l'étude de la quantification du risque cyber.

Plusieurs documents montrent un intérêt croissant pour ce type de modèle dans le monde de la cybersécurité. Chockalingam et al., 2017 présente une revue systématique des modèles de réseaux bayésiens standards en cybersécurité. Cette étude met en avant un large éventail d'adaptations, certains modèles se basant sur la connaissance d'experts, d'autres sur des données empiriques. Cette hétérogénéité se retrouve également au niveau des objectifs poursuivis.

Dans cette partie, le modèle FAIR sera introduit, puis son adaptation au cadre bayésien sera expliquée. Enfin, une discussion sur la difficulté d'adaptation de ce modèle à un cadre assurantiel sera menée.

2.2.1.1 La méthode FAIR

L'un des principaux enjeux dans le domaine de la cybersécurité est de transformer les problématiques techniques, qui préoccupent les experts (failles, vulnérabilités, processus d'attaque), en une quantification économique du risque pouvant être discutée au sein des sphères décisionnaires de l'entreprise, facilitant ainsi la prise de décision basée sur des résultats quantitatifs (FAIR Institute, n.d.).

C'est dans cet objectif qu'a été conçu le *framework* FAIR. Il a pour vocation de créer un pont entre le monde technique et le monde décisionnaire d'une entreprise. Ce *framework* se divise en deux parties : une partie taxonomique qui présente l'architecture de la division du risque cyber sous la forme d'un graphe, et une partie statistique qui permet d'appliquer la quantification sur cette taxonomie. Ce cadre rend la méthode adaptable et compréhensible dans un contexte de décision (par exemple, si ne pas *patcher* une faille A engendre un risque supplémentaire de 100€, tandis que la faille B représente un risque de 1000€, alors la priorité sera donnée à la correction de la faille B). Cette approche rend la quantification accessible aux décisionnaires et permet une prise de décision éclairée.

On peut observer sur la figure 2.13 que la taxonomie est divisée en deux parties : une partie coût et une partie fréquence, ce qui rappelle la méthodologie actuarielle classique (coût x fréquence). Les sous-catégories du graphe définissent l'ensemble des influences pour chaque catégorie.

2.2.1.2 Application des réseaux bayésiens à la méthode FAIR

Comme l'explique Wang et al., 2020, bien que le modèle de base présente de nombreuses qualités, sa structure statistique reste perfectible (seules des distributions triangulaires sont utilisées, rendant

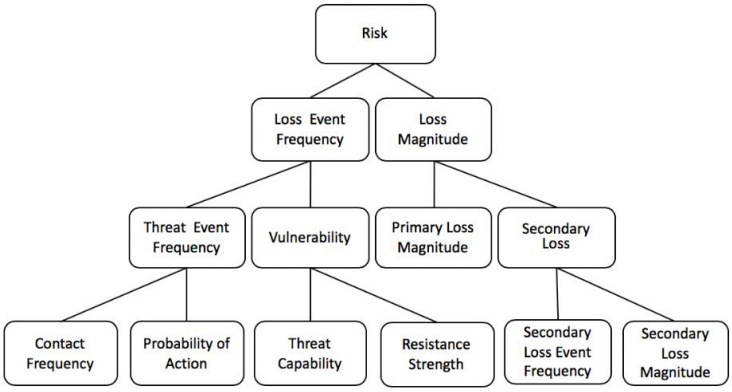


Figure 2.13: Taxonomie de FAIR - (Wang et al., 2020)

la quantification approximative). Néanmoins, il apparaît clairement, à travers la structure graphique de la taxonomie, que le modèle FAIR peut être aisément modélisé sous la forme d'un réseau bayésien.

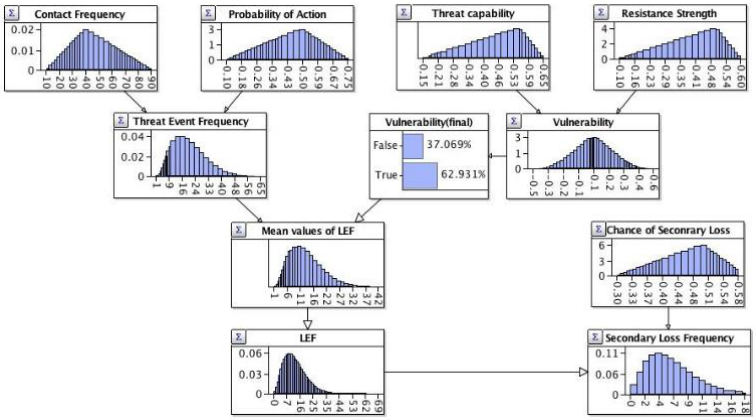


Figure 2.14: Réseau bayésien FAIR - (Wang et al., 2020)

C'est précisément ce que démontre Wang et al., 2020, qui illustre l'adaptabilité de la méthode en testant diverses distributions de probabilité pour les différentes sous-catégories. Ils mettent en avant les bons résultats et la stabilité du modèle en le comparant à d'autres méthodes d'application du cadre FAIR. La modélisation est visible sur la figure 2.14, où chaque case de la taxonomie est représentée par une variable aléatoire, formant ainsi un réseau bayésien hybride.

2.2.1.3 Intérêt et limites dans le cadre cyber-assurantiel

Ces modèles sont prometteurs dans un contexte d'entreprise. Ils permettent une meilleure connaissance du risque interne et constituent un levier clé pour la gestion des vulnérabilités.
Ils offrent la possibilité de prioriser les actions à mener (par exemple, patcher la vulnérabilité V_1 plutôt que V_2) et permettent à l'entreprise de faire des choix stratégiques, comme déterminer quelle part du risque doit être assurée et quelle part doit être conservée en interne.

Néanmoins, dans le contexte de ce mémoire, cette méthodologie présente certaines limites. Tout d'abord, elle exige une connaissance détaillée de l'entreprise et ne propose aucun moyen d'automatiser le processus, ce qui peut s'avérer peu rentable dans un contexte assurantiel. Ensuite, cette approche a été conçue pour des grandes entreprises et peut ne pas être adaptée aux petites structures, qui con-

stituent un marché dans lequel l'assurance cyber est encore peu développée. Sa vision trop **macro** rend difficile l'application d'une méthodologie unique entre différentes entreprises. Il est donc nécessaire, pour nos besoins, de recourir à une méthode permettant d'aller plus en détail.

2.2.2 Un regard sur les failles dans le réseau de l'entreprise - Une vision micro

La partie précédente présentait un modèle utilisé par les entreprises pour permettre une quantification plus précise de leur risque. Néanmoins, sa vision trop globale (vue **macro**) et ses préceptes parfois peu clairs rendent difficile l'adaptabilité de ces idées dans un cadre assurantiel. Dans cette partie, une vision plus **micro** et adaptée à l'assurance sera présentée. En deux étapes, l'accent sera mis sur la propagation de l'attaquant dans le réseau de l'entreprise et sur l'impact financier subi par l'entreprise après une perte cyber (serveur, ordinateur, etc.).

Nous nous baserons principalement sur l'étude du document de la Society Of Actuaries portant sur la quantification du risque cyber à l'aide d'une approche économique-fonctionnelle (Tatar et al., 2020). Ce document présente une approche novatrice permettant d'évaluer le risque cyber d'une entreprise dans un objectif assurantiel. Il propose une méthode ne nécessitant pas de base de sinistres, en s'appuyant uniquement sur l'étude (et la probabilisation) des attaques possibles et en estimant leur coût économique pour l'entreprise.

Dans un premier temps, les graphes d'attaque seront étudiés. Ces derniers constituent un outil mis en place par les experts en cybersécurité pour mieux comprendre la topologie d'une attaque. Ensuite, il sera montré comment, à partir de ce graphe, construire une structure probabiliste permettant d'estimer le risque de perte d'un actif lors d'une attaque. Enfin, les graphes de dépendances seront introduits, offrant un moyen de transformer une perte cyber en perte économique pour l'entreprise.

2.2.2.1 Graphe bayésien d'attaque - Probabiliser l'évolution de l'attaquant dans l'entreprise

Le chapitre 1 présentait au lecteur la démarche suivie par un attaquant pour s'infiltrer dans un système. Celui-ci exploite des vulnérabilités, qu'elles soient humaines ou techniques, afin d'acquérir des privilèges et de passer d'une cible à une autre (1.1.3). Les connaissances en cybersécurité permettent d'obtenir une liste de ces vulnérabilités (CVE — voir 1.1.3.2) ainsi que des programmes qu'elles affectent (CPE — voir 1.1.3.4). De plus, les scores CVSS permettent d'évaluer ces vulnérabilités à l'aide de critères de dangerosité et d'exploitabilité (voir 1.1.3.4).

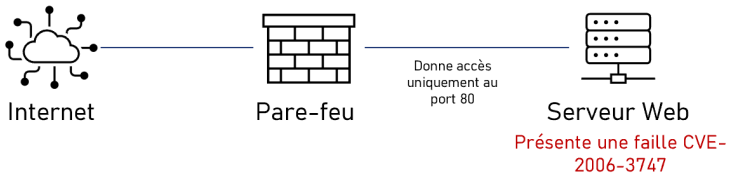
L'objectif de cette section est de franchir le pas assurantiel en transformant l'ensemble de ces connaissances en une quantification de la probabilité de progression de l'attaquant au sein de l'entreprise.

Dans un premier temps, les graphes d'attaque seront examinés, permettant de relier des cibles en fonction de leurs vulnérabilités. Ensuite, la transformation de ce graphe en un graphe probabiliste sera présentée à travers l'introduction des graphes bayésiens d'attaque.

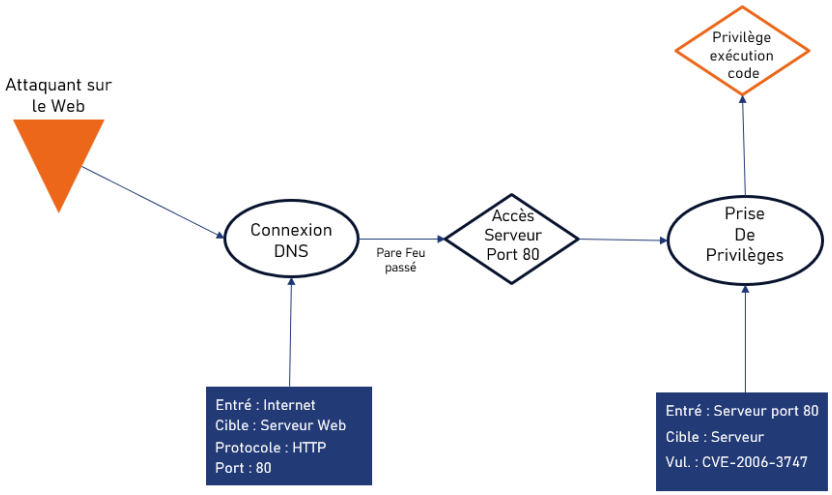
Les graphes d'attaque Comme expliqué dans Tatar et al., 2020, les réseaux d'entreprises (qui peuvent être assimilés à des graphes) sont composés de centaines de nœuds (ordinateurs, routeurs, serveurs, etc.). Compter simplement les vulnérabilités des composants d'un réseau n'est pas une méthode efficace pour évaluer les risques cyber. En effet, de nombreuses vulnérabilités ne peuvent pas être exploitées immédiatement, car les défenses multicouches empêchent les attaquants d'accéder directement à l'hôte visé. De plus, certaines vulnérabilités ne sont pas exploitables du tout. Pour atteindre leur cible, les attaquants doivent analyser la topologie du réseau et utiliser les nœuds compromis pour accéder à de nouvelles cibles accessibles à partir de ces derniers (Mensah, 2019).

Estimer les chemins d'attaque (l'ordre d'exploitation des vulnérabilités pour atteindre une cible) permet d'avoir une vision plus claire des routes que pourra emprunter l'attaquant, tant d'un point de vue cybersécurité, pour identifier les points critiques, que d'un point de vue cyberassurance, pour mieux quantifier le risque porté par l'entreprise et ainsi lui fournir des conseils adaptés. Ces routes sont constituées d'une suite d'exploitations de vulnérabilités, comme présenté dans la matrice ATT&CK.

Un formalisme, appelé **graphe d'attaque**, permet de modéliser précisément l'espace d'évolution de l'attaquant. Introduit par Phillips and Swiler, 1998, il permet de représenter, sous la forme d'un graphe orienté, les combinaisons de vulnérabilités permettant d'atteindre une cible dans le réseau. Depuis 1998, ces graphes ont suivi l'évolution de la cybersécurité, et une représentation moderne peut être observée sur la figure 2.15.



(a) Visualisation d'un réseau simple avec faille CVE



(b) Graphe d'attaque utilisant la vulnérabilité CVE

Figure 2.15: Processus d'attaque sur un réseau simple modélisé avec un graphe d'attaque

Cette présentation reprend celle introduite dans Singhal and Ou, 2011. Dans la sous-figure 2.15a, l'architecture du réseau est représentée. Celui-ci est composé d'un pare-feu et d'un serveur. Le pare-feu autorise uniquement l'accès au port 80 du serveur (port utilisé pour les connexions web), et le serveur présente une vulnérabilité.
La sous-figure 2.15b illustre le graphe d'attaque. Le triangle vert représente la position initiale de l'attaquant et les différentes cibles qu'il peut atteindre. Les ovales représentent l'exploitation de vulnérabilités ou de protocoles (comme le protocole HTTP sur le port 80). Les rectangles jaunes indiquent les configurations nécessaires à l'exploitation, tandis que les trapèzes symbolisent les conséquences de l'exploitation.

Bien que complète, cette visualisation du graphe peut devenir lourde avec l'augmentation du nombre de nœuds. Tatar et al., 2020 propose ainsi de simplifier la représentation du graphe d'attaque

en utilisant des nœuds pour symboliser l'accès de l'attaquant et des arêtes pour représenter les exploitations utilisées pour passer d'un nœud à un autre. D'autres études sur la quantification du risque cyber, comme celle de Lau et al., 2021, préfèrent conserver cette visualisation détaillée.

Génération d'un graphe d'attaque Pour générer un graphe d'attaque, il est nécessaire de disposer des éléments suivants (Tatar et al., 2020) :

1. La liste des vulnérabilités présentes sur le réseau (en lien avec les bases CVE).
2. La topologie du réseau et les différentes configurations.
3. Une base de données des différents types d'attaques et failles connues (toute l'intelligence de MITRE).

Concernant la création de ce type de graphe dans un réseau d'entreprise, plusieurs logiciels permettent d'automatiser le processus. Différentes solutions sont présentées dans Tatar et al., 2020.
Dans un contexte assurantiel, l'automatisation de cette étape réduit considérablement les coûts et le temps nécessaires lors d'une souscription. Elle permet également, par exemple, d'effectuer des scans périodiques chez l'assuré afin de mettre à jour la structure du graphe et d'obtenir une évaluation continue du risque.
Une discussion plus approfondie sur ce sujet a d'ailleurs été menée dans la sous-section 3.1.3.1.

Graphe Bayésien d'Attaque Le graphe d'attaque en tant que tel n'est pas suffisant pour probabiliser la réussite d'une attaque sur une cible. L'objectif est alors de créer une structure de probabilité autour de celui-ci en le transformant en une sorte de réseau bayésien.
Pour ce faire, Tatar et al., 2020 se base sur la théorisation du principe faite par Wang et al., 2008.

L'idée est d'associer à chaque exploitation de vulnérabilité constituant le graphe d'attaque une probabilité de réussite, sachant que les préconditions sont validées.
D'une manière théorique, si $G = (V, E)$ est le graphe orienté d'attaque (considéré comme acyclique), où V est l'ensemble des cibles et les éléments de E sont les vulnérabilités permettant de passer de V_i à V_j (à la manière de Tatar et al., 2020), alors, d'après Wang et al., 2008 :

- Chaque élément e de E est pondéré par $p(e)$, la probabilité d'exploitation réussie de la faille.
- Une fois cette structure définie, deux visions ont été notées pour le calcul du graphe :
  1. La première, celle de Tatar et al., 2020, propose que pour chaque cible V_i, il soit possible de calculer la probabilité que celle-ci soit atteinte (à condition que l'attaquant mène une attaque). Cette probabilité sera notée $\mathbb{P}_D(V_i)$. Pour ce faire, plusieurs cas :
     (a) V_i n'a qu'un seul parent V_j. Dans ce cas, en nommant e l'arête les reliant, il est défini que
     $$\mathbb{P}_D(V_i) = p(e) \times \mathbb{P}_D(V_j).$$
     (b) V_i a plusieurs parents (V_{k_j}). Dans ce cas, en nommant e_j l'arête liant V_{k_j} à V_i et $e = \text{argmax}_{1 \leq j \leq n}(p(e_j))$, alors
     $$\mathbb{P}_D(V_i) = p(e) \times \mathbb{P}_D(\bigcup_{j=1}^{n} V_{k_j}).$$

Les $(V_{k_j})_{1\leq j\leq n}$ sont mutuellement indépendants (du fait de l'acyclicité du graphe). Donc, pour deux parents, nous aurions par exemple

$$\mathbb{P}_D(V_i) = p(e) \times \mathbb{P}_D(V_{k_1} \cup V_{k_2}) = p(e) \times (\mathbb{P}_D(V_{k_1}) + \mathbb{P}_D(V_{k_2}) - \mathbb{P}_D(V_{k_1} \cap V_{k_2}))$$

$$= p(e) \times (\mathbb{P}_D(V_{k_1}) + \mathbb{P}_D(V_{k_2}) - \mathbb{P}_D(V_{k_1})\mathbb{P}_D(V_{k_2})).$$

Le choix de calculer par rapport à l'argmax est fait dans (Tatar et al., 2020). Ce choix a été peu compris et est peu discuté dans le document. Il ne sera donc pas repris dans notre modèle.

Ce type de graphe *"bayésianisé"* diffère légèrement des réseaux bayésiens par le calcul des probabilités et le sens qu'apporte une flèche entre deux arêtes. En effet, l'inférence est directe et se fait par récurrence par rapport aux nœuds parents (grâce à cette notion de argmax qui facilite les calculs), ce qui n'est pas le cas avec les réseaux bayésiens.

2. La deuxième méthode est proposée par Poolsappasit et al., 2012. Elle présente une implémentation en réseau bayésien du graphe d'attaque. Pour ce faire, la probabilité conditionnelle du nœud x est définie en fonction de l'ensemble des parents de x (noté $C(x)$), comme sur la figure 2.7. Deux logiques de combinaison de failles sont ainsi définies :

   (a) La logique OR signifiant qu'au moins une des failles doit être exploitée (au moins un des parents doit être acquis). Dans ce cas, en nommant T_j le sous-ensemble de $C(V_j)$ qui est connu comme vrai (pris) et e_{V_k} l'arête entre $V_k \in T_j$ et V_j, il est défini que

   $$\mathbb{P}(V_j|C(V_j)) = \begin{cases} 0, \text{ si } T_j = \emptyset, \\ p(\bigcup_{V_k \in T_j} e_{V_k}), \text{ sinon} \end{cases}$$

   avec

   $$p(\bigcup_{V_k \in T_j} e_{V_k}) = 1 - \prod_{V_k \in T_j} (1 - p(e_{V_k})).$$

   (b) La logique AND pour laquelle tous les parents doivent être acquis :

   $$\mathbb{P}(V_j|C(V_j)) = \begin{cases} 0, \text{ si } T_j \neq C(V_j), \\ p(\bigcap_{V_k \in T_j} e_{V_k}) = \prod_{V_k \in T_j} p(e_{V_k}), \text{ sinon.} \end{cases}$$

   La logique OR sera utilisée lorsqu'une seule cible est nécessaire pour continuer le chemin, tandis que la logique AND sera utilisée lorsque l'ensemble des cibles parentes de V_j est nécessaire.

   Après avoir défini le réseau bayésien, il est possible de calculer $\mathbb{P}(V_j)$ par inférence (voir 2.1.2.3 pour la logique de l'inférence).

Il est maintenant possible d'examiner un exemple simple pour mieux comprendre le fonctionnement et les différences entre les deux méthodes. La figure 2.16 présente un graphe d'attaque comportant trois cibles, avec la cible finale étant le serveur de données. Chaque arête est pondérée par $p(e)$, représentant la probabilité de réussite d'exploitation de la vulnérabilité. La figure présente également les tableaux du graphe bayésien créé à partir de la méthode de Poolsappasit et al., 2012, développée dans le point 2 du paragraphe précédent.

Nous pouvons alors chercher à calculer $\mathbb{P}$(Serveur Data) selon les deux méthodes présentées précédemment, en prenant $\mathbb{P}(\text{Attaquant} = \text{Vrai}) = \mathbb{P}(A) = 0.8$ (correspondant à la volonté d'attaque). Ainsi, il sera montré que les deux méthodes produisent des résultats différents et qu'un choix doit être fait.

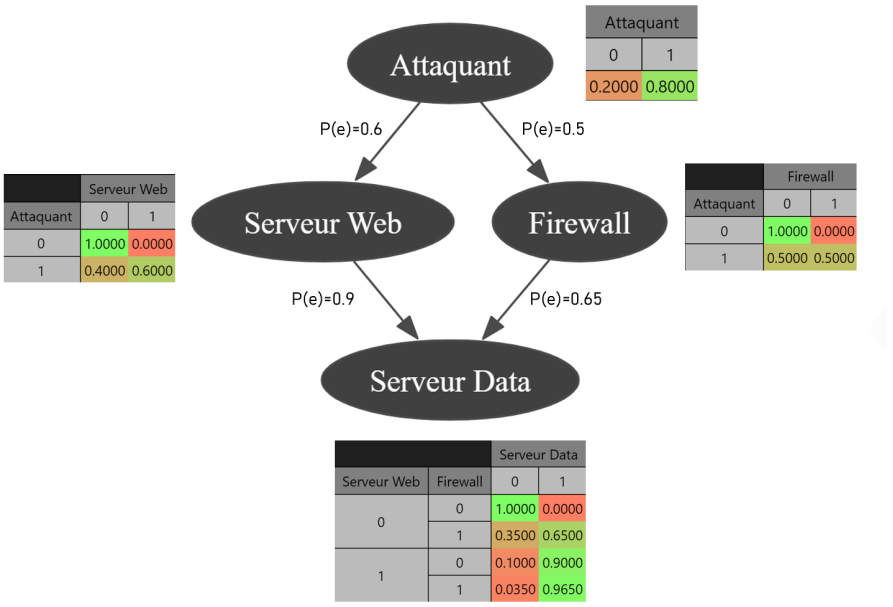


Figure 2.16: Graphe Bayésien d'attaque - Deux méthodes de calcul

1. **En utilisant la méthode de la SOA**

$$\mathbb{P}(\text{Serveur Web} = \text{Vrai}) = p(e_{A\to SW})\mathbb{P}(A) = 0.8 \times 0.6 = 0.48.$$

$$\mathbb{P}(\text{Firewall} = \text{Vrai}) = p(e_{A\to F})\mathbb{P}(A) = 0.8 \times 0.5 = 0.4.$$

Enfin, en appliquant la formule sur le Serveur Data,

$$\mathbb{P}(\text{Serveur Data} = \text{Vrai}) = p(e_{SW\to SD})(\mathbb{P}(SW) + \mathbb{P}(F) - \mathbb{P}(SW)\mathbb{P}(F)) = 0.6192.$$

2. **En utilisant l'inférence** (donc la méthode de Poolsappasit et al., 2012)

Dans un premier temps, il est nécessaire de calculer les probabilités conditionnelles pour chaque variable aléatoire (chaque nœud).

Pour illustrer la méthodologie, l'exemple de "Serveur Data" sera pris. Comme "Serveur Data" (noté SD) a deux parents ("Firewall", noté F, et "Serveur Web", noté SW), la méthode donnée précédemment peut être appliquée. Il est ainsi possible de créer le tableau des probabilités conditionnelles de SD (ici, le lien OR est supposé) :

$$\begin{cases} \mathbb{P}(SD = 1|SW = 0, F = 0) = 0 \\ \mathbb{P}(SD = 1|SW = 1, F = 0) = p(e_{SW}) = 0.9 \\ \mathbb{P}(SD = 1|SW = 0, F = 1) = p(e_F) = 0.65 \\ \mathbb{P}(SD = 1|SW = 1, F = 1) = p(e_{SW}) + p(e_F) - p(e_{SW})p(e_F) = 0.965. \end{cases}$$

Les données du tableau du Serveur Data visibles sur la figure 2.16 sont bien retrouvées. Après calcul de chaque probabilité conditionnelle, il est possible d'inférer sur le graphe et de trouver que

$$\mathbb{P}(\text{Serveur Data} = \text{Vrai}) = 0.5516.$$

Cette valeur est obtenue automatiquement par la méthode d'inférence bayésienne développée dans la section 2.1.2.3.

Comme cela a été commenté précédemment, la méthode de la SOA semble présenter certaines incohérences dans le calcul de la probabilité finale d'un élément (en particulier avec le choix de l'argmax). La différence entre les deux méthodes illustre bien une inconsistance.
La méthode de Poolsappasit et al., 2012 semble plus adaptée, car elle fournit une application rigoureuse des réseaux bayésiens théoriques à la problématique des graphes bayésiens d'attaque. L'autre méthode, présentant des zones d'ombre, il sera donc choisi de continuer avec la vision "réseaux bayésiens" dans la suite de ce mémoire.

Calcul de $p(e)$ Comment calculer les $p(e)$, les probabilités d'exploitation de failles, qui sont la base de cette quantification de la probabilité de prise d'une cible ?

Pour les vulnérabilités, Tatar et al., 2020 propose un calcul grâce au score CVSS 3.1 (le lecteur est renvoyé vers la partie 1.1.3.4 pour un développement plus approfondi de ce score). Le score CVSS 3.1 est légèrement différent du score 4.0 présenté, néanmoins la logique reste la même. À l'heure de l'écriture de ce mémoire, la majorité (si ce n'est l'ensemble) des institutions qui fournissent des scores CVSS utilisent toujours le CVSS 3.1, ce qui rend son utilisation dans le modèle logique.

Au lieu de quatre groupes, le CVSS 3.1 en présente trois (le groupe *Threat* ayant été séparé en *Threat* et *Supplemental* lors du passage de CVSS 3.1 à 4.0). La partie *subsequent system* n'existe plus dans l'impact de la métrique de base. Pour plus de clarté, regardons la figure 2.17.

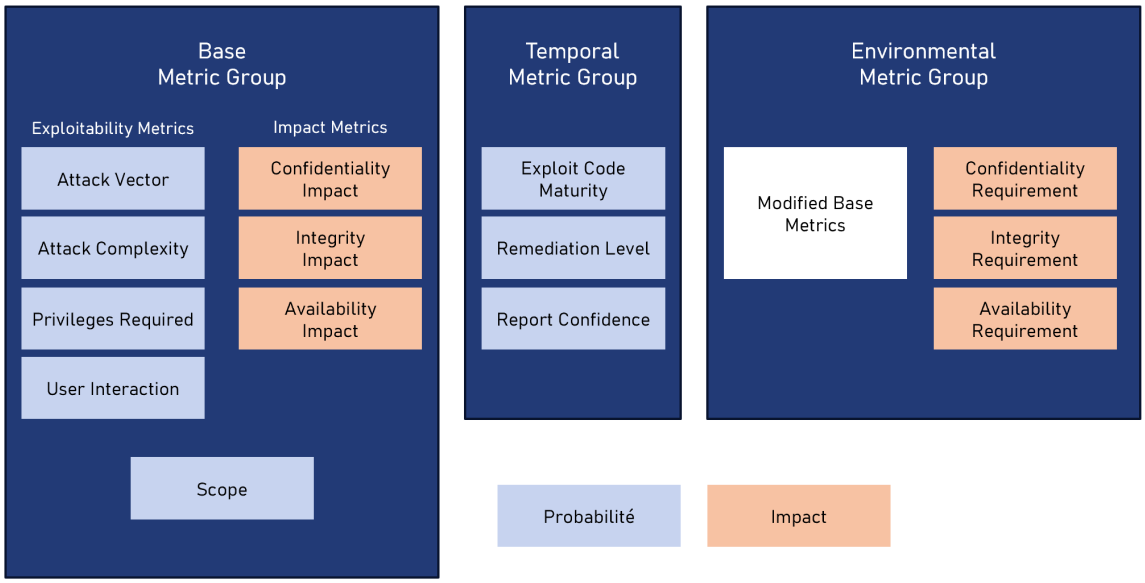


Figure 2.17: Division des métriques entre Impact et Probabilité - Tatar et al., 2020

Les métriques d'exploitabilité et temporelles (en bleu) qualifient la *facilité* avec laquelle la faille peut être exploitée. Ce sont les métriques de probabilité (dans le sens de la "chance de réussite"). Les métriques d'impact (en orange) jugent quant à elles de l'impact sur le système si la faille est exploitée. Cette partie s'intéresse au bloc de probabilité, puisque l'objectif est de quantifier la probabilité d'exploitation d'une vulnérabilité $p(e)$.

Dans le CVSS 3.1, chaque métrique reçoit un score qui dépend de la valeur qu'elle prend. Par exemple, si la métrique "User Interaction" a la valeur *Required*, alors le score associé est de 0.62. Il est ensuite possible de combiner l'ensemble de ces scores pour obtenir le score final sur 10 (qui prend en compte l'ensemble des métriques). Notons que *Scope* n'a pas de valeur en soi, mais modifie les valeurs de certaines métriques lorsqu'il est actif. La table 2.1 présente une partie de ces scores.

Métrique	Valeur Métrique	Valeur Numérique
Attack Vector / Modified Attack Vector	Network	0.85
	Adjacent	0.62
	Local	0.55
	Physical	0.2
Attack Complexity / Modified Attack Complexity	Low	0.77
	High	0.44

Table 2.1: Partie du tableau de conversion des métriques de CVSS 3.1

L'objectif est alors de transformer les valeurs numériques du groupe de métriques "probabilité" de la vulnérabilité en probabilité de réussite, donc en une valeur entre 0 et 1. Soit e_1 la vulnérabilité en question. Tatar et al., 2020, comme d'autres sources citées par le papier, propose de faire le produit de l'ensemble des valeurs. Ainsi,

$$p(e_1) = K \times \text{Attack Vector} \times \text{Attack Complexity} \times \text{Privileges Required} \times \text{User Interaction} \times \text{Exploit Code Maturity} \times \text{Remediation Level} \times \text{Report Confidence}.$$

Avec K la constante permettant de normaliser le résultat entre 0 et 1. Dans ce cas précis, $K = 2.1$.

L'attaquant n'exploitant pas uniquement des failles techniques, mais aussi des failles humaines (au sens large), Tatar et al., 2020 propose également de faire une quantification similaire pour une faille e_i qui dépend du facteur humain (par exemple le phishing pour la prise de privilèges). Il propose une approche avec des métriques (exemple : la quantité de formations de sensibilisation à la sécurité cyber au cours de l'année dernière) et une quantification numérique, similaire au score CVSS. Il fournit ainsi une liste de critères à prendre en compte.

Apports et limites En plus de permettre d'avoir une vision micro de l'évolution du risque au sein de l'entreprise, ce type de quantification du risque est **dynamique**, puisque :

- Le score CVSS dépend de l'évolution de la faille, de ses exploitations et de ses remédiations.
- Il est également dépendant de l'hygiène cyber de l'entreprise, puisque la probabilité d'exploitation des failles humaines dépend des métriques humaines.
- Il est aussi important de souligner que, même dans l'architecture du graphe, cette notion de dynamisme est présente. En effet, l'architecture du graphe d'attaque d'une entreprise très résiliente sera moins complexe, car elle possède moins de vulnérabilités exploitables, par exemple en raison d'une meilleure mise à jour de ses logiciels.
- Il est partiellement automatisé, puisque le graphe d'attaque peut être obtenu à l'aide d'outils informatiques.

Enfin, il ne nécessite pas l'existence d'un historique de sinistres et permet une quantification évolutive du risque qu'une cible soit prise.

Néanmoins, des choix doivent être faits, par exemple sur la transformation des $p(e)$ en probabilité totale, comme il a été discuté tout au long de cette section.

2.2.2.2 FDNA - Une méthodologie pour le coût avec Graphe d'Impact

La section précédente présente une méthodologie permettant d'obtenir la probabilité que différents actifs (ordinateurs, serveurs, etc.) soient capturés par l'attaquant lors de son attaque. Cependant, il n'est pas encore possible d'en tirer une perte économique. Pour ce faire, il est nécessaire de se tourner à nouveau vers le papier de la SOA (Tatar et al., 2020). Cette section présentera la notion de graphe d'impact et de quantification FDNA, dont l'objectif est de transformer une perte physique (sur un serveur, un ordinateur,...) en une perte économique pour l'entreprise.

Graphe d'Impact de Dépendances Le premier concept d'importance est celui de graphe d'impact de dépendances. Proposé par Tatar et al., 2020 et Jakobson, 2011, il offre une vision des liens de dépendance au sein d'une entreprise. Cette dernière peut être représentée par trois couches :

1. Une couche d'*actifs* (*Asset Layer*), composée de l'ensemble des machines (serveurs, ordinateurs,...), logiciels et ressources humaines faisant partie de l'entreprise. L'ensemble des actifs d'une entreprise sera formellement désigné par A par la suite.

2. Une couche de *services* (*Service Layer*), qui s'appuie sur les actifs pour permettre la création de tâches. Une connexion internet, un stockage sur le cloud, etc., sont des services utilisés par une entreprise. Cette couche sera nommée S.

3. Une couche de *business/process* (*Business Layer*), qui est la couche la plus "haute". Elle repose sur les couches précédentes et utilise un ensemble de services (et d'autres processus) pour accomplir un objectif d'entreprise. Par exemple, pour une entreprise de formation en ligne, "Proposer un service de formation en ligne" constitue un business. L'ensemble de ces éléments sera nommé B.

Le graphe d'impact $I_g = (A \cup S \cup B, E)$ est un réseau orienté de dépendances fonctionnelles, où chaque arête e_i représente les dépendances entre les différents nœuds. En effet, si $e_i = (A_1, S_2)$, alors qualitativement, S_2 a besoin de A_1 pour fonctionner correctement. Cette notion sera développée de manière plus précise dans le cadre du cyber. Les dépendances peuvent être aussi bien **intra-couches** qu'**inter-couches**. La figure 2.18 illustre un exemple de ce type de graphe.

L'introduction du Chapitre 1 (1.1.1) présentait l'idée qu'un risque cyber est un risque qui porte sur la **Confidentialité** (C), l'**Intégrité** (I) et/ou la **Disponibilité** (D) d'un ensemble de systèmes d'information. Une attaque cyber a un effet direct sur ces trois catégories au niveau de la couche des actifs, mais elle impactera également la couche de services et celle de business (par exemple, pour une attaque DDoS, la disponibilité du business est affectée). Chaque nœud de I_g peut alors être divisé en trois, comme présenté sur la figure 2.19.

Le graphe d'impact devra intégrer des liens entre les différents sous-nœuds (par exemple, un lien entre la confidentialité de A_1 et la confidentialité de S_2).

FDNA - quantifier les dépendances *Functional Dependency Network Analysis* (FDNA) est une méthode dont l'objectif est de modéliser et mesurer les relations de dépendance au sein de l'entreprise (et avec les fournisseurs) (Garvey, 2009). Elle est utilisée par Tatar et al., 2020 et appliquée au graphe d'impact.

Formellement, il s'agit d'un modèle graphique (d'où sa facilité d'application au graphe d'impact). Cette approche permet de mesurer l'effet ricochet d'une perte d'opérabilité d'un élément de l'entreprise sur le reste des éléments. Encore une fois, ce modèle peut être représenté par un graphe orienté

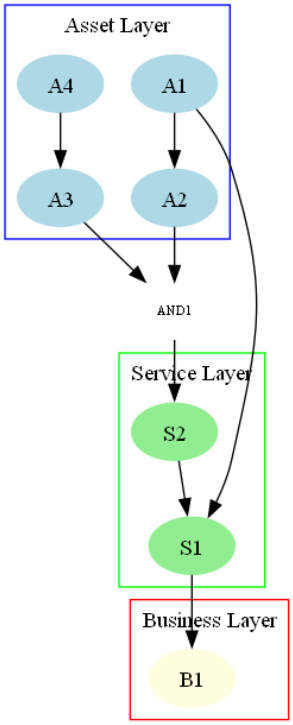


Figure 2.18: Exemple de graphe d'impact

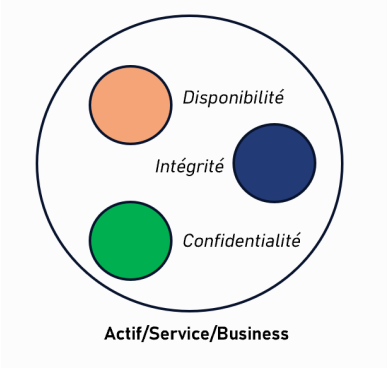


Figure 2.19: Division entre les trois catégories sur un nœud de I_g

$F_g = (V_F, E_F)$ (qui correspondra ici à I_g).
Plusieurs concepts clés sont ainsi définis dans Tatar et al., 2020, qui reprend lui-même les notions de Garvey, 2009. En particulier, pour chaque nœud V_i est défini :

1. Le **Niveau d'opérabilité**, qui correspond à la mesure de l'opérabilité d'un nœud. Nous le noterons P_i. Nous avons $0 \leq P_i \leq 100$. Si $P_i = 100$, l'opérabilité du nœud est maximale; si elle est à 0, le nœud est inopérable.

2. Le concept de **Nœud receveur** : un nœud pour lequel la liste de ses parents est non vide ($w^-(V_i) \neq \emptyset$). Cela signifie que l'opérabilité de ce nœud dépend de celle d'autres nœuds dans le graphe (par exemple, une machine vissant des boulons a besoin des boulons pour fonctionner).

3. Le concept de **Nœud d'alimentation** : un nœud pour lequel la liste de ses enfants est non vide ($w^+(V_i) \neq \emptyset$). Cela signifie que l'opérabilité des enfants dépend de celle de ce nœud.

4. Le **Niveau d'opérabilité de base** (NOB), correspondant au niveau d'opérabilité du nœud receveur lorsque le nœud d'alimentation est totalement inopérable ($P = 0$).

5. La **Force de dépendance** (FDD), qui quantifie la force avec laquelle le nœud receveur est lié au nœud d'alimentation. Pour ce faire, considérons deux nœuds V_i et V_j, avec le premier étant un nœud d'alimentation et le second un nœud receveur. Garvey, 2009 définit alors la contrainte de FDD notée $\alpha_{ij} \in [0; 1]$. Il est alors spécifié que

$$FDD_{ij} = \alpha_{ij} P_i + (1 - \alpha_{ij}) \cdot 100.$$

   Ainsi, plus α_{ij} tend vers 1, plus les variations de P_i se "retrouveront" dans P_j, et moins V_j aura une opérabilité indépendante de V_i. Notons alors que, pour l'instant, $NOB_{ij} = (1 - \alpha_{ij}) \cdot 100$.

6. La **Criticité de dépendance** (CDD) mesure la perte d'opérabilité d'un nœud sous la BOL due à l'arrêt du nœud d'alimentation (ce qui peut exprimer, par exemple, l'effet d'un arrêt prolongé d'un nœud d'alimentation causant des problèmes qu'il n'aurait pas causés autrement). Garvey, 2009 pose alors $\beta_{ij} \in [0; (1 - \alpha_{ij}) \cdot 100]$ et définit

$$CDD_{ij} = P_i + \beta_{ij}.$$

   Ainsi, pour $\beta_{ij} < (1 - \alpha_{ij}) \cdot 100$, le NOB sera plus faible que celui défini par $(1 - \alpha_{ij}) \cdot 100$, ce qui changera avec la criticité de dépendance.

7. L'**Efficacité Propre** (EP) est un concept défini par Tatar et al., 2020 pour compléter le modèle de Garvey, 2009 et l'adapter au cyber. Il s'agit de la quantification de la perte d'opérabilité d'un nœud sans perte chez les parents (par exemple, en raison d'une attaque informatique sur le nœud). Il sera noté $EP_i \in [0; 1]$.

Il est possible qu'un nœud ait plusieurs parents. Il faut donc prendre en compte cette multiplicité. Ainsi, pour V_j avec n parents $w^-(V_j) = \{V_i\}_{1 \leq i \leq n}$, Tatar et al., 2020 et Garvey, 2009 définissent :

$$FDD(V_j) = \text{moyenne}\,(FDD_{1j}, \ldots, FDD_{nj})$$

$$CDD(V_j) = \min(CDD_{1j}, \ldots, CDD_{nj}).$$

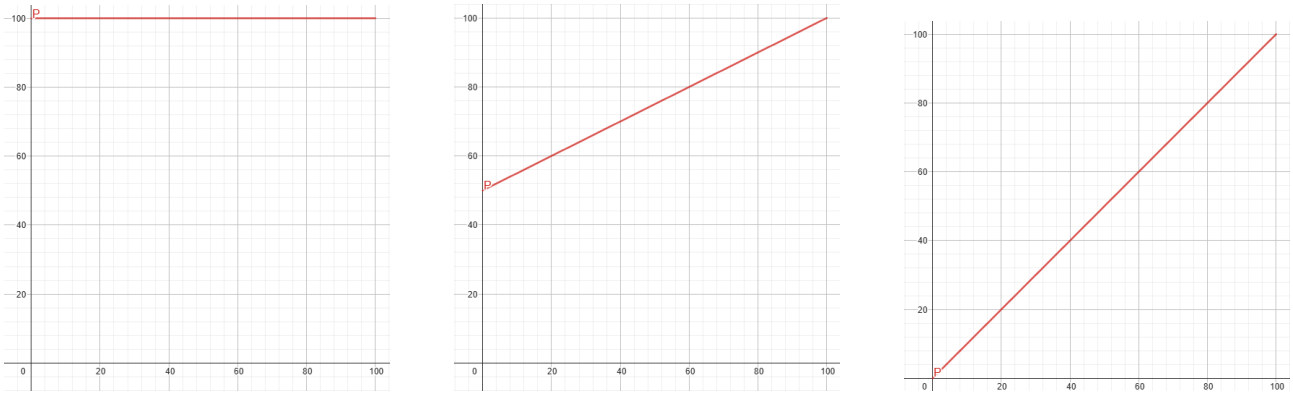
Enfin, ils posent :

$$P_j = EP_j \cdot \min(FDD(V_j), CDD(V_j)). \tag{2.4}$$

Il est ainsi possible de calculer l'opérabilité d'un nœud en fonction de ses parents et de son opérabilité interne.

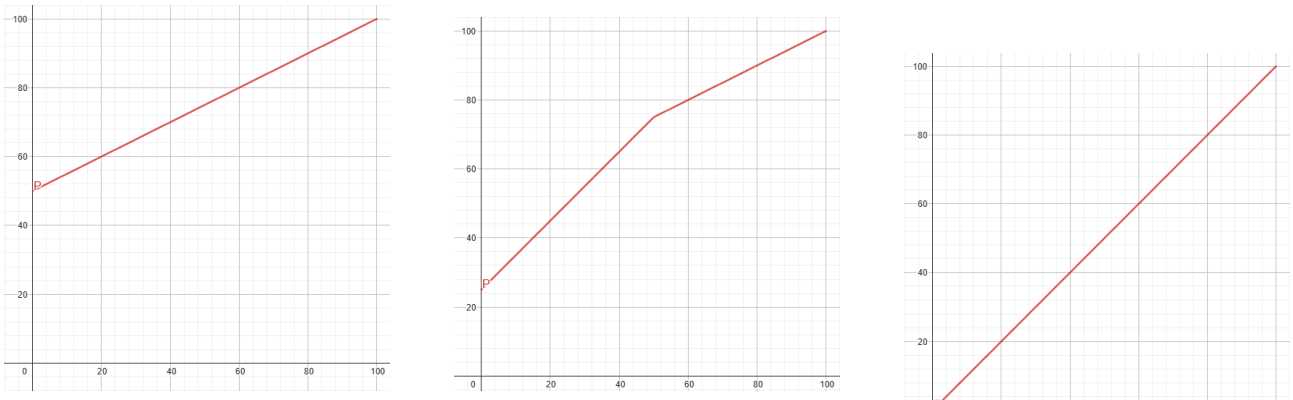
Visualisation des influences de α et β Expliqués ainsi, les concepts de α et β peuvent paraître abstraits. Pour une meilleure compréhension, il est possible de visualiser l'influence de α sur P_j en fixant β à $(1 - \alpha_{ij}) \cdot 100$ (et $EP_j = 1$). La figure 2.20 donne un aperçu de cette influence. Elle présente en abscisse P_i et en ordonnée P_j (avec un seul parent pour des raisons de simplicité).

Sans influence de β, α détermine la pente de la courbe de perte de P_j en fonction de P_i. Si $\alpha = 1$ (sous-figure 2.20c), la perte d'opérabilité est totale lorsque $P_i = 0$, alors que si $\alpha = 0$, aucune perte d'opérabilité n'est visible (sous-figure 2.20a).



(a) Cas limite : $\alpha = 0$ (b) Cas intermédiaire : $\alpha = 0.5$ (c) Cas limite : $\alpha = 1$

Figure 2.20: P_j en fonction de P_i dépendant des valeurs de α



(a) Cas limite : $\beta = (1 - \alpha) \cdot 100$ et $\alpha = 0.5$ (b) Cas intermédiaire : $\beta = 25$ et $\alpha = 0.5$ (c) Cas limite : $\beta = 0$ et $\alpha = 0.5$

Figure 2.21: P_j en fonction de P_i dépendant des valeurs de β pour α fixé

L'influence de β sur le calcul de P_j peut ensuite être visualisée en observant le graphe de P_j pour un α et un EP_j fixé. La figure 2.21 présente différents cas de figure avec $\alpha = 0.5$ et $EP_j = 1$. Les graphiques montrent que plus β est proche de 0, plus la brisure (visible sur la sous-figure 2.21b) intervient tôt dans la perte d'opérabilité de P_i et plus la courbe de perte tend vers un lien de perte total (visible sur la sous-figure 2.21c). Inversement, lorsque β tend vers $(1 - \alpha) \cdot 100$, son influence tend à disparaître. Le lecteur intéressé pourra se rendre sur ce lien Geogebra et tester par lui-même les différents paramètres.

Nœuds Constituants Dans notre graphe d'impact cyber, chaque nœud est classé en trois catégories (voir figure 2.19). Afin de prendre en compte cette caractéristique dans l'analyse FDNA, Tatar et al., 2020 s'inspire de la notion de "nœuds constituants" (Garvey, 2009) et l'adapte au contexte cyber. Ce concept suggère que chaque nœud peut être décomposé en sous-composants, qui sont interconnectés. Par conséquent, l'opérabilité d'un nœud dépendra de celle de ses nœuds constituants.

Les connexions du graphe d'impact ne se feront alors plus entre composants, mais entre sous-composants, comme le montre la figure 2.22.

Comme l'opérabilité générale (P) ne dépend pas de la Confidentialité, de l'Intégrité et de la Disponibilité de la même manière selon deux nœuds différents, la SOA propose d'attribuer un poids

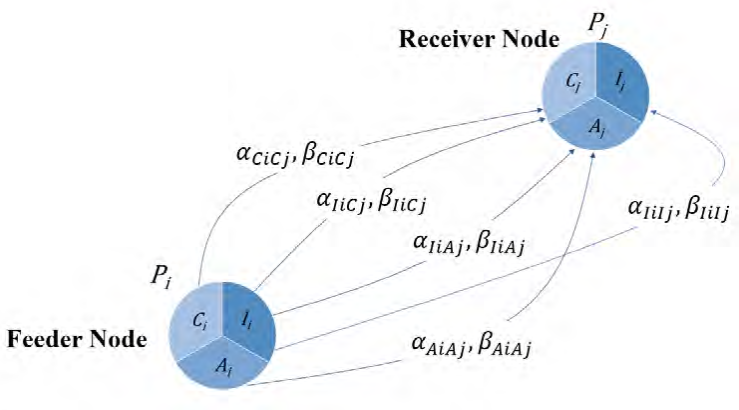


Figure 2.22: Visualisation des liens d'opérabilité entre deux nœuds constituants - Tatar et al., 2020

($w \in [0;1]$) à chacune des composantes C, I et D de manière que

$$P_i = w_{C_i} P_{C_i} + w_{I_i} P_{I_i} + w_{D_i} P_{D_i}.$$

avec la contrainte $w_{C_i} + w_{I_i} + w_{D_i} = 1$ pour chaque nœud dans I_g. Les calculs définis précédemment ne seront donc plus appliqués à P_i directement, mais à chaque nœud constituant, comme présenté sur la figure 2.22 avec les paramètres α et β.

Les portes logiques Il se peut qu'un nœud ayant plusieurs parents dans le graphe ne réponde pas à une logique de "moyenne" entre les différentes FDD, mais plutôt à une logique reposant sur des portes logiques (*AND* et *OR*). Tatar et al., 2020 propose alors de nouvelles manières de prendre en compte ces liens (les portes logiques pouvant être vues comme un nœud intermédiaire, comme sur la figure 2.18).

Logique AND Il est possible qu'un élément de l'entreprise soit dépendant de plusieurs éléments, mais que la perte d'opérabilité d'un seul d'entre eux entraîne une perte pour celui-ci. Dans ce cas, une porte logique AND peut être mise en place entre les n parents et l'enfant. La nouvelle formule pour le calcul de P_{X_j} (avec $X \in \{C, I, D\}$) est

$$P_{X_j} = EP_{X_j} \times \left(\min(\min(CDD_{X_{1j}}, FDD_{X_{1j}}), ..., \min(CDD_{X_{nj}}, FDD_{X_{nj}})\right).$$

L'enfant se comporte comme le "pire" des parents.

Logique OR Il est aussi possible que l'élément subisse une perte d'opérabilité uniquement si l'ensemble des nœuds parents sont affectés. Dans ce cas, une porte logique OR peut être mise en place entre les n parents et l'enfant. La nouvelle formule pour le calcul de P_{X_j} (avec $X \in \{C, I, D\}$) est

$$P_{X_j} = EP_{X_j} \times \left(\max(\min(CDD_{X_{1j}}, FDD_{X_{1j}}), ..., \min(CDD_{X_{nj}}, FDD_{X_{nj}})\right).$$

L'enfant se comporte comme le "meilleur" des parents.

Il est également possible de combiner plusieurs liens logiques.

Quantifier la perte d'efficacité interne EP_X Il est possible, grâce à EP_X (avec $X \in \{C, I, D\}$), de prendre en compte la perte d'opérabilité due à une cyberattaque directement sur un "Actif" (comme défini dans I_g). En effet, une attaque causera une perte sur l'opérabilité d'au moins une des catégories (C, I ou D). Cette perte peut être exprimée par EP_X. Si, durant le processus d'attaque, un actif A_i devient totalement indisponible (comme un serveur lors d'une attaque DDoS), alors il est possible d'établir que $EP_{D_{A_i}} = 0$.

Pour quantifier EP grâce aux données de cybersécurité (CVE), Tatar et al., 2020 propose une approche similaire à celle utilisée pour calculer $p(e)$ (voir section 2.2.2.1). En revenant à la figure 2.17, il est possible de constater que les métriques d'"Impact" ont été laissées de côté. En les examinant de plus près, on observe qu'elles sont divisées en trois catégories correspondant aux trois catégories CID. Ces métriques possèdent également des valeurs numériques, visibles dans le tableau 2.2.

Métrique	**Valeur Métrique**	**Valeur Numérique**	**Valeur Normalisée**
Confidentiality / Integrity / Availability	High	0.56	1
	Low	0.22	0.392
	None	0	0

Table 2.2: Métriques d'impact avec leur valeur numérique et normalisée

Une valeur "*High*" représente, par exemple, une perte totale de la métrique observée par l'utilisation de la faille.

La SOA propose de normaliser ce score (pour obtenir un résultat entre 0 et 1) et de poser, pour un actif A_i touché par une attaque utilisant la faille CVE en question en posant

$$EP_{X_{A_i}} = 1 - \text{Valeur Numérique Normalisée de X.}$$

Estimer le coût Après avoir étudié ce modèle, la question se pose toujours de savoir comment estimer le coût pour l'entreprise, puis dans un second temps le coût pour l'assureur dans le cadre d'une tarification sur un produit.

En suivant la logique du graphe d'impact, la perte financière doit s'observer sur la couche de *business* de l'entreprise, qui est la couche "économique". Le papier de la SOA propose alors d'associer à chaque nœud business un coût pour la perte de C, I ou D (complète) en fonction des différentes pertes liées au cyber (1.1.1.4).

En considérant une attaque cyber qui réduit de moitié l'opérabilité du nœud $B_i \in B$ de l'entreprise pendant d jours, et sachant que le chiffre d'affaires journalier de l'entreprise s'élève à 1000€, le coût de la disponibilité, représentant la perte d'exploitation de l'entreprise, peut être calculé comme

$$Coût_{Disp}(B_i) = 1000 \times d \times 0,5.$$

Ainsi, ce coût reflète l'impact économique de la réduction de l'opérabilité sur la rentabilité de l'entreprise.

D'autres coûts sont présentés dans le papier. Les auteurs exposent l'ensemble des pertes possibles liées au cyber, comme mentionné dans la partie (1.1.1.4). Ils associent ensuite à chacune de ces pertes une composante C, I et D afin d'établir un lien entre le graphe d'impact à la couche business et la perte économique. Pour déterminer les valeurs journalières de ces différentes pertes, lorsqu'une estimation n'est pas possible, le papier recommande de recourir à un avis d'expert.

Mise en commun du graphe d'attaque et du graphe d'impact Il est enfin possible d'associer le graphe d'attaque avec le graphe d'impact en notant que l'ensemble des "actifs" du graphe d'impact correspond également aux nœuds du graphe d'attaque.

La méthode consiste à calculer la probabilité totale de chaque nœud $\mathbb{P}(V_j)$ (chaque actif) et à diminuer son efficacité propre en évaluant la perte en cas de compromission. Par exemple, si l'actif devient totalement inopérable sur la confidentialité, la perte est de 100 s'il est capturé par l'attaquant. Cette perte est ensuite multipliée par la probabilité totale que l'actif soit compromis, selon la méthodologie décrite dans la section (2.2.2.1). Supposons que la probabilité de compromission de l'actif est de $\mathbb{P}(V_j) = 0,27$, alors la perte d'opérabilité en termes de confidentialité sera de $100 \times 0,27 = 27$, et l'efficacité restante sera $EP_{I_j} = 100 - 27 = 73$.

Une fois l'ensemble des EP_X des cibles du graphe d'attaque calculé, la perte économique peut être obtenue en appliquant le graphe d'impact et la méthodologie FDNA.

Apport et limites L'apport principal de cette modélisation est qu'elle permet de quantifier la perte à une échelle micro à partir de la modélisation des probabilités d'attaques sur chaque actif de l'entreprise (via le graphe d'attaque). Elle représente ainsi une quantification du coût sans nécessité d'historique de sinistres.

L'estimation des coûts liés à la perte totale d'opérabilité des nœuds business reste tout de même approximative dans Tatar et al., 2020, le papier proposant un avis d'expert pour estimer certains de ces coûts. Néanmoins, les coûts liés aux pertes d'exploitation peuvent être considérés comme des pertes de disponibilité et sont relativement faciles à estimer grâce au chiffre d'affaires de l'entreprise.

Enfin, l'objectif du papier est de proposer des méthodes sans réelle application directe à une police d'assurance, il n'est donc pas exhaustif. La suite de ce mémoire modifiera et appliquera la méthodologie et les concepts techniques (graphe d'attaque, FDNA) à une assurance **perte d'exploitation** cyber.

2.3 Développement du Modèle Dynamique pour la Perte d'Exploitation

Dans la section précédente, une méthodologie permettant d'appliquer des modèles graphiques à une quantification cyber dynamique a été étudiée en détail.
Cette vision approfondie, qui prend en compte la méthodologie réelle sous-jacente au "risque cyber" (des attaques sur des cibles intermédiaires jusqu'à l'atteinte de la cible principale), nécessite néanmoins quelques ajustements.

Dans un premier temps, l'estimation unique du coût, sans distribution, ne permet pas de bien cerner le risque, car peu d'informations sont disponibles sur la répartition du risque. De plus, le framework présenté dans le papier ne détaille pas suffisamment les prémices de l'attaque (lorsque l'attaquant n'a pas encore choisi d'attaquer l'entreprise) et demeure trop général concernant les pertes subies et l'application du modèle en général. En effet, une méthodologie unique est appliquée à la fois aux pertes d'exploitation et aux pertes juridiques.

Cette partie s'intéressera donc à la mise en place d'un modèle dans le cadre de la tarification d'une assurance **perte d'exploitation**. En effet, selon Cambridge Centre for Risk Studies, 2016, la garantie pertes d'exploitation est l'une des plus courantes, étant incluse dans 69% des produits. S'intéresser à cette garantie en particulier permet ainsi de couvrir un grand nombre de polices. Le modèle se basera sur les préceptes techniques étudiés dans la partie précédente tout en permettant d'obtenir une distribution du risque pour l'entreprise (donc à l'échelle de l'"assuré"). Des modifications

y seront apportées et justifiées. L'objectif de cette partie n'est pas de fournir un modèle parfait, mais bien de démontrer qu'il est possible d'appliquer l'ensemble des connaissances acquises dans un cadre assurantiel concret. Certaines zones d'ombre seront identifiées et donneront lieu à des perspectives de recherches futures.

Dans un premier temps, cette section rappellera les aspects de la perte d'exploitation dans le contexte cyber, tout en présentant comment adapter un modèle graphique, tel que celui vu précédemment, à ce type de pertes. Dans un second temps, le modèle sera présenté dans son ensemble. Enfin, deux modules du modèle seront étudiés en détail.

2.3.1 Perte d'exploitation - Contexte du modèle

Nos recherches se concentrent sur la perte d'exploitation dans le contexte cyber et sur l'adaptation des modélisations présentées dans la partie précédente à ce cadre. Cette section aborde ainsi certains aspects de la perte d'exploitation qu'il sera nécessaire de prendre en compte dans le modèle.

Dans un premier temps, la perte d'exploitation sera redéfinie. Ensuite, certaines considérations liées à la création d'un modèle de quantification individuelle (entreprise par entreprise) dans le contexte de la perte d'exploitation seront discutées.

2.3.1.1 La perte d'exploitation cyber

La perte d'exploitation a été brièvement étudiée dans le Chapitre 1. Dans la section (1.1.1.4), elle est définie comme une perte directe de la productivité de l'entreprise due à une attaque cyber.

En effet, lorsqu'une attaque compromet la disponibilité (D dans les métriques CID (1.1.1)) de certaines cibles de l'entreprise, celle-ci n'a plus la capacité de fournir certains services à leur plein potentiel. Pendant toute la durée de l'attaque ainsi que la période de remédiation (période durant laquelle l'entreprise répare ses systèmes endommagés), son fonctionnement est réduit. Le graphe FDNA peut permettre de quantifier cette perte à une échelle journalière, comme cela sera présenté dans la partie dédiée à la remédiation du modèle.

La garantie perte d'exploitation protège alors l'entreprise contre ce risque (avec la possible existence de limites et de franchises) en lui remboursant la perte subie. Nous souhaitons construire un modèle de tarification individuelle suivant les préceptes établis (dynamisme, absence de nécessité d'une base de sinistres, possibilité de suivi et de conseil des assurés) et utilisant les modélisations abordées dans la partie précédente. Il est donc nécessaire d'adapter le modèle graphique à la perte d'exploitation. Pour ce faire, certains aspects de cette perte doivent être pris en compte.

2.3.1.2 Adapter un modèle graphique à la perte d'exploitation

La perte d'exploitation est une perte particulière dans le sens où son coût total augmente avec le temps τ d'incapacité de l'entreprise. Le modèle doit donc prendre en compte cette période post-attaque, durant laquelle l'entreprise ne fonctionne que partiellement et s'améliore progressivement jour après jour jusqu'à retrouver son opérabilité initiale. Cela ajoute une dimension temporelle à la situation post-attaque de l'entreprise.

Il ne s'agit pas d'un état binaire (malade puis sain), mais d'une transition progressive : chaque fois qu'un système est réparé, l'opérabilité de l'entreprise est estimée comme étant améliorée.

Cette notion n'est pas abordée dans l'étude de Tatar et al., 2020. Il est donc nécessaire de l'intégrer dans notre modèle en tenant compte d'une période de remédiation, durant laquelle l'entreprise ne

fonctionne pas à plein régime, mais réduit puis augmente progressivement sa production jour après jour.

La section suivante présentera un modèle prenant en compte cette particularité en l'intégrant dans une vision du risque *micro* (graphe d'attaque et FDNA).

2.3.2 Présentation du modèle

L'idée portée par ce modèle est d'intégrer le papier de Tatar et al., 2020 au contexte de la tarification individuelle d'une police d'assurance perte d'exploitation. Le modèle modifie l'approche établie dans le document en sectionnant le mécanisme d'attaque en trois étapes : volonté d'attaque, processus d'attaque et remédiation, afin d'estimer correctement le risque. Il modifie également la vision de l'attaque en réalisant des simulations du processus d'attaque et de la remédiation, permettant ainsi d'obtenir non seulement une valeur moyenne, mais aussi une répartition du risque pour l'entreprise, ce qui constituera une quatrième étape du modèle.

Les domaines d'applicabilité de ce type de modèle seront discutés dans le Chapitre 3 de ce mémoire, où plusieurs cas pratiques seront menés sur un portefeuille fictif.

Le modèle est conçu pour "suivre" un attaquant ciblant une entreprise, avec des étapes similaires à celles de la matrice ATT&CK. En adoptant cette méthodologie, l'objectif est de mieux estimer le risque en se rapprochant au maximum du réalisme des mécanismes des attaques cyber.

Dans un premier temps, une vision d'ensemble du modèle, accompagnée d'un schéma récapitulant son architecture, sera présentée. Chaque partie du modèle sera ensuite discutée plus en détail afin d'en comprendre le fonctionnement.

2.3.2.1 Une vision d'ensemble

Dans cette section, l'architecture globale du modèle mis en place sera présentée. La figure 2.23 décrit schématiquement les étapes du modèle permettant d'arriver au tarif de la prime pure individuelle pour un assuré. La figure présente ainsi un modèle en quatre étapes clés.

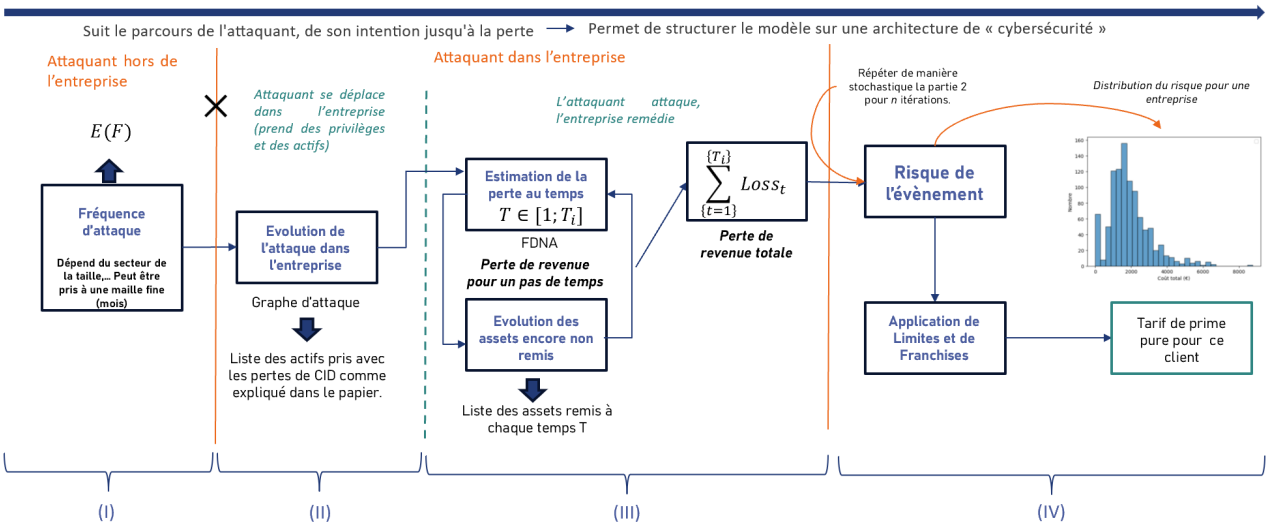


Figure 2.23: Architecture du modèle - Un modèle graphique à simulations

L'idée du modèle est de pouvoir suivre l'attaquant dans sa tentative d'attaque, depuis sa volonté

d'attaquer l'entreprise i de notre portefeuille jusqu'à sa réussite ou son échec. L'ensemble des connaissances développées dans les sections précédentes de ce chapitre permet d'appliquer des outils au niveau de détail des actifs et des failles, tels que le graphe d'attaque probabilisé et le FDNA. Cette vue **micro** est d'autant plus intéressante que la majorité des connaissances en cybersécurité se trouvent à ce niveau de détail.

Le modèle est donc divisé ainsi :

(I) L'attaquant se trouve à l'extérieur de l'entreprise et décide d'attaquer l'assuré i. Afin d'étudier le risque, il est nécessaire d'obtenir la fréquence (ou la loi de fréquence) d'un tel événement. Ainsi, la variable aléatoire représentant cette fréquence sera notée F. L'espérance de la **fréquence d'attaque** sera notée $\mathbb{E}(F)$.

La modélisation de cette partie est hors de notre domaine d'étude ; nous nous intéresserons donc peu à l'estimation précise de F, bien que cela puisse faire l'objet d'une étude complémentaire.

Néanmoins, il a été présenté au Chapitre 1, section 1.1.1.5, que le risque d'attaque dépend de la taille et du secteur de l'entreprise en question.

Il est important de noter que cette partie détermine également la granularité de notre tarification. Les parties suivantes, qui estiment la réussite de l'attaquant sous condition d'attaque, sont indépendantes de l'exposition de l'entreprise.

D'un point de vue assurantiel, afin d'obtenir une prime dynamique évoluant en fonction des périls de l'année, les contrats ne devraient plus être annuels, mais plutôt d'une granularité plus fine (par exemple, mensuelle). Dans ce cas, $\mathbb{E}(F)$ représente la fréquence d'attaque moyenne d'un assuré sur un mois. Il est possible d'envisager des études plus approfondies sur la dépendance de la fréquence d'attaque en fonction d'autres critères cyber, tels que l'augmentation des attaques dans d'autres entreprises du même secteur ou encore la présence de virus en vente sur le dark web.

Cette section du modèle est indépendante du reste et peut être considérée à part. $\mathbb{E}(F)$ sera multipliée par le résultat de la perte obtenue par la suite afin d'obtenir une quantification complète du risque, selon une approche fréquence x coût des attaques. Ne pas prendre en compte cette première étape reviendrait à commettre l'erreur de supposer qu'une entreprise subit une attaque par mois (ou par an, selon la granularité choisie).

(II) Une fois la volonté de l'attaquant prise en compte, il est possible de représenter l'attaque dans son aspect probabiliste. C'est à ce stade que le modèle fait appel à des simulations. Chaque simulation est considérée comme une attaque indépendante de l'entreprise par l'attaquant.

La partie (II) correspond à la quantification de l'évolution de l'attaquant dans l'entreprise. Le Chapitre 1 présentait une méthodologie d'attaque selon laquelle l'attaquant se déplace dans le réseau de l'entreprise afin d'accéder au maximum de cibles et de privilèges. Le Chapitre 2 introduisait un graphe d'attaque décrivant l'ensemble des chemins empruntables par l'attaquant ainsi qu'une structure de probabilité.

L'objectif ici est que, grâce à une méthodologie d'attaque développée plus loin dans cette section (2.3.2.2), l'attaquant obtienne une liste de cibles attaquées avec succès L_c, en exploitant les vulnérabilités du graphe d'attaque. Cette liste dépend des probabilités de passage d'un nœud au suivant, conférant ainsi un aspect probabiliste à l'attaque. Par exemple, si l'attaquant a 90% de chances d'exploiter une vulnérabilité directement depuis la surface d'attaque, l'actif correspondant apparaîtra dans la liste dans environ 90% des observations. Chaque élément de L_c est ensuite diminué de son opérabilité interne (C, I ou D), suivant la méthodologie de Tatar

et al., 2020. Ce point sera détaillé ultérieurement, car la probabilité $p(e)$ n'est plus fixe entre les simulations, mais devient une observation d'une loi.

L'intérêt de cette section réside dans la possibilité de calculer un grand nombre d'attaques en fonction de la topologie du graphe d'attaque et de la structure de probabilité induite par le réseau bayésien. De manière informelle, à chaque simulation, l'attaquant démarre en dehors de l'entreprise et tente de s'infiltrer en suivant le graphe d'attaque.

(III) Une fois que l'attaquant détient l'ensemble des cibles qu'il a pu capturer, il lance l'attaque. Cela suit la logique introduite par MITRE dans la figure 1.8, où l'attaquant capture un maximum d'accès et de machines ("Accès et élévation") avant de lancer l'attaque ("Exfiltration et attaque").

Il est donc nécessaire d'estimer le coût pour l'entreprise, d'abord à l'échelle journalière, puis à l'échelle globale. Cela est représenté par la partie (III). Comme expliqué dans la section précédente (2.3.1.2), la perte d'exploitation s'étale sur plusieurs unités de temps. Le modèle considère donc d'abord le coût lié à la diminution d'opérabilité observée pour chaque actif de L_c, en utilisant le graphe d'influence et la méthodologie FDNA.

Ensuite, une simulation représentant la **remédiation** de l'entreprise est effectuée. On considère qu'à chaque instant T, un actif capturé peut se remettre de l'attaque (être réparé), et la perte journalière P_T est recalculée en fonction de la liste des actifs restants L_c (sans compter les actifs remis), jusqu'à ce que la liste soit vide.

La perte totale de production pour cette simulation est alors calculée en prenant

$$C_i = \sum_{T=1}^{T_i} P_T.$$

où T_i est le premier instant où l'ensemble des actifs sont remis en état, une variable aléatoire qui sera discutée plus en détail dans la section 2.3.2.3.

(IV) Le modèle effectue n simulations d'attaque/remédiation. Grâce à cela, il est capable de fournir une répartition du risque pour l'entreprise.

Après application des limites et des franchises, il est possible de tarifier la prime pure de l'entreprise, en prenant en compte $\mathbb{E}(F)$.

La figure 2.23 donne une représentation visuelle de l'architecture du modèle. Le graphique de distribution du risque est une véritable sortie du modèle pour une entreprise test.

Dans les sous-parties suivantes, la prise de cible et la remédiation seront présentées en détail (parties II et III). Une discussion portera sur l'application de Tatar et al., 2020, les choix de modifications réalisés dans le contexte des simulations ainsi que les zones non explorées.

2.3.2.2 La prise de cibles

La section précédente présentait l'architecture globale du modèle, sans s'attarder sur les détails techniques de chacune des parties. Cette section étudiera la partie II. La méthodologie et l'adaptation du papier seront présentées. Nous verrons comment le réseau bayésien d'attaque a été adapté pour obtenir une liste de cibles et comment $p(e)$ a été modifié par rapport à la définition de Tatar et al., 2020. Les raisons de ces changements ainsi que la cohérence des résultats par rapport au modèle de base (et aux bases techniques) seront discutées.

Discussions préliminaires Le modèle de Tatar et al., 2020 utilise les graphes bayésiens d'attaque, comme présenté en section 2.2.2.1, pour quantifier la probabilité qu'un actif soit pris par l'attaquant lors d'une attaque. Pour estimer le coût, discuté en section 2.2.2.2, ils considèrent la diminution de l'opérabilité de l'actif en fonction de l'impact de la faille sur celui-ci et de la probabilité que cela arrive, en multipliant les deux. Ils appliquent ensuite le graphe d'impact aux actifs diminués en efficacité propre et obtiennent une perte économique moyenne unique.

Le modèle proposé dans ce mémoire vise à permettre une distribution du risque pour un assuré. Pour ce faire, il fera appel à de multiples simulations. Au lieu d'avoir une unique liste d'actifs avec les probabilités globales, qui peuvent être inférées grâce au réseau bayésien ou grâce à la méthodologie du papier, comme discuté en section 2.2.2.1, l'idée est de simuler à chaque itération une nouvelle liste suivant les probabilités locales calculées dans le réseau bayésien d'attaque. Le choix de faire appel à un réseau bayésien pour l'inférence et le calcul des probabilités locales ayant déjà été discuté lors de la présentation des méthodes de calcul, il ne sera pas rediscuté ici.

Le paragraphe suivant détaillera, étape par étape, comment créer une simulation pour obtenir cette liste L_c.

Création d'une simulation et obtention de la liste L_c. Il est considéré dans cette simulation que l'attaquant s'en prend à notre assuré i (la fréquence d'attaque moyenne étant calculée en partie I avec $\mathbb{E}(F)$). Voici, étape par étape, comment se déroule une simulation :

1. **Obtention du graphe d'attaque** Comme dans le papier de Tatar et al., 2020, la base du calcul est le graphe d'attaque.
   Le graphe d'attaque d'une entreprise du portefeuille doit être considéré comme une donnée d'entrée du modèle. La possibilité d'obtenir ces graphes de manière automatique a été discutée en partie 2.2.2.1, plusieurs outils le permettent, ce qui donne la possibilité de l'obtenir pour chaque assuré de manière périodique. Ce graphe, pour un assuré, sera nommé G. La figure 2.24 donne un exemple de ce type de graphe pour une très petite entreprise. Les A_i représentent les actifs et $X\_$ représente l'attaquant aux portes de l'entreprise.

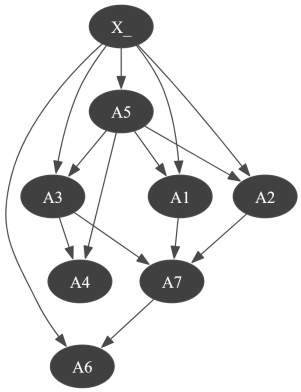


Figure 2.24: Exemple de graphe d'attaque pour une petite entreprise

   La surface d'attaque au temps 0 (avant l'attaque) est $S = (A_5, A_3, A_1, A_2, A_6)$.

2. **Calcul des $p(e_i)$** Tatar et al., 2020 présentait un calcul unique pour la probabilité d'exploitation réussie d'une faille (détaillé dans la section 2.2.2.1).
   La prise en compte des métriques de probabilité du score CVSS présente une corrélation avec la probabilité de succès. Néanmoins, elles ne représentent pas la probabilité exacte. Le choix d'une probabilité unique donnant un unique résultat peut être jugé trop peu détaillé.

Pour pallier ce problème, le modèle développé considère que $p(e_i)$ est une variable aléatoire suivant une loi $\mathcal{L}$. À chaque simulation, le graphe d'attaque sera pondéré par de nouvelles observations de $p(e_i)$. Toutefois, afin de conserver la logique de la quantification du papier de Tatar et al., 2020 et l'utilisation du score CVSS, qui est corrélé à cette probabilité, il sera considéré que la probabilité de chaque arête, telle que calculée dans le papier (désormais notée $\overline{p(e_i)}$), sera liée à un paramètre de la loi.
Cette approche a été discutée lors d'une réunion avec M. Tatar.

Dans ce mémoire, il sera considéré que $\mathcal{L}$ suit une loi normale $\mathcal{N}(\mu, \sigma)$. Le calcul de μ et σ sera réalisé à partir des paramètres suivants :

(a) La probabilité que $p(e_i)$ soit hors de l'intervalle $[0; 1]$ doit être négligeable ($\leq \lambda$). Dans ce mémoire, $\lambda = 0.1\%$. Étant donné que la loi normale a un domaine de définition sur $\mathbb{R}$ et que $p(e_i)$ doit être compris entre $[0; 1]$, cette contrainte doit être introduite.

(b) Par souci de prudence, nous souhaitons que $\overline{p(e_i)}$ ne soit pas la moyenne (car dans ce cas, il y aurait autant de fois où la probabilité de réussite est en dessous qu'au-dessus), mais qu'elle soit un paramètre limite. Ainsi, nous imposons que $\mathbb{P}(p(e_i) \leq \overline{p(e_i)}) = \nu$, où ν est un paramètre de prudence. Des recherches complémentaires sur ν pourraient être menées, par exemple en tenant compte de l'évolution du score CVSS de la faille e_i dans le temps. Dans le contexte de ce mémoire, ν a été fixé à 0.3, ce qui signifie que la probabilité qu'une observation de $p(e_i)$ soit inférieure à $\overline{p(e_i)}$ est de 0.3. La figure 2.25 présente un exemple du paramétrage d'une loi pour $p(e_i)$ concernant une faille CVE (e_i) dont le score CVSS de probabilité est de 0.35.

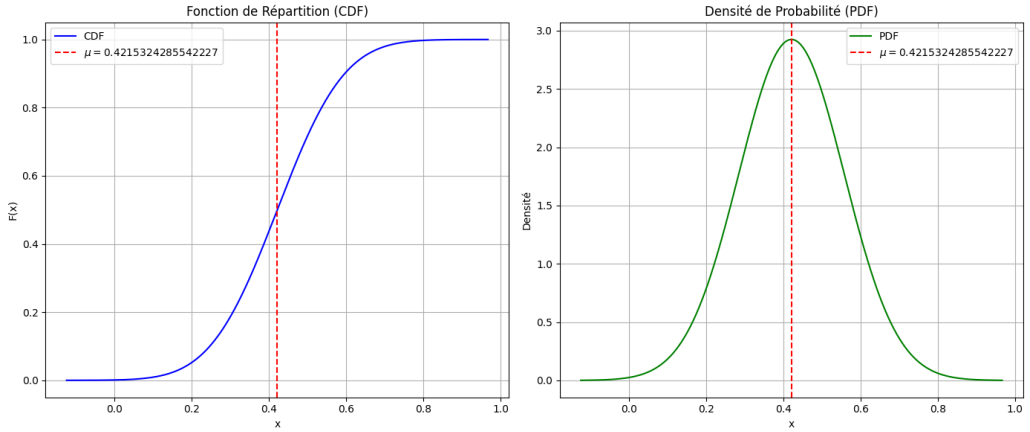


Figure 2.25: Fonction de répartition et densité de probabilité de $p(e_i)$, avec $\lambda = 0.1\%$ et $\nu = 30\%$

Le fait que $\overline{p(e_i)} \leq \mu$ est rassurant en termes de prudence du modèle. ν est un paramètre à garder en tête pour de possibles tests de sensibilité.

Le choix de cette loi n'est pas unique. Des études plus poussées sur le type de loi s'adaptant le mieux à ce cas de figure doivent être menées pour plus de précision (voir la sous-section 3.3.1.3 pour plus de détails). La question de la loi est importante, mais elle sort du cadre de ce mémoire, tout comme celle de l'étude du paramètre ν. L'objectif n'est pas de fournir un modèle finalisé, mais d'avancer des pistes de modélisation pour une tarification assurantielle dans un contexte de manque de données de sinistre.

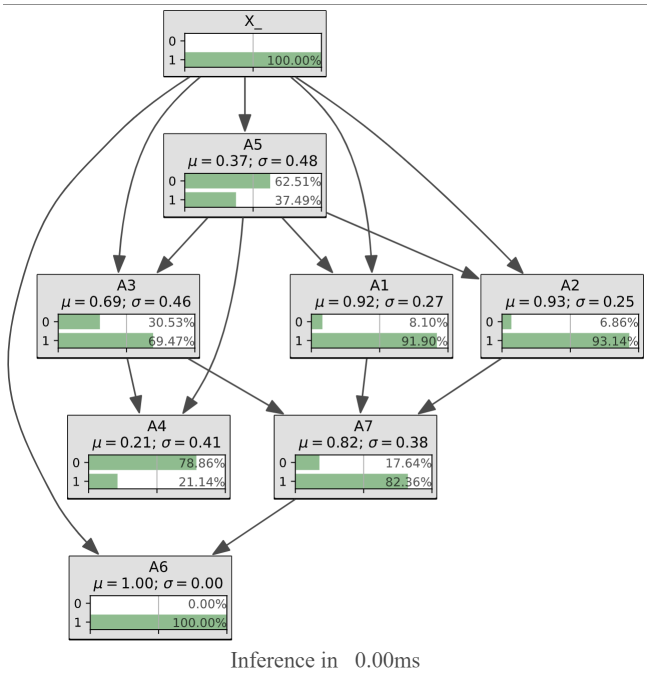
3. **Création du réseau bayésien d'attaque** Il suffit d'appliquer la méthodologie de Poolsappasit et al., 2012, vue dans la partie 2.2.2.1, afin d'obtenir les probabilités locales pour le graphe.

Dans ce mémoire, seule la logique OR sera utilisée pour des raisons de simplicité. Néanmoins, il est possible et raisonnable d'adapter le modèle à une logique AND ou même à une logique intermédiaire. Le choix dépend en réalité du graphe d'attaque en entrée du modèle et des besoins de la tarification.

Un exemple de probabilité conditionnelle est illustré dans la figure 2.26a, qui représente une simulation pour l'actif A_7 de notre exemple précédent. Il peut être observé que, pour cet exemple, $\mathbb{P}(A_7 = 1 \mid A_1 = 0, A_2 = 1, A_3 = 0) = 0.58$.

			A7	
A3	A2	A1	0	1
0	0	0	1.0000	0.0000
		1	0.4733	0.5267
	1	0	0.4214	0.5786
		1	0.1995	0.8005
1	0	0	0.6255	0.3745
		1	0.2960	0.7040
	1	0	0.2636	0.7364
		1	0.1248	0.8752

(a) Probabilités conditionnelles pour l'actif A_7



(b) Inférence sur le graphe d'attaque de la figure 2.24

Figure 2.26: Probabilités locales et inférence sur le graphe d'attaque

Grâce à ce graphe, il est possible d'obtenir les probabilités globales de prise de chaque actif par inférence. La figure 2.26b présente un exemple d'inférence du réseau bayésien d'attaque basé sur l'exemple de la figure 2.24. Dans cet exemple, il est possible de voir que la probabilité que l'attaquant obtienne A_4 lors d'une attaque est de 0.21.

4. **Obtenir la liste des actifs capturés pour cette simulation** Chaque simulation est une attaque indépendante sur l'entreprise i, chaque attaque n'ayant pas forcément la même topologie (l'attaquant pouvant réussir là où, dans une autre simulation, il avait échoué, et inversement). Chaque simulation j produit donc une liste d'actifs capturés par l'attaquant différente, notée L_{c_j}. Pour obtenir cette liste, le modèle suit la topologie du graphe d'attaque. Il part de la racine (la position de l'attaquant X_-), puis prend la liste des enfants (dans notre exemple $w^+(X_-) = (A_5, A_3, A_1, A_2, A_6)$).

Pour chaque enfant, la probabilité conditionnelle est connue grâce au réseau bayésien et, en considérant la liste des actifs déjà capturés, il est possible d'estimer la probabilité de capture. Le processus continue en suivant les enfants des enfants jusqu'à ce que l'ensemble du graphe soit parcouru (il arrive parfois que l'attaquant soit "bloqué" et que la probabilité de continuer jusqu'à l'actif final soit nulle). Cette méthodologie est conçue pour fournir, en moyenne, des résultats similaires à l'inférence bayésienne afin d'assurer la cohérence du modèle. De plus, elle

permet d'introduire une logique réaliste dans la méthodologie d'attaque, en retraçant les étapes suivies par l'attaquant sur le graphe à chaque simulation.

Une fois la liste obtenue, les effets sur chaque EP_X de chaque actif capturé sont calculés selon la méthodologie développée en 2.2.2.2, avec

$$EP_{X_{A_i}} = 1 - \textit{Valeur Numérique Normalisée de } X.$$

La perte est, elle aussi, transformée en le résultat d'une variable aléatoire, de manière similaire à $p(e_i)$ et pour les mêmes raisons.

Après avoir obtenu une liste d'actifs capturés (L_{c_j}) pour la simulation j et les valeurs EP_X pour chacun d'entre eux, il est nécessaire d'estimer la perte d'exploitation correspondante. Le paragraphe suivant discutera de la **remédiation**, c'est-à-dire de l'étape III, au cours de laquelle les actifs sont restaurés et le coût total de la perte d'exploitation pour cette attaque est quantifié.

2.3.2.3 La remédiation

Une fois l'attaque opérée, les systèmes altérés par l'attaquant et les *business* de l'entreprise diminués, il est nécessaire de quantifier la perte économique liée à la **perte d'exploitation**, qui constitue notre centre d'intérêt.

Le modèle présenté dans ce mémoire reprend le graphe FDNA, fonctionnant avec les nœuds constituants C, I et D, comme expliqué dans la section 2.2.2.2.
Comme présenté dans Tatar et al., 2020, la perte d'activité est uniquement liée à la composante de Disponibilité des nœuds de business. Par conséquent, seule cette composante sera étudiée dans le cadre du modèle proposé dans ce mémoire.

Il est considéré que le chiffre d'affaires de l'entreprise est réparti entre chaque nœud de business de l'entreprise (B_i). Cette répartition peut être égale ou non, en fonction des spécificités de l'entreprise. Ainsi, un nœud de business détient un chiffre d'affaires journalier $CA_{jour}(B_i)$ qui est égal à x lorsque l'opérabilité du nœud est à 100%. Ce chiffre d'affaires diminue proportionnellement à l'opérabilité du nœud constituant de Disponibilité de l'entreprise. Par exemple, si $P_{D_{B_i}} = 0.5$, alors le chiffre d'affaires durant cette période de baisse d'opérabilité est de $0.5 \cdot x$.

Le paragraphe suivant expliquera comment la perte d'exploitation est calculée en utilisant cette notion de chiffre d'affaires journalier.

Méthodologie de calcul de perte d'exploitation pour notre modèle Le modèle FDNA permet d'évaluer la perte d'opérabilité des nœuds de business à partir de la liste des actifs capturés lors de la simulation j (notée L_{c_j}), ainsi que des pertes d'efficacité propre (EP) calculées pour chacun de ces actifs. L'EP, discutée dans le paragraphe "FDNA - quantifier les dépendances" (2.2.2.2), constitue un élément clé de cette estimation. Il reste alors à estimer la durée d'incapacité de l'entreprise τ et l'évolution de la perte journalière durant cette période (celle-ci pourrait être constante puis s'arrêter à τ, auquel cas seul τ serait à étudier).

Dans ce mémoire, une modélisation similaire à celle des modèles SIR sera considérée, comme détaillée dans le mémoire de Peyrat, 2022, où la liste d'infectés est connue à $t = 0$ et correspond à L_{c_j}.

Il sera ainsi supposé qu'à chaque instant t (jour), chaque actif a une probabilité ρ de passer d'infecté à remis. Cette probabilité est considérée identique pour chaque actif, et chaque actif infecté est indépendant dans sa remédiation.

En d'autres termes, le temps de remédiation de chaque actif (τ_i) suit une loi géométrique $\mathcal{G}(\rho)$, donc $\tau \sim \max(\tau_1, ..., \tau_k)$ pour L_c contenant k actifs. Par indépendance, cela donne

$$\mathbb{P}(\tau \leq u) = \mathbb{P}(\max(\tau_1, ..., \tau_k) \leq u) = \mathbb{P}(\cap_{i=1}^{k} \tau_i \leq u)$$

$$\mathbb{P}(\cap_{i=1}^{k} \tau_i \leq u) = \prod_{i=1}^{k} \mathbb{P}(\tau_i \leq u) = \left(\sum_{j=0}^{u} ((1-\rho)^{j-1}\rho) \right)^k = (1-(1-\rho)^u)^k \,.$$

Bien qu'il soit possible de calculer la loi de τ, les simulations seront privilégiées afin de retirer progressivement des actifs de la liste L_c et ainsi modéliser l'évolution de la perte dans le temps.

L'objectif est de simuler une entreprise réparant progressivement ses actifs et retrouvant son exploitation après une attaque cyber. Toutefois, ρ est un paramètre nécessitant une étude approfondie, et la méthodologie doit faire l'objet de tests et de modifications dans des conditions réelles. Des recherches supplémentaires doivent être menées dans un contexte autre que celui de ce mémoire.

L'algorithme 1 présente la méthode de calcul de la perte d'exploitation totale en suivant cette approche.

Algorithm 1 Calcul de la perte de production

$P \leftarrow 0$
Tant que L_c n'est pas vide **faire**
    FDNA(L_c)
    Pour chaque b **dans** B **faire**
        $P \leftarrow P + \text{CA}(b) - \text{CA}(b) \times P\_D(b)$
    Fin Pour
    Pour chaque a **dans** L_c **faire**
        Si uniform$(0,1) < \rho$ **alors**
            Enlever a de L_c
        Fin Si
    Fin Pour
Fin Tant que

L_c est la liste des actifs, B la liste des nœuds de *business*, FDNA(.) est la fonction de calcul des FDNA en fonction des actifs infectés (et de leurs EP respectifs), et CA(b) correspond au chiffre d'affaires journalier du *business* b.
La perte pour b est alors la valeur de CA(b) diminuée de CA$(b) \cdot P_{D_b}$, correspondant à la valeur journalière réduite par l'attaque du *business* b.

Pour finaliser le modèle, il suffit d'effectuer les étapes (II) et (III) n fois afin d'obtenir une distribution des pertes de l'entreprise. La moyenne empirique de cette distribution peut être calculée, servant d'estimation de l'espérance via la simulation de Monte-Carlo.

Dans ce chapitre, nous avons présenté un cadre théorique permettant une tarification en assurance cyber malgré les nombreux défis du domaine, en particulier le manque de données de sinistre. Le choix des modèles graphiques a été privilégié en raison de leur versatilité et de leur adaptabilité aux deux mondes, cyber et assurantiel. Pour ce faire, un cadre théorique a été établi, intégrant des connaissances issues de la cybersécurité (graphes d'attaque, CVSS, etc.) dans un modèle de tarification assurantielle (section 2.3).

Compte tenu de l'étendue des recherches encore nécessaires dans ce domaine, nous nous sommes concentrés sur une architecture permettant la création d'un modèle intégrant ces connaissances dans le contexte de la garantie de perte d'exploitation. De nombreux points restent à explorer, néanmoins, l'objectif est de fournir une voie d'évolution pour de futurs travaux.

Chapter 3

Application à la perte d'exploitation pour PME

L'objectif de ce chapitre est de discuter d'un cadre applicatif au modèle construit. Ce cadre vise à schématiser l'étendue des possibilités assurantielles offertes par l'utilisation du modèle tel que conçu dans la section 2.3. De la quantification à différents *stress tests*, nous étudierons plusieurs cas d'application pouvant apporter une plus-value à l'assureur, tant dans la connaissance du risque que dans l'ouverture à de nouveaux marchés, notamment auprès des PME recherchant plus qu'une simple assurance.

Les données cruciales à l'utilisation du modèle (graphe d'attaque, graphe d'impact et probabilité d'attaque) n'étant pour l'instant pas disponibles à l'échelle d'un portefeuille, il sera question dans les prochaines parties de la création d'un **portefeuille fictif**, qui sera rendu aussi réaliste que possible. La récupération de ces données dans un cadre réaliste, par exemple lors de la souscription, constituera également un point de discussion important.

Ce chapitre sera placé dans le contexte d'un assureur cyber proposant un contrat d'assurance fournissant une garantie perte d'exploitation à des **PME**. Le choix de l'assureur de se spécialiser dans ce secteur, ainsi que l'intérêt qu'auraient les PME à souscrire ce type de contrat, seront détaillés tout au long de ce chapitre à travers la présentation des opportunités offertes par une quantification dynamique et micro (à l'échelle des vulnérabilités) du risque.

Dans une première partie, la méthodologie de création d'un portefeuille assuré fictif sera abordée. Un parallèle avec la faisabilité de la récupération des graphes d'attaque et d'impact dans un cadre réel sera établi. Les raisons du choix de spécialisation pour les PME seront également approfondies. Dans une deuxième partie, la quantification dynamique dans le cadre du portefeuille assuré sera examinée. Nous présenterons l'application du modèle au portefeuille ainsi que sa capacité à faire évoluer la prime en fonction du risque du moment. L'échelle temporelle d'un contrat cyber sera évoquée, et les bénéfices de l'implication de l'assureur dans la prévention seront discutés. Des stress tests cyber seront également réalisés à l'aide du modèle, offrant ainsi d'autres perspectives d'utilisation en dehors de la tarification. Enfin, dans une dernière partie, une discussion sur les prochaines étapes du développement et de l'adaptation de ce type de modèles sera menée.

3.1 Création d'un portefeuille assuré

Cette section étudie la création d'un portefeuille assuré permettant l'utilisation du modèle développé dans la partie 2.3. Pour que le modèle fonctionne correctement pour une entreprise assurée, il est nécessaire de fournir en entrée :

1. Le graphe d'attaque composé de l'ensemble des actifs A de l'entreprise, ainsi que des liaisons par des vulnérabilités pouvant évoluer dans le temps, comme discuté dans la section 3.2.2.

2. Le graphe d'impact avec sa pondération FDNA, défini par $I_g = (E, V, Val_\alpha, Val_\beta)$.

3. Étant donné le cadre de l'assurance perte d'exploitation, il est nécessaire d'obtenir un chiffre d'affaires pour chaque nœud de la couche *business*, comme discuté dans la section 2.3.2.3.

4. Enfin, dans un cadre réaliste, il convient également de disposer de l'ensemble des informations nécessaires à la création du modèle pour obtenir $\mathbb{E}(F)$.

La cohérence du portefeuille repose également sur la sélection des secteurs, la proportion de ceux-ci dans le portefeuille, ainsi que sur la taille des entreprises constituantes. L'ensemble des points évoqués sera discuté dans la suite de cette section.

3.1.1 Secteurs des entreprises

3.1.1.1 Une diversité de stratégies d'assurance

Dans le cadre de la constitution d'un portefeuille d'assurance dédié aux PME, plusieurs approches stratégiques peuvent être envisagées par l'assureur. L'objectif est de gérer au mieux les risques tout en répondant aux besoins spécifiques des entreprises et en restant attractif sur le marché choisi.

Une première option pourrait être de se concentrer sur des secteurs où des accords de branche ou des conventions collectives existent déjà, facilitant l'intégration des entreprises. Cela permettrait de standardiser l'offre d'assurance et les contrats proposés, tout en développant une connaissance spécifique du milieu qui, à terme, pourrait permettre de mieux conseiller les PME et ainsi offrir une grande valeur ajoutée.

D'un autre côté, l'assureur peut chercher à diversifier son portefeuille en répartissant les risques entre plusieurs secteurs d'activité. Par exemple, certains secteurs, comme la santé, sont plus exposés aux cybermenaces (voir figure 1.6), mais en les associant à des secteurs à moindre risque et avec une dépendance cyber plus faible, tels que l'agriculture, l'assureur pourrait réduire son exposition globale à des sinistres majeurs.

Il est également envisageable pour l'assureur de se spécialiser dans certains secteurs à risque élevé, tels que la santé ou les services financiers, qui sont souvent en première ligne face aux cybermenaces. Cette spécialisation peut offrir un avantage concurrentiel grâce à une expertise pointue dans la gestion des risques spécifiques à ces industries.

Enfin, les contraintes géographiques et les régulations locales peuvent influencer les choix stratégiques de l'assureur. Certaines régions imposent des normes plus strictes pour des secteurs stratégiques, tels que les infrastructures critiques, par exemple l'énergie ou les transports (comme discuté en partie 1.2.1.2). Dans ce cas, se concentrer sur ces secteurs pourrait offrir un positionnement avantageux sur un marché soumis à des régulations exigeantes et donc potentiellement à un besoin accru d'assurance.

Il est également à noter, comme il sera discuté en section 3.1.4.2, que l'intérêt des PME pour le produit vendu est primordial. Certaines PME seront, par exemple, moins affectées par les pertes d'exploitation, leur chiffre d'affaires ne dépendant pas d'une régularité des ventes tout au long de l'année. Une perte de disponibilité des réseaux informatiques aura donc peu d'influence sur leur productivité annuelle. Cela ne signifie pas que ces structures sont moins cyber-vulnérables en général, mais plutôt qu'elles ne subissent pas le risque de perte d'exploitation aussi fortement que d'autres.

Ces différentes stratégies permettent à l'assureur de définir la composition sectorielle de son portefeuille en fonction de ses priorités, qu'il s'agisse de diversification, de spécialisation ou de maximisation des opportunités de croissance.

3.1.1.2 Méthodologie générale pour le portefeuille fictif

Concernant le portefeuille fictif de ce mémoire, il n'était a priori pas question de discuter de stratégie d'assurance dans le domaine. L'idée générale pour la répartition des secteurs dans le portefeuille était donc d'adopter la répartition sectorielle des entreprises françaises, offrant ainsi une stratégie "neutre" de la part de l'assureur.

L'obtention de cette répartition est possible grâce à la base SIRENE, qui détaille le secteur de chaque entreprise française en activité. Il est ensuite envisageable de filtrer cette base afin de ne conserver que les PME ainsi que leur secteur d'activité. Il était ainsi prévu d'associer à chaque entreprise i de notre portefeuille un secteur d'activité aléatoire suivant cette répartition.

Il est à noter que plusieurs niveaux de granularité existent pour le secteur d'activité selon la nomenclature de l'INSEE. En particulier, la maille "Section" offre une division des activités suffisamment détaillée sans être trop spécifique. La figure 3.1 présente la répartition des PME en fonction du secteur à la maille "section".

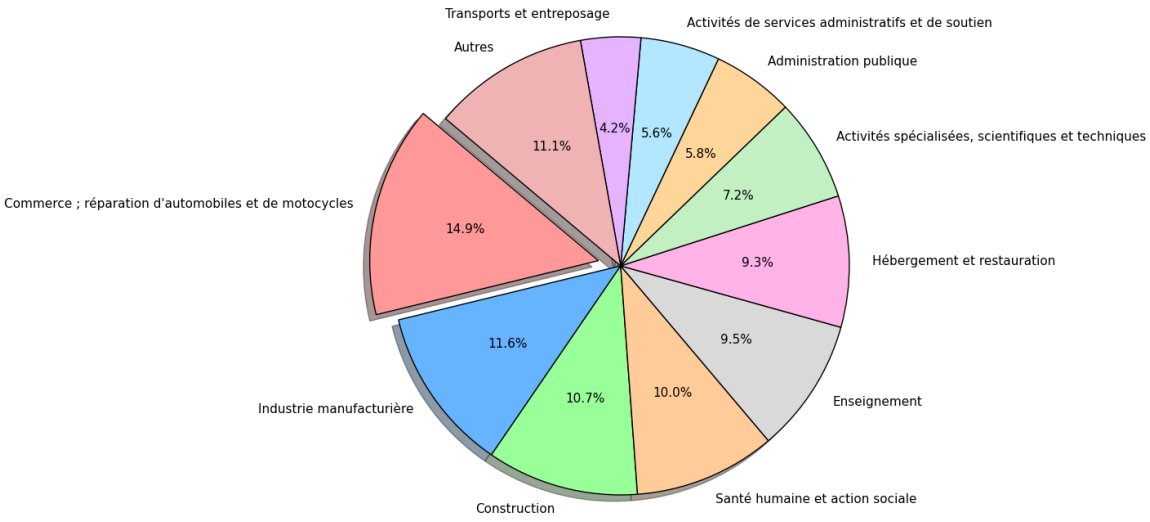


Figure 3.1: Répartition des secteurs à la maille "section" pour les PME françaises

Néanmoins, comme cela sera détaillé dans la partie 3.1.4, le facteur limitant de la création du graphe d'impact contraindra la sélection à un nombre restreint de secteurs. En conséquence, il ne sera pas possible, dans le cadre de ce mémoire, d'obtenir une répartition pleinement représentative des secteurs des PME françaises.

3.1.2 Taille et Chiffre d'Affaires des entreprises

Une fois que la constitution sectorielle du portefeuille fictif est décidée, réfléchissons à la taille et au chiffre d'affaires des différentes entreprises du portefeuille. Ces informations sont doublement importantes pour constituer le portefeuille :

1. Premièrement, le chiffre d'affaires de l'entreprise sera divisé entre les différents nœuds de *business* lors de la création du graphe d'impact (voir section 3.1.4).
2. Deuxièmement, l'architecture du réseau informatique d'une entreprise dépend de sa taille (itaia, 2023). En particulier, plus l'entreprise est grande, plus elle tendra à avoir du matériel informatique pour ses employés, ce qui augmente la surface d'attaque. Le graphe d'attaque est dépendant de ces actifs ; il est donc imaginable qu'il soit influencé par la taille (voir 3.1.3).

Il faut donc veiller à être réaliste dans le choix des tailles et chiffres d'affaires pour avoir un portefeuille cohérent.

3.1.2.1 Pourquoi les PME ?

Dans ce mémoire, le choix a été fait de se concentrer sur le secteur des PME. D'après le décret n°2008-1354, une PME est une entreprise dont l'effectif est inférieur à 250 personnes et dont le chiffre d'affaires annuel n'excède pas 50 millions d'euros (ou dont le total du bilan n'excède pas 43 millions d'euros). Plusieurs raisons pourraient pousser un assureur à se concentrer sur ce marché :

1. Premièrement, comme présenté dans la section 1.2.2.3, le secteur des grandes entreprises est déjà grandement développé, avec 94 % des grandes structures couvertes par une assurance cyber (AMRAE, 2024). Un assureur avec une offre d'accompagnement (comme il sera discuté dans la section 3.2.3) aura doublement du mal à entrer dans ce secteur, puisqu'il est déjà bien établi et que les grandes entreprises détiennent, pour la plupart, une connaissance cyber interne qui serait redondante avec une possible offre d'accompagnement.
   Le secteur des PME est, quant à lui, très peu développé (moins de 10 % des entreprises souscrivent à un contrat, toujours selon AMRAE). L'assurance manque d'attractivité, et l'utilisation d'un modèle comme celui développé dans le chapitre 2 pourrait permettre un accompagnement et donc une offre plus attractive (comme discuté dans la partie 1.2.4).
2. Deuxièmement, le modèle lui-même est adapté à cette taille de structure. Une vision *micro*, comme celle développée, devient inefficace pour les trop grandes structures. Elle demande une analyse très poussée des actifs, liaisons et du fonctionnement d'une entreprise, qui devient de plus en plus complexe à mesure que la structure grandit. Le temps de calcul est aussi un facteur limitant pour des entreprises de plusieurs milliers d'employés, puisque ce type de modèle est coûteux en ressources.
3. Enfin, sur le marché français et international, cette position stratégique est déjà tentée par plusieurs assureurs (ou courtiers) fournissant des services similaires avec un accompagnement aux entreprises. En France, Stoïk couvre les TPE, PME et ETI, et à l'étranger, Coalition se positionne sur le même secteur. Ce type de taille est donc potentiellement viable comme cible. De plus, se concentrer sur des petites entreprises permet à l'assureur de limiter le risque de déséquilibre dans un portefeuille où une seule entreprise détient un contrat avec une couverture trop importante par rapport aux autres.

Dans la prochaine section, la répartition des tailles d'entreprises et des chiffres d'affaires dans le portefeuille sera abordée.

3.1.2.2 Détermination du chiffre d'affaires et de la taille d'entreprise dans le portefeuille fictif

Un grand nombre de données et d'études sont disponibles sur les PME françaises. Elles seront utilisées pour créer des lignes cohérentes dans notre portefeuille et avoir des entreprises représentatives du marché français.

Tailles d'entreprises La base SIRENE, en plus de détailler les secteurs des entreprises françaises (comme discuté en section 3.1.1.2), détient aussi des informations sur leur taille. En effet, elle présente pour chaque entreprise une colonne *trancheEffectifsEtablissement*. Après avoir téléchargé la base sur le site de SIRENE en choisissant le critère de sélection "PME" (sans les TPE), il est possible d'analyser la répartition des catégories de tailles sur les 237 000 lignes de la base filtrée. Plus particulièrement, il est possible d'obtenir la répartition réelle des PME françaises en fonction des différents secteurs, comme présenté sur la figure 3.2.

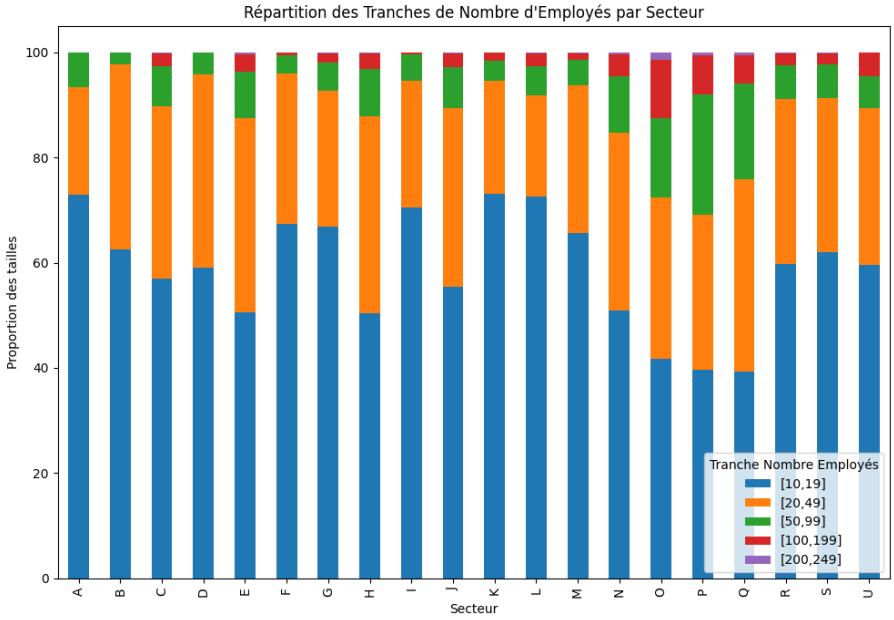


Figure 3.2: Répartition des tranches de nombre d'employés par secteur

Chaque lettre représente un secteur à la maille "Section" de la nomenclature d'activités française (INSEE, 2024). La base SIRENE est une représentation actualisée et fidèle des entreprises françaises. Ainsi, en supposant que le portefeuille de l'assureur respecte la proportion française des entreprises, il est possible d'affecter à chaque ligne (l_i) du portefeuille une tranche de nombre d'employés. Ces lignes ayant déjà un secteur défini par la section 3.1.1.2, il suffit alors d'attribuer à chaque ligne une tranche aléatoire (t_i) suivant la proportion visible dans la base pour ce secteur (figure 3.2).

À notre connaissance, il n'existe pas de statistiques plus précises, accessibles au public, sur la répartition du nombre d'employés. L'hypothèse sera donc faite que le nombre exact d'employés d'une entreprise suit une loi uniforme entière sur sa tranche. Une entreprise appartenant à la tranche (t_i) $[10; 19]$ aura donc un nombre d'employés (assimilé à une v.a. N_i) qui suit $N_i \sim \mathcal{U}(10, 19)$. Chaque ligne l_i étant une observation, le nombre d'employés de l'entreprise i, noté n_i, sera pris comme une observation de $\mathcal{U}(t_i)$.

Chiffre d'affaires des entreprises La base SIRENE ne dispose pas d'informations sur le chiffre d'affaires des entreprises. Cette donnée est difficilement accessible au public à l'échelle des entreprises. Néanmoins, l'INSEE (Institut National de la Statistique et des Études Économiques) dispose de ces informations via le dispositif Esane. Cet institut réalise des études pouvant être utiles dans le cadre de ce mémoire.

L'étude de l'INSEE (Kremp and Sklénard, 2019) fournit un aperçu sectoriel de la répartition par quantile de la productivité par équivalent temps plein (approximé par employé), appelée **productivité apparente du travail**. La figure 3.3 présente cette répartition pour les entreprises organisées en groupe. Bien que le graphique ne se concentre pas spécifiquement sur les PME, un autre graphique de l'étude montre que la répartition de la productivité apparente du travail des PME est semblable à celle des entreprises en général (figure 8 - Kremp and Sklénard, 2019). L'hypothèse forte sera faite que cette similarité se maintient à l'échelle des secteurs. La répartition au niveau de l'entreprise (et non de l'unité légale) sera celle retenue.

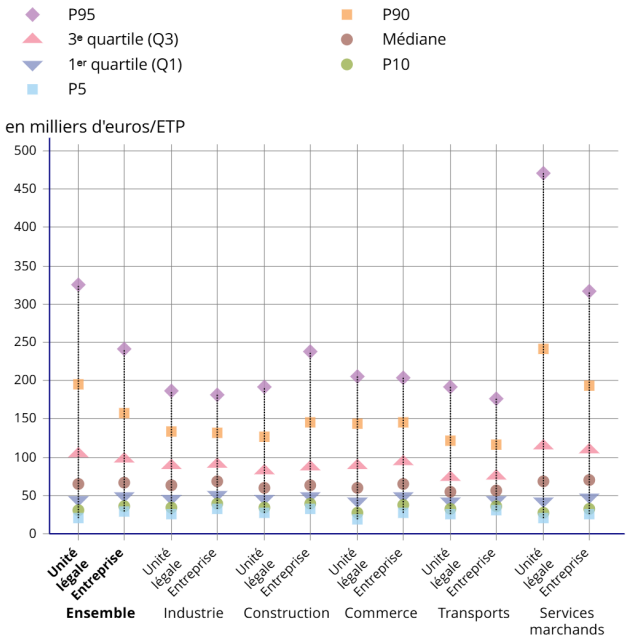


Figure 3.3: Distribution par secteur d'activité de la productivité apparente du travail mesurée au niveau de l'unité légale et au niveau de l'entreprise - INSEE

Chaque ligne (l_i) obtiendra donc une **productivité apparente du travail** (p_i) aléatoire en fonction de son secteur, suivant la répartition présentée par l'étude INSEE. La productivité de l'entreprise i sera alors calculée comme

$$prod_i = p_i \times n_i.$$

L'INSEE fournit dans plusieurs de ses études le *taux de valeur ajoutée* moyen par secteur, défini par la relation taux de valeur ajoutée $= \frac{\text{VA}_{\text{ht}}}{\text{CA}}$ (INSEE, 2016). Il est ainsi possible d'estimer un chiffre d'affaires pour l'entreprise i du portefeuille en utilisant la formule

$$ca_i = prod_i \times \text{taux de valeur ajoutée}(s_i).$$

La méthodologie pour déterminer le secteur, la taille et le chiffre d'affaires de chaque ligne i est ainsi établie. La section suivante s'intéressera à la création d'un graphe d'attaque cohérent pour chaque entreprise du portefeuille.

3.1.3 Création du Graphe d'attaque

Dans cette section, il sera question de l'obtention des graphes d'attaque pour chaque entreprise du portefeuille. Dans un premier temps, les procédures réalistes permettant d'obtenir ce graphe (possiblement applicables lors de la souscription d'un nouvel assuré) seront discutées. Dans un second temps, la procédure utilisée pour obtenir des graphes d'attaque dans le cadre du portefeuille fictif sera évoquée.

3.1.3.1 Procédure réaliste

D'un point de vue réaliste, pour obtenir le graphe d'attaque d'une entreprise, il faut (Mensah, 2019) :

1. Avoir la liste des vulnérabilités dans le réseau (dans chaque actif).
2. Avoir la topologie du réseau concerné ainsi que les configurations effectuées.
3. Utiliser un algorithme ou logiciel prenant en entrée les deux premiers points afin d'obtenir un graphe d'attaque.

Plusieurs logiciels existent pour scanner un réseau et identifier les différentes vulnérabilités présentes dans les éléments de ce réseau (ordinateurs, pare-feu, routeurs, etc.). En particulier, Nessus est un logiciel utilisé à cet effet ; il fournit non seulement les vulnérabilités mais aussi la topologie du réseau en question (Tatar et al., 2020). Les sorties de ce type de logiciel sont souvent utilisées comme entrées pour la création d'un graphe d'attaque.

Une fois les informations (1) et (2) collectées, plusieurs solutions permettent de créer le graphe d'attaque. TVA (Topological Analysis of Network Attack Vulnerability), MulVal (Multihost, Multi-stage, Vulnerability Analysis) et NetSPA sont souvent cités comme des outils permettant d'atteindre ce résultat (Tatar et al., 2020; Mensah, 2019). Ces solutions reposent sur différents types d'algorithmes techniques, dont plus de détails sont fournis dans la thèse de Mensah, 2019.

Il est ainsi théoriquement possible d'obtenir de manière automatique les différents éléments menant à la création des graphes d'attaque nécessaires au modèle. Du point de vue de l'assureur, cela pourrait passer par une installation unique (à la souscription) de solutions permettant le scan du réseau chez l'assuré. Cela permettrait la création de graphes d'attaque ainsi que l'envoi des données directement à l'assureur de manière périodique, assurant ainsi une mise à jour continue des données d'entrée et un suivi dynamique du risque de l'assuré. Évidemment, cette procédure pourrait s'avérer coûteuse. Néanmoins, son automatisation et l'apport en termes de connaissance du risque pourraient justifier l'investissement initial, la mise à jour étant peu contraignante comparée à l'installation initiale. La faisabilité technique est un premier pas vers l'applicabilité, mais des études supplémentaires sont nécessaires, notamment des tests dans des conditions réelles d'assurance et des analyses de coûts plus approfondies.

Dans le contexte de ce mémoire, nous n'avons toutefois pas les moyens de mettre en place ces solutions sur un nombre d'entreprises suffisant (ni même sur une seule) pour obtenir des résultats réels concernant les graphes d'attaque de notre portefeuille. La logique présentée dans cette partie sera néanmoins appliquée. Un réseau sera simulé pour chaque entreprise, des vulnérabilités CVE seront attribuées à chaque actif et un graphe d'attaque sera ensuite construit. Le détail de la méthodologie sera développé dans les sections suivantes.

3.1.3.2 Réseau d'entreprise

La première étape de notre méthodologie consiste à simuler un réseau d'entreprise pour chaque ligne i. Ce réseau est composé de divers actifs (ordinateurs, serveurs, etc.) et possède une topologie définie. L'idée sous-jacente à cette simulation est que l'attaquant exploite les "routes" établies par la topologie du réseau pour tirer parti des vulnérabilités d'un actif afin d'en compromettre un autre (Mensah, 2019).

Les actifs composant le réseau Un réseau d'entreprise se compose de plusieurs types d'actifs : ordinateurs, serveurs, etc. Spacey, 2023 propose une liste de 41 exemples d'actifs informatiques susceptibles de constituer un réseau. À partir de cette liste, il est possible de distinguer différentes catégories d'actifs. Dans le cadre de la création du graphe d'attaque (comme discuté dans la section 2.2.2.1, en suivant l'approche de Tatar et al., 2020), seuls les actifs de type *host* sont pris en compte. Cela inclut les actifs physiques (ordinateurs) ainsi que ceux qui peuvent être représentés par des actifs physiques (par exemple, les serveurs virtuels). Le graphe d'attaque se concentre sur ces actifs pour établir les connexions et identifier les vulnérabilités, ce qui justifie de traiter le problème à cette échelle.

Deux grandes catégories d'actifs ont été retenues dans le cadre de ce mémoire pour créer la liste d'actifs d'un réseau :

1. Les actifs de réseau, c'est-à-dire ceux qui permettent la transmission et le tri des informations au sein du réseau (Gomersall, 2024).
2. Les actifs terminaux, c'est-à-dire ceux qui reçoivent ou envoient les informations à travers le réseau (Axonius, n.d.).

Les actifs sélectionnés pour chaque catégorie sont répertoriés dans le tableau 3.1.

Catégorie	Actifs
Actifs de Réseau	Pare-feu, routeur, switch, routeur wifi
Actifs Terminaux	Ordinateur (PC), serveur (de données et web)

Table 3.1: Différents actifs réseau considérés dans ce mémoire

Ces actifs ont été sélectionnés en raison de leur rôle essentiel dans la structure et le fonctionnement d'un réseau. La majorité des sources consultées s'accordent sur l'importance de ces actifs dans la composition d'un réseau. Pour les besoins de ce mémoire, l'hypothèse est faite que cet ensemble représente les actifs présents dans un réseau typique. Toutefois, dans un contexte réel, la liste serait beaucoup plus étoffée et inclurait un plus grand nombre d'actifs.

La Topologie du réseau Lors de la création d'un réseau d'entreprise, une architecture spécifique est choisie en fonction des besoins et des objectifs de l'entreprise. Chaque architecture réseau définit la disposition physique et logique des nœuds (actifs) ainsi que la manière dont les données circulent entre eux. Les topologies courantes incluent le bus, l'étoile, l'anneau, le maillage et des configurations hybrides. Les différentes architectures sont illustrées à la figure 3.8. Chaque topologie présente des avantages et des inconvénients en termes de coûts, de gestion, de fiabilité et d'évolutivité, influençant la performance globale du réseau (Cohen, 2023).

Pour simplifier la création du réseau, celui-ci adoptera une structure de type multi-étoiles (voir la figure 3.9). Ce type de réseau est le plus répandu, ce qui justifie son choix, en particulier dans le

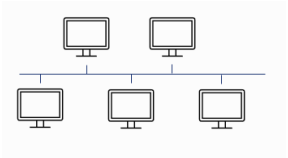


Figure 3.4: Architecture en bus

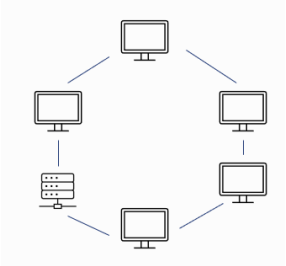


Figure 3.5: Architecture en anneau

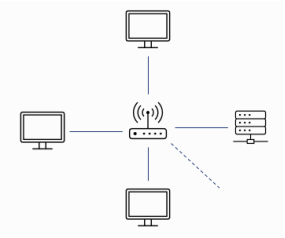


Figure 3.6: Architecture en étoile

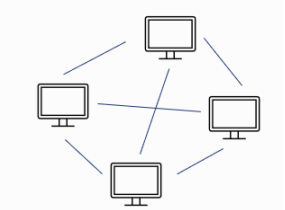


Figure 3.7: Architecture en maillage

Figure 3.8: Différentes architectures de réseau

cadre d'une gestion client-serveur où les ordinateurs de l'entreprise se connectent à un élément central (serveur) pour obtenir des ressources (EGE, 2024).

Création aléatoire d'un réseau Plusieurs hypothèses seront prises pour la création aléatoire des réseaux des entreprises du portefeuille :

1. La taille du réseau dépend de la taille de l'entreprise, en particulier de son nombre d'employés. Il sera fait l'hypothèse que le nombre d'ordinateurs de l'entreprise i est exactement égal au nombre d'employés (n_i).

2. Les switches et routeurs Wi-Fi serviront de points d'accès pour connecter les ordinateurs au réseau. Le routeur principal jouera le rôle de passerelle entre l'entreprise et Internet. Enfin, les *firewalls* seront positionnés à divers points du réseau afin de protéger certains actifs critiques, tels que les serveurs.

3. Plus une entreprise est grande, plus la probabilité qu'elle possède des serveurs est élevée. Pour une PME, il est possible d'imaginer des structures avec 0, 1, 2 ou 3 serveurs (un serveur dédié interne et un serveur externe pour le web, par exemple). Un plus grand nombre de serveurs est possible mais doit rester peu probable. C'est pour cela que le choix a été fait de modéliser le nombre de serveurs par une loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda = 2\left(\frac{n_i}{250}\right) + 1$. Ainsi, pour une petite entreprise (proche de 10 employés), la moyenne du nombre de serveurs sera proche de 1 et, pour une entreprise plus grande, cette moyenne sera proche de 3.

4. Le nombre de switches est lui aussi aléatoire et suit une loi uniforme

$$\mathcal{U}(\max(1, n_i//50), \max(1, n_i//25)).$$

   En effet, le nombre de switches pour une petite entreprise (10 employés) est souvent de 1, alors que pour une entreprise de 250 employés, il est généralement compris entre 4 et 9. Ce calcul est basé sur le nombre de ports par employé et le nombre moyen de ports sur un switch (FS Community, 2022).

5. La quantité de routeurs Wi-Fi ne croît pas de manière proportionnelle à la taille de l'entreprise. Une entreprise de 10 employés aura généralement 1 routeur, tandis qu'une entreprise de 250 employés en aura 2 ou 3, en fonction des choix techniques de l'organisation (Reddit, 2019). Pour modéliser ce choix, le nombre de routeurs Wi-Fi sera exprimé par $1 + k$, où k est la réalisation d'une variable aléatoire suivant une loi de Poisson $\mathcal{P}\left(\frac{n_i}{250}\right)$.

6. Un seul routeur sera considéré pour assurer la connexion avec Internet.

7. Le nombre de firewalls est proportionnel au nombre de switches et suit une loi uniforme

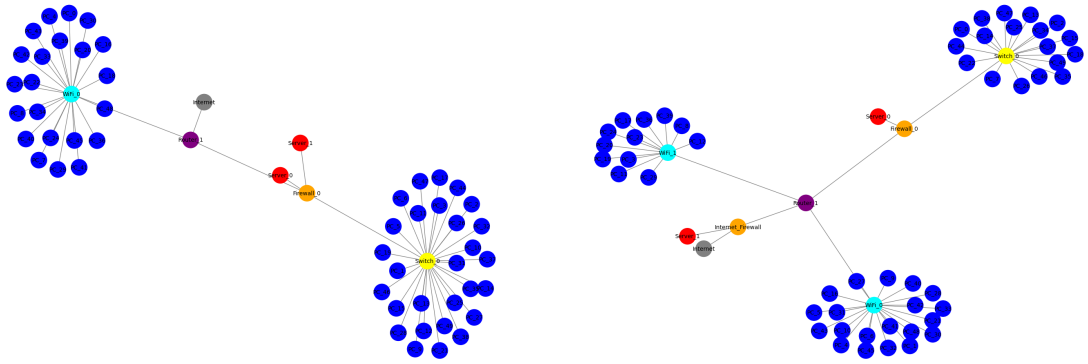
$$\mathcal{U}(\max(1, n_{\text{switches}}//3), \max(1, n_{\text{switches}}//2)).$$

   De plus, il y a 50 % de chances qu'un firewall supplémentaire soit ajouté à l'entrée du réseau pour renforcer la sécurité (notion de double barrière) (Prevost, 2023).

8. Pour des raisons de sécurité, les serveurs ne partageront pas le même switch que les ordinateurs. Un firewall les séparera.

Le choix des lois de Poisson a été fait car celles-ci sont adaptées aux processus de comptage. Lorsque le choix d'une loi uniforme a été privilégié, c'est parce que les informations semblent indiquer un intervalle de valeurs plutôt qu'une valeur préférentielle pour ce type d'actif. Grâce à ces hypothèses, un réseau peut être généré aléatoirement pour chaque entreprise. La figure 3.9 présente une génération aléatoire d'un réseau pour 50 employés (suivant les hypothèses établies). Les nœuds du réseau sont représentés par des couleurs distinctes : les ordinateurs sont en bleu foncé, les serveurs en rouge, les firewalls en orange, les switches en jaune, le routeur en violet et les routeurs Wi-Fi en bleu clair.

Il est à noter qu'il est possible de forcer la simulation d'une entreprise moins résiliente en réduisant le nombre de firewalls et de switches, ou en plaçant les serveurs sur les mêmes switches que les ordinateurs.



(a) Première simulation (b) Deuxième simulation

Figure 3.9: Simulation de deux réseaux d'entreprises de 50 employés

3.1.3.3 Attribution des CVE et Graphe d'Attaque

Cette partie présente la méthodologie d'attribution des vulnérabilités CVE aux différents actifs du réseau. Cela correspond à la première étape de la procédure réaliste pour créer un graphe d'attaque, c'est-à-dire obtenir la liste des vulnérabilités pour chaque actif (voir section 3.1.3.1). La méthodologie se divise en deux étapes. D'abord, une bibliothèque spécifique de failles CVE est créée pour chaque type d'actif (présenté dans le tableau 3.1). Ensuite, les failles pour chaque actif de l'entreprise sont sélectionnées.

Obtention d'une bibliothèque de vulnérabilités par type d'actif Les vulnérabilités présentes sur les différents actifs dépendent des programmes/services installés et/ou du *hardware* de ceux-ci.

La section 1.1.3.4 a déjà abordé les CPE, un moyen permettant d'identifier de manière unique les programmes vulnérables à une faille donnée. Pour obtenir une bibliothèque cohérente de failles, il n'est donc pas pertinent de créer une bibliothèque unique pour tous les actifs. Une telle approche pourrait, par exemple, associer des vulnérabilités spécifiques aux PC avec des serveurs. De plus, cette méthodologie ne prendrait pas en compte la fréquence d'apparition des failles. Si une vulnérabilité concerne un programme très peu utilisé, il serait préférable de la retrouver moins fréquemment dans les vulnérabilités du réseau.

Pour améliorer le réalisme dans la répartition des vulnérabilités, la méthodologie suivante a été appliquée pour la création de la bibliothèque :

1. Dans un premier temps, chaque type d'actif a été divisé selon les différents types de programmes/services qu'il peut utiliser. Le tableau 3.2 présente l'exemple de l'ordinateur.

Actif	Types de Programmes
Ordinateur (PC)	Système d'exploitation
	Navigateur Internet
	Antivirus
	Logiciel de partage d'information
	Logiciel de productivité
	Logiciel de communication

Table 3.2: Types de programmes pour l'ordinateur

2. Une fois la division effectuée, une étude de marché des différents types de programmes a été réalisée afin de déterminer les logiciels les plus utilisés pour chaque catégorie. Par exemple, le tableau 3.3 présente les parts de marché des différents systèmes d'exploitation selon l'étude Statista, 2024b.

Type de programme	Programmes	Part de marché
Système d'exploitation	Windows	75,51%
	Mac OS	15,06%
	Linux	3,16%
	Chrome OS	1,23%
	Autre	5,02%

Table 3.3: Parts de marché des systèmes d'exploitation des ordinateurs

3. Pour chaque programme, la liste des vulnérabilités affectant celui-ci est récupérée via l'API de la NVD (*National Vulnerability Database*, voir section 1.1.3.4). Pour garantir un certain réalisme, seules les CVE dont la date de mise à jour est postérieure à 2022 sont prises en compte, afin d'éviter d'inclure des vulnérabilités affectant des versions obsolètes du programme.

Ainsi, une bibliothèque de vulnérabilités est constituée, associée à une répartition des proportions des programmes dans chaque type d'actif.

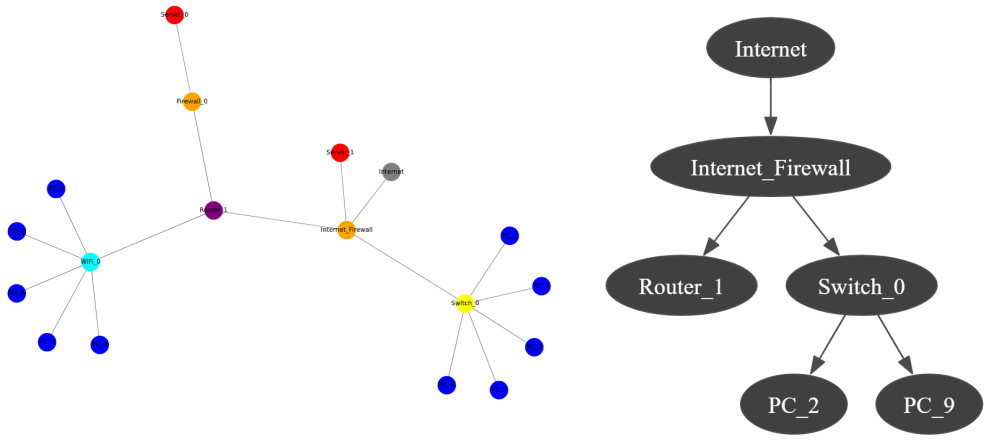
Répartition des failles par actif de l'entreprise La répartition des failles dans chaque actif de l'entreprise i se déroule en trois étapes :

1. Tout d'abord, une liste de programmes est associée à chaque actif. L'hypothèse retenue est qu'un actif donné dispose d'un unique programme par type de programme. Cette hypothèse se justifie par le fait qu'un ordinateur n'a généralement qu'un seul système d'exploitation, par exemple. Le choix aléatoire du programme est réalisé en fonction de sa part de marché. Ainsi, il est attendu que, dans l'ensemble du portefeuille, un peu plus de 75% des ordinateurs aient Windows comme système d'exploitation (voir tableau 3.3).
2. Ensuite, le nombre de failles existantes par actif est déterminé aléatoirement selon une loi de Poisson $\mathcal{P}(\lambda)$, avec λ fixé a priori à 3. Les programmes affectés par des failles sont sélectionnés par un tirage aléatoire avec remise, basé sur la loi de Poisson.
3. Enfin, les failles associées à chaque programme sont sélectionnées aléatoirement depuis la bibliothèque de failles.

En suivant ce procédé, il est possible de générer un faux scan de réseau complet pour une entreprise i. L'architecture du réseau est produite, et les failles pour chacun des actifs sont attribuées. Pour compléter cette partie de la simulation, le graphe d'attaque sera ensuite simulé.

Création de l'architecture du graphe d'attaque L'hypothèse est faite que les failles identifiées pour chaque actif permettent de construire le graphe d'attaque en fonction de l'architecture du réseau.

La création du graphe pour l'entreprise i s'effectue en parcourant le réseau informatique depuis Internet jusqu'aux nœuds terminaux. Pour chaque lien (a, b) existant dans le réseau, il est vérifié si le nœud fils b contient des vulnérabilités. Si c'est le cas, un lien entre a et b est ajouté au graphe d'attaque, et la procédure continue récursivement pour b. Si b ne présente aucune vulnérabilité, ses nœuds fils n'apparaîtront pas dans le graphe d'attaque. Si b présente plusieurs vulnérabilités, un seul lien sera créé. Toutefois, le calcul de $\overline{p(e)}$ (la probabilité d'exploitation réussie de la faille e) prendra en compte une logique OR, comme expliquée en section 2.2.2.1 lors de la description de la création du réseau bayésien d'attaque. La figure 3.10 illustre un réseau d'une entreprise de 10 employés ainsi que le graphe d'attaque associé.



(a) Réseau de l'entreprise simulé (b) Graphe d'attaque obtenu

Figure 3.10: Exemple de simulation pour 10 employés

Bien évidemment, les hypothèses de simulation s'accumulent et influencent le réalisme du graphe obtenu. Un véritable graphe d'attaque est généralement bien plus dense et complexe. Toutefois, cette problématique ne se poserait pas dans un cadre assurantiel réaliste où les données sont facilement accessibles, comme discuté en section 3.1.3.1.

3.1.4 Création du graphe d'impact

Dans cette section, nous aborderons l'obtention des graphes d'impact pour chaque entreprise du portefeuille. Dans un premier temps, nous discuterons des procédures réalistes permettant d'obtenir ce graphe, notamment lors de la souscription. Ensuite, une analyse sera effectuée à partir de plusieurs entretiens menés auprès de dirigeants de PME dans le cadre de ce mémoire. Enfin, l'application au portefeuille fictif sera étudiée.

3.1.4.1 Procédure réaliste

Le graphe d'impact est la clé de voûte du passage de la perte "cyber" à la perte économique. La cohérence des différentes méthodes d'obtention de celui-ci est donc cruciale pour assurer le bon fonctionnement du type de modèle présenté dans ce mémoire.

Cependant, l'utilisation de graphes d'impact reste aujourd'hui un sujet très théorique. Dans Tatar et al., 2020, peu de détails sont donnés sur les méthodologies de création dans un contexte réaliste. Ce sujet est brièvement évoqué dans Bahşi et al., 2018, où des pistes de méthodologies semi-automatiques sont suggérées à partir d'études connexes. Toutefois, leur utilisation semble limitée à certaines parties du graphe dans des cas spécifiques. L'utilisation de la connaissance d'experts pour construire l'architecture de ce graphe est également discutée. La principale faiblesse de cette méthode réside dans le risque d'oubli de certains nœuds ou dépendances critiques. Néanmoins, elle demeure, à notre connaissance, la méthode la plus accessible à ce jour.

Une recherche approfondie est nécessaire dans ce domaine afin de rendre l'obtention du graphe d'impact parfaitement applicable et robuste (Bahşi et al., 2018). Des pistes existent néanmoins, et l'expertise humaine peut constituer une solution intermédiaire en attendant l'élaboration de méthodes plus fiables.

L'assureur pourrait, lors de la souscription, fournir un questionnaire à l'entreprise afin de recueillir les informations nécessaires à la création du graphe d'impact. Ce questionnaire devrait permettre une construction progressive et structurée du graphe. Dans le cadre de ce mémoire, un tel questionnaire a été élaboré, non seulement dans l'objectif de proposer une méthodologie de création de graphe d'impact, mais également comme base de la section 3.1.4.2. Ce questionnaire est disponible en annexe (A.2).

3.1.4.2 Entretiens

Afin de mieux comprendre le secteur des PME, ainsi que l'influence des cybermenaces sur leur activité — notamment en ce qui concerne la perte d'exploitation — et d'élaborer un graphe d'impact cohérent, nous avons mené des entretiens avec trois dirigeants de PME opérant dans des secteurs différents.

Ainsi, la dirigeante d'un atelier d'architecture dans le secteur de la construction, le PDG d'une école d'enseignement supérieur privé dans le domaine de l'éducation, ainsi que le dirigeant d'une entreprise manufacturière dans le secteur industriel ont été interrogés. Lors de ces entretiens, les contours du graphe d'impact ont été délimités, tout en discutant des répercussions des risques cyber sur leur activité. Par ailleurs, l'intérêt qu'ils pourraient porter à une assurance cyber a été abordé, en particulier si celle-ci proposait des services annexes proactifs, tels que des mesures de prévention en cas de failles.

Plusieurs informations très intéressantes ont émergé de ces entretiens.

1. Il apparaît clairement que le graphe d'impact doit être adapté à chaque entreprise.

   Au fil des discussions avec les différents dirigeants, il a été constaté que la structure des services et du *business* des entreprises est étroitement liée à leur fonctionnement interne. Chaque entreprise possède ses propres spécificités, ce qui se reflète dans son graphe d'impact, contribuant ainsi à sa valeur ajoutée. Plus particulièrement, deux entreprises opérant dans le même secteur peuvent avoir des graphes d'impact très différents. Cette observation a été confirmée par tous les professionnels interrogés. Par exemple, pour le dirigeant de l'école privée d'enseignement supérieur, son établissement se distingue par une spécialisation en médecine animale, avec de nombreuses pratiques et recherches réalisées sur des animaux, une spécificité ne se retrouvant pas dans d'autres écoles. De son côté, la dirigeante de l'atelier d'architecture a souligné que la taille de la structure, ainsi que la nature des activités (maçonnerie, charpenterie, architecture, etc.), influencent largement la couche de services du graphe et le modèle de l'entreprise. Enfin, pour le dirigeant de l'entreprise manufacturière, la spécialisation de son entreprise joue un rôle crucial dans l'organisation des processus et l'architecture globale de l'entreprise.

2. D'un point de vue des cybermenaces, toutes les entreprises ne sont pas exposées de la même manière.

   Par exemple, pour le dirigeant de l'école privée, la notion de perte d'exploitation due à une cyberattaque lui semblait négligeable, étant donné que le modèle économique de son établissement repose sur des frais payés en une seule fois au début de l'année. Toutefois, il reconnaît que les pertes structurelles liées à une cyberattaque, notamment la suppression des données, pourraient être considérables, car cela impliquerait de reconstituer tous les processus, les supports de cours, etc. Néanmoins, il ne perçoit pas la menace cyber comme fatale pour la survie de son entreprise. Il explique également que la faible dépendance de son école aux technologies numériques, avec un enseignement majoritairement présentiel et de nombreuses pratiques sur les animaux, réduit l'impact potentiel d'une cyberattaque. Il illustre son propos en prenant l'exemple d'une école informatique, qui serait bien plus vulnérable à ce type de problème. Cet exemple appuie notre constat concernant la variabilité de l'exposition aux cybermenaces au sein d'un même secteur.

   En revanche, le dirigeant du secteur industriel a identifié la cybermenace comme l'une des plus grandes menaces pour son entreprise. Ayant subi une attaque par ransomware en 2023, il a une expérience directe de l'impact dévastateur de ce type d'attaque. Pour lui, la perte d'exploitation représente un risque majeur, mais c'est surtout la destruction des données stockées sur ses serveurs qui pourrait entraîner une perte totale de l'entreprise, menant à une fermeture quasi inévitable.

3. Les PME ne perçoivent pas toujours l'ampleur réelle des cybermenaces.

   Lors de notre discussion sur l'attaque subie par son entreprise, le dirigeant du secteur industriel a souligné que la menace cyber est souvent mal comprise et sous-estimée dans les PME. Dans son cas, bien qu'il ait conscience de l'existence de cette menace, il a fallu attendre l'attaque pour qu'il prenne pleinement la mesure des pertes qu'elle pouvait engendrer. En raison de la numérisation complète de ses processus (plans manufacturiers sur tablettes, employés utilisant des ordinateurs, structure réseau avec serveurs, etc.), son entreprise était en réalité extrêmement vulnérable à ce type de risque. Ce constat fait écho à la section 1.1.4.1, où la tendance des PME à minimiser l'importance des risques cyber a été discutée.

4. Un besoin d'assurer différemment.

   Lors des entretiens, les dirigeants ont été interrogés sur leur perception de l'assurance cyber, en présentant le cadre de ce mémoire ainsi que le modèle basé sur les failles. Une discussion particulièrement révélatrice s'est tenue avec le dirigeant de l'industrie manufacturière. Il a

exprimé que l'idée d'associer l'assurance avec des services de conseil représente une réelle valeur ajoutée pour les PME comme la sienne. En effet, n'ayant pas de service informatique interne, il trouve très intéressant qu'un contrat d'assurance inclue également des conseils sur la gestion des risques cyber. Cette observation renvoie à la section 1.2.4, où l'importance croissante de l'accompagnement dans les contrats d'assurance pour répondre aux besoins spécifiques des PME a été discutée.

5. Concernant les graphes d'impact, une série de questions basées sur celles de l'annexe A.2 a été utilisée. Chaque dirigeant y a répondu, ce qui a permis de récupérer l'architecture des graphes, notamment celui du secteur industriel. Ces informations serviront de base pour la section 3.1.4.3. Nous avons également pu identifier les points forts et les limites de cette méthode pour obtenir le graphe d'impact.

   D'une part, l'entretien permet une véritable discussion avec les dirigeants et offre une compréhension approfondie du fonctionnement interne de leur entreprise. Cependant, la structure d'un graphe d'impact est difficile à expliquer, notamment la couche des services, qui peut être interprétée de différentes manières. Bien que les dirigeants connaissent très bien leur entreprise, il peut être difficile d'entrer dans des détails quantitatifs extrêmement précis, comme les paramètres α et β, lors des entretiens. De plus, le sujet est très théorique et les exemples concrets sont rares, ce qui peut rendre l'application de ces concepts quelque peu abstraite.

   Néanmoins, cette méthode a fourni des indications précieuses sur les liens d'opérabilité au sein des différentes entreprises, ce qui constitue un point de départ prometteur pour la construction des graphes d'impact.

3.1.4.3 Application au portefeuille fictif

La création de graphes d'impact pour chaque ligne de notre portefeuille constitue l'ultime étape avant de pouvoir appliquer le modèle et analyser son comportement en fonction de différentes situations extérieures, comme cela sera fait dans les sections 3.2 et 3.2.4. Les entretiens ont révélé qu'il est complexe de générer des graphes d'impact réalistes pour un groupe d'entreprises, chaque entreprise ayant un fonctionnement propre. Par conséquent, il ne sera pas possible, dans ce mémoire, de créer un graphe d'impact par secteur, comme cela avait été envisagé dans la section 3.1.1.2. Des recherches approfondies (ou des entretiens) sont nécessaires pour bien comprendre le fonctionnement spécifique de chaque entreprise afin d'établir une architecture de graphe adéquate.

Néanmoins, afin de réaliser les tests, deux graphes types seront créés. L'hypothèse retenue est que toutes les entreprises du portefeuille suivront l'une ou l'autre de ces architectures. Le premier graphe, construit à partir des entretiens, sera utilisé pour représenter le secteur industriel. Les paramètres α et β seront choisis aléatoirement, en tenant compte des informations qualitatives recueillies lors des entretiens. Par exemple, si un lien a été jugé très important, la valeur de α sera sélectionnée de manière aléatoire, mais proche de 1. Ainsi, le secteur de l'industrie sera inclus dans le portefeuille.

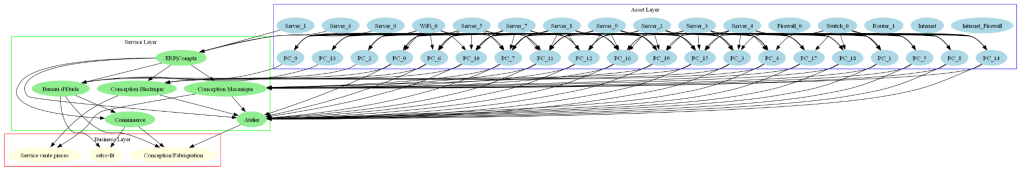
Pour plus de diversité, il a également été décidé d'étudier le fonctionnement d'une entreprise de commerce en ligne, étant donné que ce type d'entreprise est particulièrement sensible aux pertes d'exploitation dues à des cyberattaques. En effet, leur chiffre d'affaires dépend directement de leurs ventes en ligne, qui reposent sur des infrastructures informatiques. Le secteur du commerce en ligne sera donc représenté dans le portefeuille, et une répartition proportionnelle entre ces deux secteurs sera effectuée en fonction de la répartition sectorielle, telle que visible dans la figure 3.1.

Pour la couche d'actifs du graphe d'impact de l'entreprise i, il est nécessaire de prendre en compte le réseau simulé préalablement (qui sera utilisé en entrée de la procédure de création aléatoire). Cela permet d'associer à chaque actif un type spécifique, par exemple, si l'entreprise possède différents types

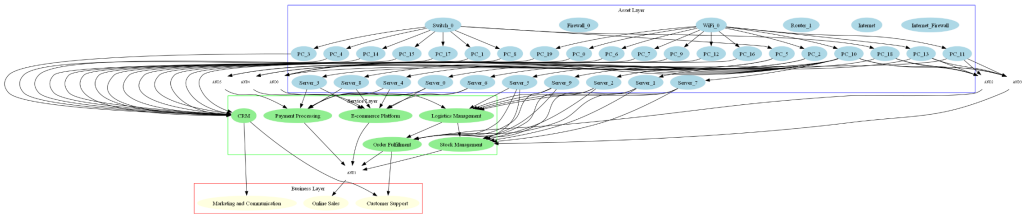
d'employés ou de serveurs, et de définir les dépendances correspondantes. La génération du graphe d'impact s'intègre ainsi en fin de parcours de la génération du portefeuille fictif.

Enfin, en ce qui concerne la perte d'exploitation, une proportion du chiffre d'affaires sera associée à chaque nœud business. Le chiffre d'affaires sera calculé selon la méthode décrite dans la sous-section 3.1.2.2, et la proportion ajustée en fonction des informations recueillies lors des entretiens (en particulier pour le secteur industriel).

La figure 3.11 présente deux graphes d'impact générés aléatoirement à partir des architectures établies : l'un pour le secteur industriel et l'autre pour le commerce en ligne, chacun comprenant 20 employés. Les trois couches y sont représentées par des couleurs distinctes : la couche d'actifs est en bleu, la couche de services en vert, et la couche business en rouge. L'impact de ces différentes architectures sera analysé dans les sections suivantes de ce mémoire.



(a) Graphe d'impact généré pour une entreprise dans l'industrie



(b) Graphe d'impact généré pour une entreprise dans le commerce en ligne

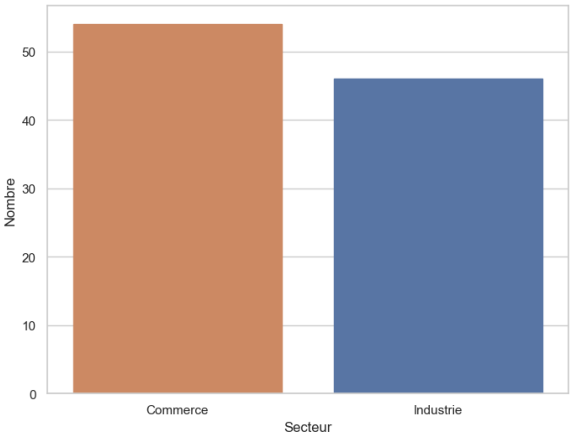
Figure 3.11: Exemples de génération aléatoire de graphes d'impact

3.2 Quantification Dynamique dans le cadre du portefeuille PME

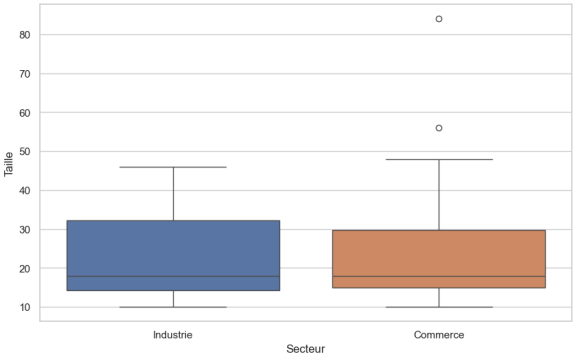
Dans cette section, le modèle sera appliqué à un portefeuille créé grâce à la méthodologie développée dans la section 3.1. Dans un premier temps, une discussion sera menée sur les résultats du modèle appliqué au portefeuille, en tenant compte des failles simulées selon la méthodologie de la sous-section 3.1.3 (assimilée au temps $t = 0$). Ensuite, l'évolution temporelle des primes pures du portefeuille sera examinée, avec une attention particulière portée au concept d'évaluation "dynamique". Enfin, l'engagement de l'assureur dans la réduction du risque sera discuté. Tout au long de cette section, des liens avec des situations réelles seront établis, et les limites potentielles à prendre en compte lors de l'analyse des résultats et des conclusions seront soulignées.

Avant de commencer cette section, revenons brièvement sur la création du portefeuille. Un portefeuille composé de 100 entreprises ($n = 100$) a été retenu, principalement en raison de contraintes techniques liées à la machine utilisée pour réaliser ce mémoire. Ces limitations empêchent d'exécuter le modèle avec la rapidité nécessaire pour effectuer un nombre suffisant de simulations (1 000 par entreprise dans la suite de cette section), tout en gérant un portefeuille de taille beaucoup plus im-

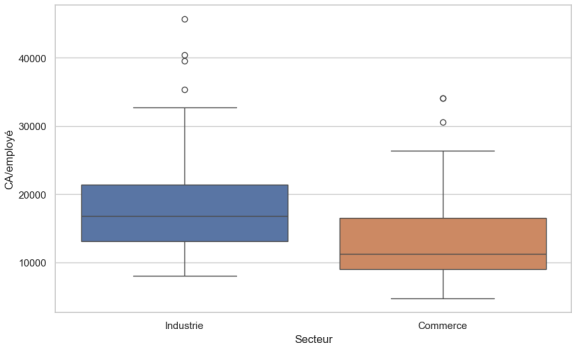
portante.
Comme précisé dans la partie 3.1.4.3, deux secteurs sont étudiés dans la suite de ce mémoire. La répartition suit les proportions relatives de la figure 3.1, représentant la répartition des secteurs français à l'échelle des PME. La répartition du portefeuille entre ces deux secteurs est visible sur la figure 3.12a.



(a) Répartition des secteurs du portefeuille fictif



(b) Répartition du nombre d'employés en fonction du secteur



(c) Répartition du chiffre d'affaires par employé en fonction du secteur

Figure 3.12: Diverses informations sur le portefeuille fictif.

De plus, la taille des entreprises a été calculée selon la méthodologie présentée dans la partie 3.1.2.2 et est visible pour les deux secteurs sur la figure 3.12b. Il en est de même pour le chiffre d'affaires par employé, dont la répartition par secteur est illustrée sur la figure 3.12c.

3.2.1 Application du modèle au portefeuille

Cette sous-section se penchera sur l'application du modèle à un portefeuille fictif. Dans un premier temps, les résultats du modèle seront analysés à l'échelle d'une entreprise. Nous procéderons ensuite à une comparaison des résultats entre plusieurs entreprises, en mettant en lumière l'influence du graphe d'attaque. Enfin, une perspective plus globale sera adoptée en étudiant le portefeuille dans son ensemble, avec une attention particulière portée à l'impact du secteur d'activité et du chiffre d'affaires.

3.2.1.1 Résultats de l'application du modèle à une entreprise

Cette partie étudiera la sortie du modèle pour une entreprise du portefeuille. Cela permettra d'introduire un premier exemple concret et de discuter de l'application du modèle présenté dans la

sous-section 2.3.2.

Comme discuté dans la partie 2.3.2.3, le calcul de la perte est effectué sur k simulations (ici $k = 1000$). Pour chaque itération, une simulation d'attaque j est effectuée, ce qui permet d'obtenir une distribution de la perte pour l'entreprise i en regardant les pertes à chaque simulation (C_j^i). Il suffit ensuite de prendre la moyenne de ces différentes simulations pour avoir une approximation de l'espérance de perte pour l'entreprise. Enfin, la prime pure de cette entreprise (sans franchises ni limites) peut s'exprimer comme

$$\pi_i = \mathbb{E}(F_i)\frac{1}{k}\sum_{j=1}^{k} C_j^i.$$

Avec F_i qui est la variable aléatoire représentant la fréquence d'attaque de l'entreprise i dans l'intervalle de temps de la prime (année, mois, ...). Cet élément a été présenté lors de la sous-section 2.3.2. Il est rappelé que l'étude de cette fréquence est hors du cadre de ce mémoire. Ce qui est noté ici $\mathbb{E}(F_i)$ pourrait par exemple être le résultat d'un autre modèle. La différence avec un coût-fréquence classique est qu'ici le modèle est centré sur une attaque et non un sinistre (une attaque pouvant échouer et ne causer aucun sinistre, donc C_j^i peut être nul). L'hypothèse est que la fréquence d'attaque est bien plus simple à estimer que la fréquence de sinistre. Cette idée sera discutée plus en détail dans la section 3.3.

Résultats du modèle L'exemple de l'entreprise *ENT0Com* (première entreprise du portefeuille) sera pris pour analyser le résultat donné par le modèle. Cette entreprise du secteur du commerce compte 16 employés et réalise un chiffre d'affaires de 201 464 €. Son graphe d'attaque est visible sur la figure 3.13.

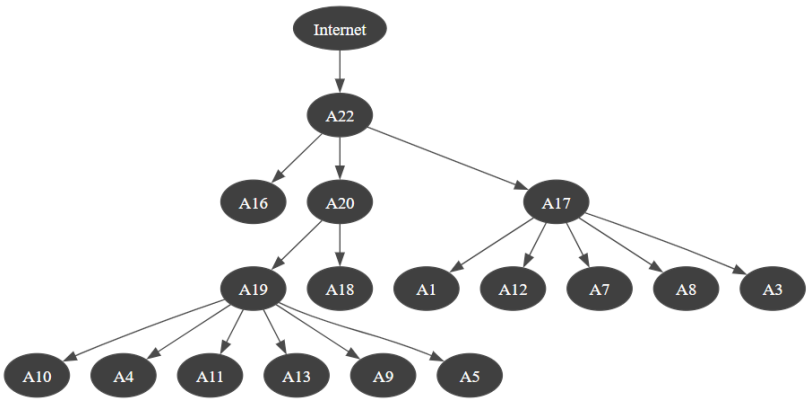


Figure 3.13: Graphe d'attaque de l'entreprise *ENT0Com*

Après l'application du modèle à cette entreprise, 1000 valeurs de C_j^i sont obtenues. La distribution des pertes (sachant que l'entreprise est attaquée) peut être visualisée en consultant l'histogramme des pertes de la figure 3.14.

L'observation de l'histogramme suggère une distribution des coûts à queue lourde. Il est notable que près de 7% des attaques n'entraînent aucun coût. Étudier cette valeur pourrait constituer un indicateur pertinent pour l'assureur afin d'évaluer la robustesse des entreprises présentes dans son portefeuille.

Avec ce résultat, il est aussi possible de calculer la moyenne des coûts. Pour l'entreprise étudiée, cette valeur est de 1 902 €, soit 1% du chiffre d'affaires. Cela peut s'analyser comme la valeur moyenne

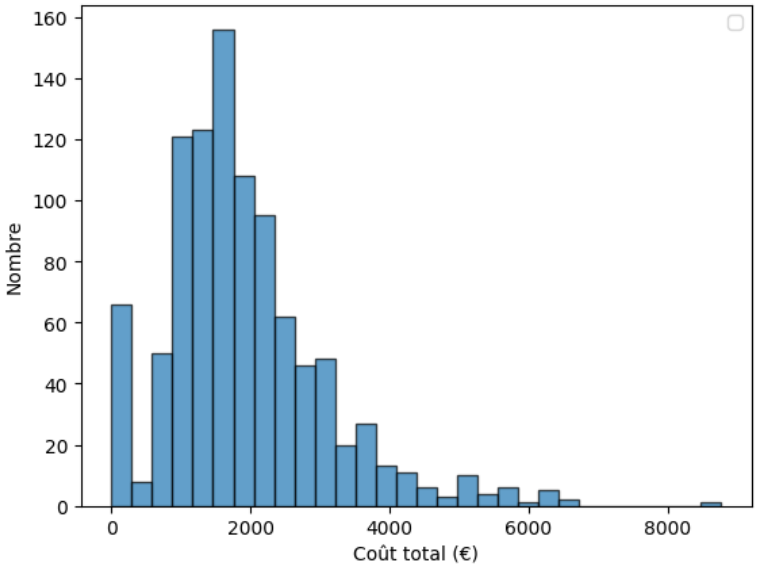


Figure 3.14: Histogramme des coûts pour l'entreprise *ENT0Com*

perdue par attaque subie par l'entreprise. Une comparaison avec les valeurs réelles sera effectuée au niveau du portefeuille dans la section 3.2.1.3.

Il est important de souligner que les résultats quantitatifs du modèle ne doivent en aucun cas être interprétés comme reflétant la réalité, et aucune conclusion définitive ne peut être tirée quant à l'adéquation du modèle avec le monde réel. Le modèle repose sur un ensemble de facteurs qui ont été simulés ou supposés. Bien qu'un effort considérable ait été fourni pour établir des hypothèses cohérentes et des simulations réalistes, les résultats obtenus ne visent qu'à illustrer des applications potentielles du modèle dans un contexte actuariel.

En particulier, le paramètre ρ du modèle, qui représente la probabilité de passer d'un état infecté à un état remis (lors de la phase de remédiation), joue un rôle important dans les résultats du modèle. Pour la suite de ce mémoire (et les résultats présentés précédemment), l'hypothèse est faite que $\rho = 0.1$. Cette valeur signifie que chaque jour, un actif a 10% de chance d'être remis. Il est néanmoins possible d'étudier l'évolution de la valeur de la prime en fonction de ce paramètre en prenant l'exemple de l'entreprise *ENT0Com*.

Sensibilité du modèle par rapport à ρ Ce paragraphe étudiera les résultats du modèle en fonction de ρ. Pour cela, trente valeurs de ρ seront prises de manière équirépartie entre 0 et 1. La figure 3.15 présente les résultats du modèle, avec la valeur de ρ en abscisse et le coût total (ou sévérité de l'attaque) en ordonnée. Pour chaque ρ, une boîte à moustaches illustre la distribution des coûts observés lors des différentes simulations. Cela permet de visualiser non seulement la valeur moyenne des pertes, mais aussi la dispersion des résultats pour chaque cas.

Avant d'analyser ces résultats, la forme générale de la courbe peut être prédite en réfléchissant au rôle de ρ. En effet, plus ρ tend vers 0, plus la probabilité qu'un actif infecté soit remédié à chaque pas de temps diminue (donc plus le temps d'infection est long). Puisque le coût total est la somme du coût de chaque jour infecté, nous pouvons donc intuiter (en nommant C^i la perte moyenne pour l'entreprise i lors de la simulation) que

$$C^i(\rho) \underset{\rho \to 0^+}{\to} +\infty.$$

Cela signifie que si la probabilité de remédiation est quasi nulle, les actifs restent infectés indéfiniment, et le coût total de l'attaque devient infini pour l'entreprise.

De la même manière, lorsque ρ tend vers 1, il est attendu de C^i que

$$C^i(\rho) \underset{\rho \to 1^-}{\to} D_i^1.$$

En effet, lorsque la probabilité que l'entreprise se rétablisse lors du premier pas de temps est de 1, seul le coût associé à ce premier pas sera significatif. Cette quantité est nommée ici D_i^1.

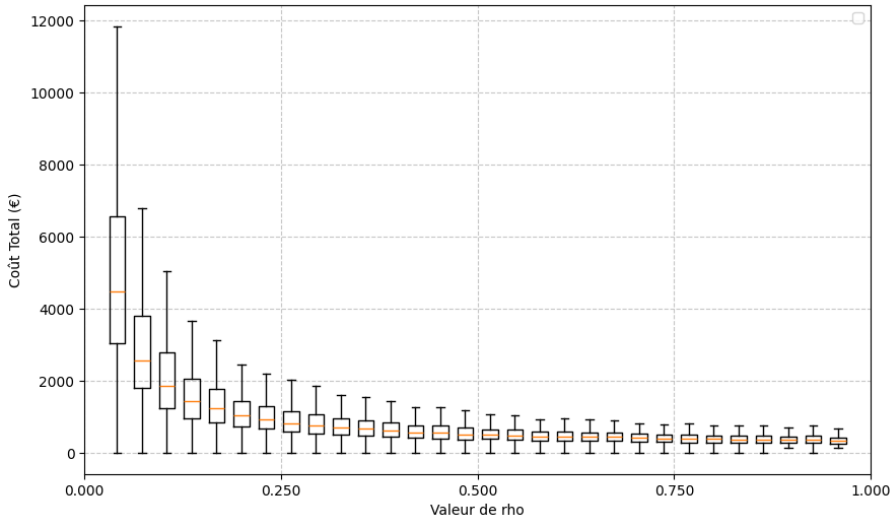


Figure 3.15: Coûts totaux en fonction de ρ pour l'entreprise *ENT0Com*

Cette tendance est particulièrement visible sur la figure 3.15. Il apparaît clairement que pour des valeurs de ρ proches de 0, la moyenne des coûts totaux augmente fortement. L'allure de la figure, formée par l'ensemble des valeurs, semble alors tendre vers $+\infty$, comme cela avait été prédit théoriquement.

Un autre phénomène observable est l'élargissement de la boîte à moustaches, ce qui indique une augmentation de la variance à mesure que ρ diminue. Cette observation est cohérente : pour des valeurs de ρ faibles, il existe une grande disparité entre deux attaques, où l'une permet à l'attaquant de capturer un actif supplémentaire par rapport à l'autre. Cet actif supplémentaire, s'il est pris en compte, augmente considérablement le coût, surtout lorsque la probabilité de remédiation entre deux périodes est faible (puisque cet actif supplémentaire restera plus longtemps non remédié).

Ainsi, lorsque ρ diminue, la différence entre les coûts résultant de deux attaques distinctes se renforce, entraînant une augmentation de la variance. De plus, cette variance, tout comme la moyenne, semble également tendre vers $+\infty$ lorsque ρ s'approche de 0.

Cette étude démontre que les résultats du modèle sont extrêmement sensibles à la valeur de ρ et qu'une analyse approfondie du temps de remédiation des différents actifs dans une entreprise est nécessaire pour que les résultats donnés par le modèle soient cohérents avec la réalité.

Discussion sur la tarification individuelle À partir des résultats du modèle, il devient possible de réfléchir à la tarification d'un contrat de perte d'exploitation pour une entreprise du portefeuille en se basant sur la prime pure. En plus des notions de limites et de franchises, que l'assureur peut ajuster pour moduler son exposition au risque, la principale force de ce type de modèle réside dans sa

capacité à évaluer le risque de manière continue. Puisque le modèle n'a pas besoin de sinistres pour fonctionner, il peut être relancé régulièrement afin de suivre l'évolution du risque pour l'entreprise. Cette capacité est discutée dans la section 3.2.2.

Comme le souligne Hillairet and Lopez, 2022, l'échelle temporelle du risque cyber est relativement courte. Grâce à ce modèle, l'assureur a la possibilité de choisir la durée de validité de son contrat, pouvant ainsi opter pour un maillage plus fin que la traditionnelle période annuelle. Seule la valeur de $\mathbb{E}(F_i)$ est affectée par un changement de durée, les simulations reposant sur l'hypothèse qu'une attaque se produit. Il devient donc envisageable de proposer un contrat mensuel, avec une prime recalculée chaque mois, permettant à l'assureur de mieux évaluer le risque "instantané" de son portefeuille.

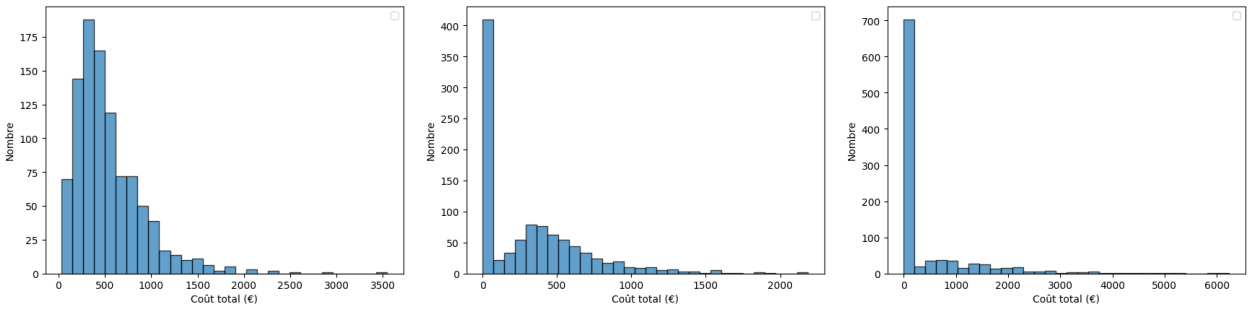
Pour illustrer cette idée, il est possible d'évoquer la faille EternalBlue, qui a permis l'attaque WannaCry (comme mentionné dans le paragraphe 1.1.1.1). Cette faille, découverte en avril 2017, n'aurait pas pu être anticipée par un contrat ayant évalué les risques en début d'année, ce qui aurait empêché l'assureur d'ajuster la prime en conséquence ou encore le provisionnement.

Cette possibilité d'appliquer le modèle à une échelle de temps plus faible répond à la question soulevée concernant l'évolutivité du risque et les limites de la modélisation classique, abordée dans la section 1.2.4.

3.2.1.2 Différences de résultats entre entreprises

Cette section abordera les différences de comportements des entreprises face au modèle, en mettant en évidence l'impact du graphe d'attaque sur les résultats. Cela conduira à discuter d'une limite inhérente à la méthode de création du portefeuille fictif.

Pour illustrer cette hétérogénéité des entreprises, il est possible de comparer les résultats de différentes entreprises similaires en taille et en secteur. En particulier, il est possible d'observer la figure 3.16.



(a) Histogramme des coûts pour l'entreprise *ENT26Com*

(b) Histogramme des coûts pour l'entreprise *ENT20Com*

(c) Histogramme des coûts pour l'entreprise *ENT15Com*

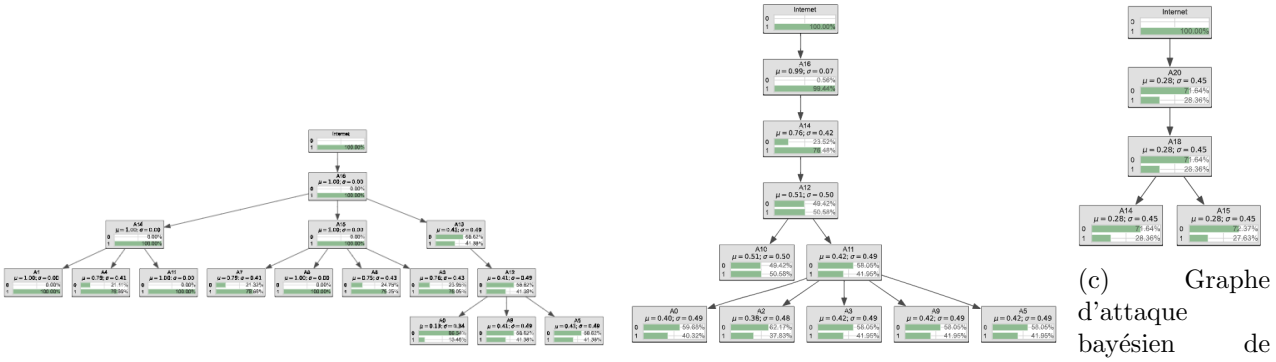
Figure 3.16: Comparaison de la répartition des coûts entre trois entreprises

Les trois entreprises présentées appartiennent toutes au secteur du commerce, ce qui leur confère un graphe d'impact similaire. De plus, leur taille est comparable, avec respectivement 12, 10 et 14 employés. Cependant, il est évident que la répartition des coûts (et donc des risques) varie d'une entreprise à l'autre.

Cette différence dans l'évaluation de chaque entreprise constitue également l'une des forces du modèle. Ce dernier est spécifiquement adapté à chaque entreprise et évalue le risque en se basant sur le graphe d'attaque. Ainsi, bien que ces trois entreprises paraissent similaires, elles ne présentent en réalité pas le même risque cyber, et le modèle est capable de saisir cette nuance grâce à l'analyse du

graphe d'attaque.

En particulier, l'observation des graphes d'attaque permet de mieux comprendre ces différences de forme. La figure 3.17 présente les graphes d'attaques bayésiens de chaque entreprise avec l'inférence effectuée.



(a) Graphe d'attaque bayésien de l'entreprise *ENT26Com*

(b) Graphe d'attaque bayésien de l'entreprise *ENT20Com*

(c) Graphe d'attaque bayésien de l'entreprise *ENT15Com*

Figure 3.17: Les différents graphes d'attaques de trois entreprises

En observant la figure, il est clairement constatable que les graphes d'attaques des différentes entreprises diffèrent à plusieurs niveaux, notamment en ce qui concerne le nombre d'actifs accessibles par l'attaquant (par exemple, l'entreprise 15 n'en compte que 4, tandis que l'entreprise 26 en a 15). De plus, la topologie du graphe varie également : celui de l'entreprise 26 se divise bien plus tôt que celui de l'entreprise 20, offrant à l'attaquant une plus grande opportunité de capturer un actif et donc de créer un coût pour l'entreprise.

Ces différences permettent en partie d'expliquer les variations de forme observées sur la figure 3.16. Par exemple, le fait que le graphe d'attaque de l'entreprise 20 soit plus étendu en longueur par rapport à celui de l'entreprise 26 complique l'accès aux serveurs et ordinateurs pour l'attaquant, ce qui explique les 400 simulations aboutissant à un coût nul. De plus, si l'on observe l'actif A12 dans le graphe, celui-ci n'a que 50% de probabilité d'être compromis, ce qui pourrait également expliquer pourquoi environ 400 simulations n'engendrent aucun coût.

Concernant l'entreprise 15, la petite taille de son graphe d'attaque explique le faible coût moyen observé. De plus, la faible probabilité de capturer l'actif A20 (seulement 28%) explique les 700 simulations ayant abouti à un coût nul.

Pour vérifier cette propriété sur l'ensemble du portefeuille, il est possible d'étudier le coût moyen des entreprises en fonction d'une mesure de la "complexité" du graphe. Il n'existe pas de mesure unique de complexité pour un graphe, chaque mesure captant un aspect spécifique de cette complexité. Ainsi, plusieurs mesures ont été testées pour évaluer leur impact sur le coût en pourcentage du chiffre d'affaires associé à cette complexité. Les résultats montrent que, pour chacune des mesures, la dépendance semble être linéaire. Nous avons donc appliqué une régression linéaire simple pour chacune d'entre elles, et nous avons observé les valeurs de R^2. Le tableau 3.4 résume les différentes mesures testées ainsi que les valeurs de R^2 associées.

Certaines mesures, telles que la largeur maximale et l'élargissement du graphe, ont été développées pour tester les idées évoquées plus tôt dans cette partie, notamment l'impact de la largeur du graphe sur le coût. L'élargissement du graphe, quant à lui, évalue la rapidité avec laquelle le graphe s'étend ; par exemple, le graphe de l'entreprise ENT26Com s'étend plus rapidement à l'horizontale que celui de ENT20Com. Pour ce faire, la largeur du graphe est mesurée à chaque niveau (c'est-à-dire le nombre de

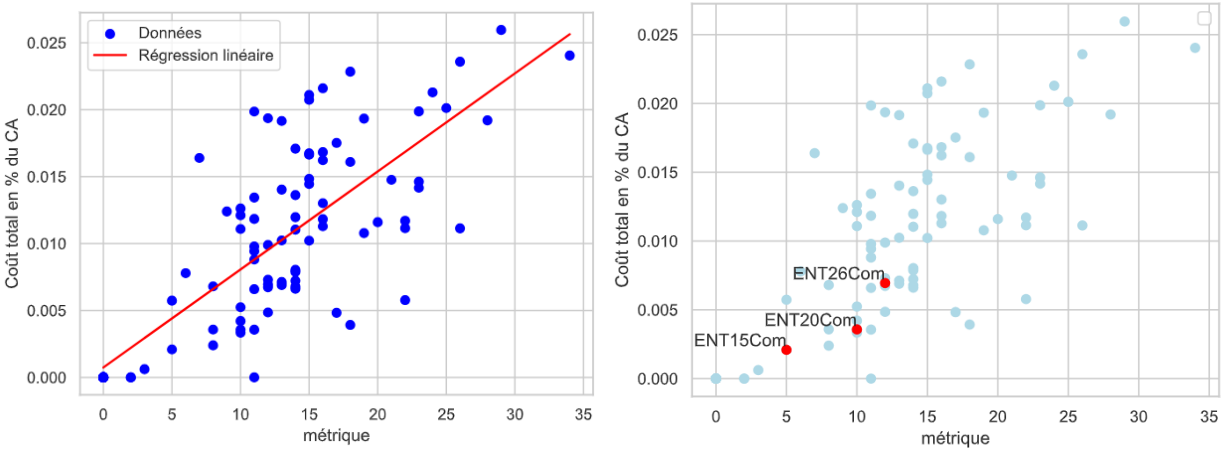
Métrique	Valeur	R^2
Nombre de nœuds	n	0.603
Nombre d'arêtes	a	0.598
Somme	$n + a$	0.600
Longueur du graphe	l	0.595
Largeur maximale du graphe	L_{max}	0.421
Élargissement du graphe (*)	e_g	0.565

Table 3.4: Tableau des métriques et R^2 associé

nœuds à chaque profondeur) et une moyenne est calculée en pondérant davantage les premiers niveaux que les suivants. Cela permet d'obtenir une notion de "vitesse" d'expansion du graphe en longueur. Toutefois, comme le montre le tableau, cette mesure sous-performe par rapport à la simple évaluation du nombre total de nœuds.

Une fois ces différentes mesures de complexité définies, une régression linéaire multiple par sélection exhaustive a été effectuée dans le but d'obtenir la meilleure combinaison possible. Toutes les combinaisons de métriques ont alors été testées.

À l'issue de l'analyse, la combinaison ayant obtenu les meilleures performances est celle qui combine la longueur et la largeur maximale du graphe avec un $R^2_{adj} = 0.652$ (à comparer avec les valeurs du tableau 3.4) et des p-valeurs extrêmement significatives pour les deux paramètres. En effet, cette approche intègre les éléments discutés dans le paragraphe précédent, soulignant l'importance de la longueur et de la largeur du graphe. La figure 3.18a illustre les résultats de la régression linéaire utilisant cette métrique.



(a) Régression linéaire sur la métrique choisie et le pourcentage de CA

(b) Position des entreprises 26, 20 et 15 sur le graphique

Figure 3.18: Présentation du chiffre d'affaires en pourcentage en fonction de la métrique choisie

La relation linéaire est clairement observable, bien qu'elle ne soit pas parfaite du fait de divers facteurs pouvant influencer le coût pour l'entreprise, comme les métriques CVE spécifiques des failles. La figure 3.18b illustre la position des entreprises précédemment étudiées sur le graphique. L'ordre des entreprises correspond à nos attentes, avec l'entreprise *ENT15Com* se situant en bas sur les deux dimensions, tandis que l'entreprise *ENT26Com* se positionne au sommet.

Ce résultat de régression linéaire est particulièrement intéressant s'il peut être appliqué à une

base de graphes d'attaque réels. En effet, cela permettrait de proposer une première estimation du risque pour une entreprise en se basant uniquement sur une analyse topologique de son graphe d'attaque. De plus, l'observation selon laquelle le risque augmente avec la largeur du graphe fait écho aux résultats dans le domaine de la cybersécurité, où la propagation latérale des attaques constitue une problématique majeure (Rieß-Marchive, 2022).

Cette section atteste que le modèle est effectivement capable d'estimer le risque sans nécessiter d'historique de sinistres en observant le graphe d'attaque. Les résultats obtenus semblent cohérents avec la taille et la topologie du graphe bayésien d'attaque, comme le montre la régression linéaire. Cette capacité à évaluer le risque à partir du graphe d'attaque répond à la problématique du manque de données fiables sur les sinistres dans le domaine de la cyberassurance, tel qu'évoqué dans la section 1.2.4.

Limites du portefeuille fictif L'analyse des résultats obtenus pour les différentes entreprises du portefeuille fictif a révélé que 20% d'entre elles affichent un coût moyen par attaque de 0 €, ce qui indique une protection théorique parfaite contre la perte d'exploitation. Cependant, comme discuté tout au long du Chapitre 1, un risque nul n'existe pas. Par conséquent, obtenir un coût moyen nul pour certaines entreprises n'est pas un objectif souhaité pour notre modèle. Même une entreprise résiliente sur le plan cyber, avec une architecture réseau sécurisée et peu de vulnérabilités, devrait afficher un coût moyen non nul par tentative d'attaque.

Dans ce mémoire, l'objectif a été de concevoir des graphes d'attaque cohérents, tout en intégrant une part d'aléatoire et de diversité. Cependant, les hypothèses sur lesquelles repose ce graphe ont conduit à une représentation beaucoup plus "simple" que celle d'un graphe d'attaque réel. Par conséquent, la surface d'attaque de l'attaquant est souvent réduite à un seul actif au temps $t = 0$. Si cet actif ne présente aucune faille lors de la simulation, l'entreprise peut apparaître comme incapturable, indépendamment de la sécurité des autres actifs. Cela crée une situation où l'entreprise semble être un mur infranchissable. Tenter de résoudre ce problème entraînerait une complexification des hypothèses et n'apporterait en réalité pas plus d'informations pour ce mémoire que la création du graphe d'attaque actuel.

Ce type de problématique illustre la sensibilité du modèle au graphe d'attaque. Cependant, comme discuté dans la section 3.1.3.1, il existe des méthodes automatiques pour générer ce type de graphe, et ce problème est alors limité à ce mémoire et ne résulte pas d'une limite méthodologique.

3.2.1.3 Analyse des résultats à la maille du portefeuille

Dans cette partie, les résultats du modèle seront analysés à l'échelle de l'ensemble du portefeuille. Cela permettra d'obtenir une vue d'ensemble des résultats fournis par le modèle en matière de tarification, ainsi que d'examiner les différences entre les secteurs et l'influence du graphe d'impact. Comme aucune connaissance préalable n'a été développée sur $\mathbb{E}(F_i)$, les résultats se concentreront, comme précédemment, uniquement sur la partie coût.

Dans un premier temps, examinons le comportement du coût au niveau du portefeuille et selon les différents secteurs. La figure 3.19 présente plusieurs boîtes à moustaches à partir desquelles il est possible de tirer des informations clés.

La somme totale des coûts moyens pour le portefeuille s'élève à **416 513 €**, ce qui correspond à un coût moyen par entreprise de 4 165,13 €. Cependant, la figure 3.19 montre une grande disparité entre les secteurs du commerce et de l'industrie, tant en termes de répartition (plus étendue pour le secteur de l'industrie) que de coût moyen (plus élevé pour le secteur de l'industrie). Le tableau 3.5

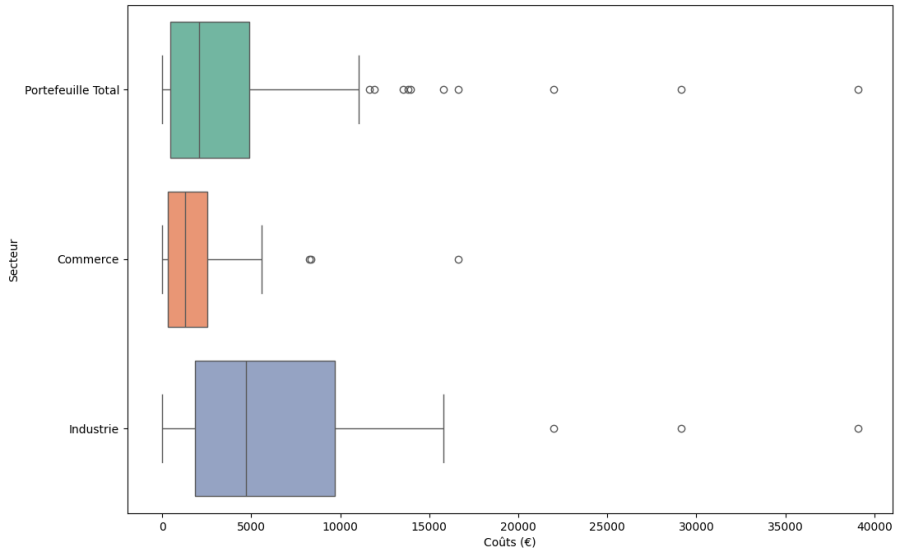


Figure 3.19: Présentation des coûts sur le portefeuille, les entreprises industrielles et commerciales

récapitule les valeurs clés.

	Coût total	**Nombre dans la base**	**Coût moyen**
Portefeuille total	416 513 €	100	4 165,13 €
Commerce	113 846,73 €	54	2 108,27 €
Industrie	302 666,25 €	46	6 579,70 €

Table 3.5: Récapitulatif des coûts du portefeuille

Cette analyse semble suggérer que le secteur de l'industrie est plus sensible à la perte d'exploitation cyber (en supposant qu'une attaque ait lieu) que le secteur du commerce. Cependant, cette conclusion n'est pas entièrement correcte. Le secteur de l'industrie pourrait inclure des entreprises de taille plus grande ou ayant des chiffres d'affaires plus élevés, ce qui pourrait biaiser les résultats. En effet, la figure 3.12c montre une répartition plus élevée du chiffre d'affaires par employé dans le secteur industriel, tandis que la figure 3.12b indique une répartition similaire du nombre d'employés entre les deux secteurs. Cela suggère qu'il pourrait y avoir un biais dans cette analyse.

Pour éviter de biaiser la conclusion, il est possible d'examiner la prime en fonction du pourcentage du chiffre d'affaires. La figure 3.20 présente plusieurs boîtes à moustaches qui permettront de vérifier les conclusions obtenues.

La moyenne des coûts en pourcentage du chiffre d'affaires sur le portefeuille est de **0.946%**. Néanmoins, la figure 3.20 présente encore une dissimilarité entre le secteur du commerce et celui de l'industrie, tant sur la répartition que sur la valeur moyenne. Le tableau 3.6 récapitule les valeurs clés.

	Pourcentage du CA moyen
Portefeuille total	0.946%
Commerce	0.62%
Industrie	1.32%

Table 3.6: Récapitulatif des coûts d'une attaque sur le portefeuille en pourcentage du chiffre d'affaires

L'industrie se révèle donc plus sensible que le commerce aux pertes d'exploitation causées par une

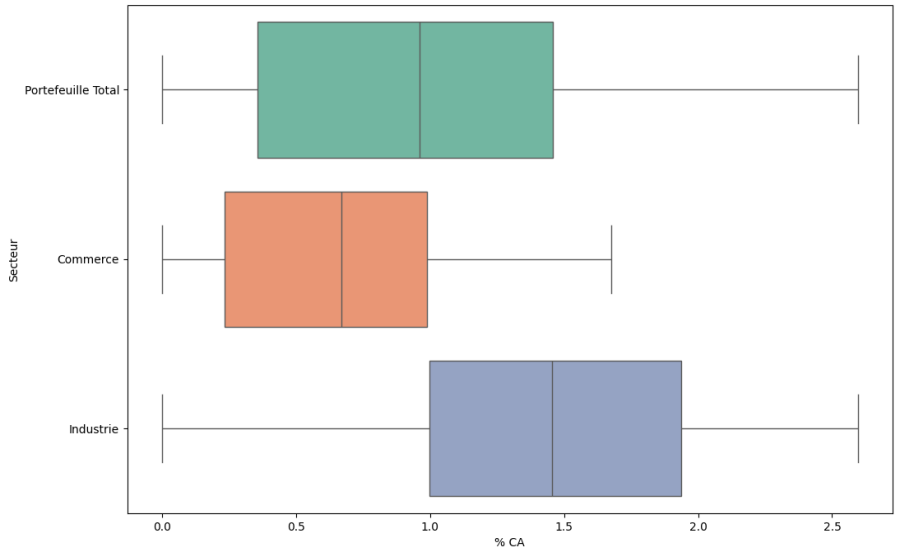


Figure 3.20: Présentation des coûts en pourcentage du chiffre d'affaires sur le portefeuille, les entreprises industrielles et commerciales

cyberattaque (toujours en supposant qu'une attaque ait lieu). Le fait que le pourcentage moyen du chiffre d'affaires soit plus élevé dans ce secteur, ainsi que le décalage des boîtes à moustaches vers des valeurs plus élevées, indique qu'il s'agit d'une tendance générale parmi les entreprises de chaque secteur, et non d'une influence exercée par certains *outliers* susceptibles de tirer les moyennes vers le bas ou vers le haut.

Les deux secteurs suivent la même méthodologie pour la création des graphes d'attaque. Autrement dit, seul le graphe d'impact différencie les deux secteurs dans la méthodologie de création du portefeuille. Cette analyse montre ainsi que le graphe d'impact influence effectivement les résultats du modèle, ce qui était prévisible compte tenu de la fonction assignée à ce graphe.

Discussion sur les valeurs obtenues La question peut se poser quant au réalisme des résultats du modèle avec les paramètres choisis et le portefeuille fictif. En particulier, il est possible de comparer les valeurs obtenues à des résultats réels afin de comprendre si l'évaluation est du bon ordre de grandeur.

Selon AMRAE, 2024, le taux de prime pour les PME en France varie entre 0,1% et 0,6%, les entreprises de taille moyenne ayant généralement un taux de prime plus élevé que les petites.

Cependant, le taux de 0.946% obtenu pour le portefeuille n'est pas directement comparable à ces valeurs. D'une part, le taux de prime est calculé en fonction de la somme assurée, tandis que notre pourcentage est basé sur le chiffre d'affaires de l'entreprise. On pourrait toutefois supposer que la compagnie d'assurance simulée fixe le chiffre d'affaires de l'entreprise comme limite de couverture (ce qui ne changerait pas le résultat, aucun sinistre n'ayant dépassé la valeur du chiffre d'affaires de l'entreprise sur notre portefeuille).

En plus de cette première hypothèse, il est important de rappeler que la valeur de 0.946% ne prend pas en compte la fréquence d'attaque $\mathbb{E}(F_i)$. Pour obtenir une approximation, on peut se référer au baromètre de Cesin, 2024, qui indique qu'en 2022, 45% des entreprises sondées ont subi une cyberattaque significative. Par conséquent, pour les besoins de cette analyse, nous pouvons estimer $\mathbb{E}(F_i) = 0.45$. En partant de ce postulat, le taux de prime pure moyen de notre portefeuille est de

$$\pi = 0.45 \times 0.946\% = 0.43\%.$$

Ce chiffre se situe dans les tranches réalistes proposées par les assurances cyber, selon l'étude de AMRAE, 2024.

Il est essentiel de préciser que ce résultat ne prouve en aucun cas l'adéquation du modèle à la réalité. D'autres paramètres, tels que le S/P des contrats ou l'ensemble des garanties offertes par les contrats de l'étude LUCY, devraient également être pris en compte pour tirer des conclusions plus approfondies. Cependant, ces chiffres fournissent un comparatif permettant d'évaluer la cohérence de futures études avec une meilleure connaissance de $\mathbb{E}(F_i)$.

Cette section a permis d'analyser les résultats de l'application du modèle au portefeuille. À l'échelle des entreprises, la sensibilité du modèle à différents paramètres a été étudiée, en discutant de la tarification individuelle et en examinant l'influence de la topologie du graphe d'attaque sur les résultats. À l'échelle du portefeuille, les coûts moyens modélisés, l'influence du secteur et du graphe d'impact, ainsi que la cohérence des résultats obtenus par rapport au marché ont été analysés.

3.2.2 Étude de l'évolution temporelle des résultats du modèle

Dans le chapitre 1, le caractère évolutif du risque cyber a été largement discuté. Le besoin de suivre continuellement ce risque, mentionné dans la sous-section 1.2.4, découle directement de cette réalité. En particulier, l'évolution des vulnérabilités au sein des systèmes informatiques entraîne des changements dans le niveau de risque pour l'entreprise, qui ne seraient pas pris en compte par un modèle d'évaluation se fondant uniquement sur les sinistres passés. Par exemple, il a été évoqué dans la sous-section 3.2.1.1 la faille EternalBlue, qui a permis l'attaque WannaCry, et dont l'apparition (et exploitation) a considérablement modifié le risque des entreprises aux systèmes vulnérables à celle-ci.

L'objectif de cette section est de démontrer que le modèle présenté est "dynamique", c'est-à-dire capable d'évaluer les variations du risque du portefeuille en fonction de l'évolution des menaces extérieures. Pour cela, une faille critique fictive sera introduite dans les actifs du portefeuille, et l'évolution du coût (en supposant qu'une attaque a lieu) sera analysée. Deux types d'actifs seront ciblés successivement, afin de montrer que l'impact d'une faille dépend également de la nature de l'actif et de son rôle au sein du réseau de l'entreprise.

Pour simuler l'évolutivité des menaces, le portefeuille fictif (et en particulier le graphe d'attaque de chaque entreprise) sera considéré comme issu des informations collectées par l'assureur auprès de ses assurés au temps $t_1 = 0$. Au temps t_2, une faille critique est découverte sur une version spécifique d'un type d'actif. Pour cette étude, l'impact d'une faille touchant Windows 10 sera comparé à celui d'une faille affectant les firewalls des marques Cisco et Palo Alto. Le choix de Cisco et Palo Alto permet d'obtenir une proportion d'actifs touchés similaire à celle de Windows, facilitant ainsi la comparaison. La faille fictive (illustrée à la figure 3.21) est donc aléatoirement ajoutée à 67,59% des actifs concernés, ce qui correspond à la proportion d'appareils sous Windows 10 parmi les PC sous Windows en 2024 (Caillebotte, 2024). Le graphe d'attaque est ensuite mis à jour pour chaque actif, en suivant la méthodologie développée dans la section 3.1.3.3.

Nom	CVE-0000-00000
Score CVSS	9.8
Vecteur	CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:H/I:H/A:H

Figure 3.21: Faille découverte sur les actifs

3.2.2.1 Faille sur les *Firewalls*

Lors de la construction de notre portefeuille fictif, et plus particulièrement lors de la création du réseau de l'entreprise (cf. section 3.1.3.2), il a été attribué au *firewall* un rôle central. Il protège les connexions et permet d'isoler l'ensemble de l'entreprise, ainsi que certaines de ses parties, des menaces extérieures. Il est donc attendu que l'ajout de cette vulnérabilité ait un impact notable sur la quantification du risque dans le portefeuille.

En appliquant le modèle au nouveau portefeuille, un coût total moyen de **447 093,17 €** est obtenu, soit une augmentation moyenne de 7,34% pour l'ensemble du portefeuille. Cette hausse est légèrement plus marquée dans le secteur du commerce, avec une augmentation moyenne de 9%, contre 6,7% dans le secteur de l'industrie. Cependant, comme l'illustre la figure 3.22, l'impact du secteur sur cette variation en pourcentage semble relativement faible.

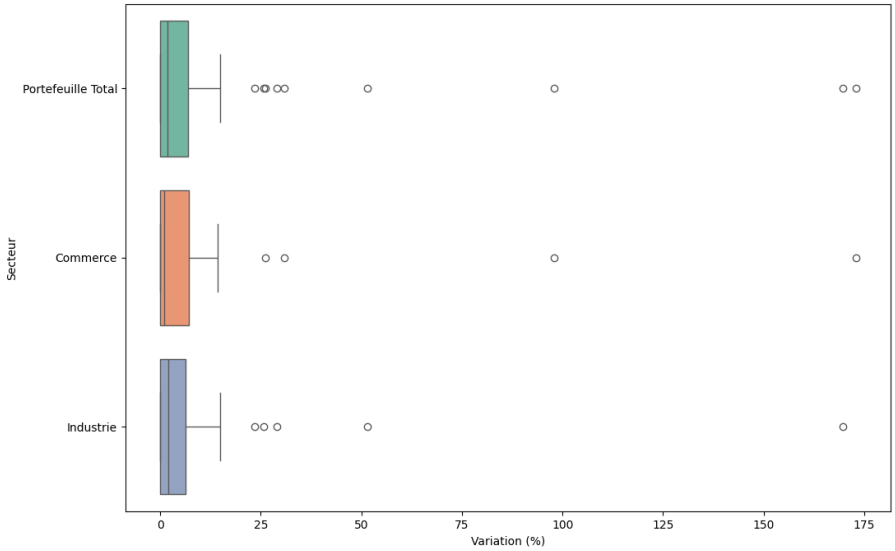


Figure 3.22: Boîtes à moustaches des variations des coûts moyens en pourcentage

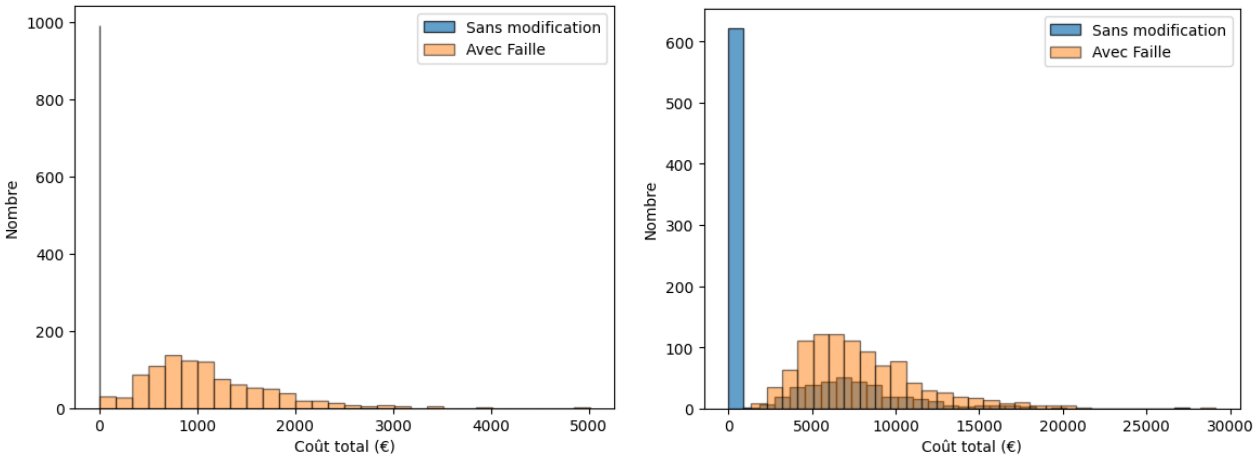
Nous pouvons également consulter le tableau 3.7 pour observer les cinq plus fortes augmentations en pourcentage au sein du portefeuille.

Nom	Augmentation (%)
ENT62Ind	inf %
ENT79Com	172,86 %
ENT88Ind	169,68 %
ENT53Com	98,00 %
ENT97Ind	51,49 %

Table 3.7: Plus grandes augmentations du coût moyen après l'apparition de la faille critique dans le portefeuille

Notamment, la première ligne présente une augmentation infinie en pourcentage. Pour expliquer cette hausse, il est possible de se référer à la figure 3.23a, qui montre les distributions des coûts avant et après l'apparition de la faille. À t_1, l'entreprise n'encourait aucun coût lié à la perte d'exploitation cyber (ce type de situation a été discuté dans la sous-section 3.2.1.2). L'introduction de la faille au temps t_2 a créé une brèche dans cette protection, révélant ainsi l'ensemble des vulnérabilités jusque-là

cachées dans le réseau. Cette notion d'exploitation successive des failles a par ailleurs été développée dans la sous-section 1.1.3.1.



(a) Comparaison des histogrammes de coûts pour l'entreprise *ENT62Ind*

(b) Comparaison des histogrammes de coûts pour l'entreprise *ENT79Com*

Figure 3.23: Comparaison de l'évolution des coûts pour les deux entreprises ayant les plus fortes augmentations

Un autre exemple de forte augmentation en pourcentage est illustré par la figure 3.23b, qui présente l'entreprise *ENT79Com* avec une augmentation moyenne de 172% des coûts. Au temps t_1, la proportion de coûts nuls représentait près de 60% des cas, elle est quasiment nulle à t_2. Pour comprendre cette variation, il faut examiner le graphe d'attaque bayésien inféré et noter que l'actif A45 (un pare-feu) présentait des failles avec une faible probabilité d'exploitation, ce qui entraînait l'arrêt de l'attaque dès la première étape (l'attaquant ne pouvait pas franchir le premier pare-feu). Avec l'ajout de la nouvelle vulnérabilité au temps t_2, cette probabilité a augmenté, permettant une plus grande pénétrabilité de l'entreprise et, par conséquent, une hausse significative du coût moyen.

3.2.2.2 Faille sur les PC

À titre de comparaison, et pour illustrer l'influence du type d'actif sur l'augmentation du coût moyen, l'expérience a été réitérée en utilisant une faille identique, mais cette fois sur un système Windows (c'est-à-dire sur un actif PC).

En appliquant le modèle au portefeuille modifié, un coût total moyen de **421 657 €** est obtenu, ce qui correspond à une augmentation de seulement 1,23%. Rappelons qu'avec la faille sur le *firewall*, l'augmentation était de plus de 7%. Cette différence est due à la position moins centrale des ordinateurs dans notre simulation de réseau. Étant des actifs terminaux, ils ne sont pas affectés par des scénarios comme ceux illustrés dans les figures 3.23, où la découverte d'une faille expose d'autres actifs. Ces situations ne peuvent pas se produire avec la topologie de réseau définie dans la partie 3.1.3.2.

En conclusion, cette section a démontré la capacité du modèle à s'adapter dynamiquement aux variations des risques liés à l'évolution des menaces extérieures. L'ajout de failles critiques fictives, que ce soit sur des *firewalls* ou des systèmes Windows, a permis d'illustrer l'impact différencié en fonction du type et du rôle des actifs au sein d'une entreprise. La position des actifs dans le réseau influence les résultats du modèle et donc fortement l'augmentation du coût moyen, avec des hausses significativement plus élevées pour les failles touchant des actifs centraux comme les *firewalls*.

Cette capacité est un élément clé dans le contexte de l'assurance cyber, notamment en ce qui concerne la réduction des menaces et la prévention, comme évoqué dans la section 1.2.4. Ce sujet sera approfondi dans la prochaine section.

3.2.3 Implication de l'assureur dans la diminution du risque

Lors des entretiens avec les dirigeants de PME, un point clé a été soulevé à plusieurs reprises. Comme discuté dans la sous-section 3.1.4.2, les dirigeants interrogés ont exprimé un manque flagrant de connaissances en matière de cybersécurité au sein de leurs entreprises. Ce point avait aussi été abordé de manière plus globale dans la section 1.1.4. Souvent, même pour des entreprises fortement dépendantes du numérique, comme l'entreprise du secteur industriel interrogée, la gestion de la cybersécurité est déléguée à des prestataires externes. Ces prestataires n'interagissent pas directement avec les assureurs, ce qui oblige l'entreprise à jouer le rôle d'intermédiaire entre les aspects techniques du prestataire et les besoins de l'assurance en termes d'informations pour évaluer le risque.

Pour le dirigeant du secteur industriel, l'assurance cyber pour les PME bénéficierait grandement de la fourniture d'un service intégré combinant prestation technique en cybersécurité et assurance. Cela permettrait à l'assureur de garantir la fiabilité et la robustesse des entreprises qu'il assure, tout en ayant une meilleure connaissance du risque. Quant à l'assuré, il pourrait déléguer complètement cet aspect technique, souvent complexe et chronophage.

L'aspect conseil offert par l'assureur aux assurés est également largement promu dans le domaine de la cyberassurance comme un moyen efficace de réduire le risque (Hillairet and Lopez, 2022). Comme mentionné dans la section 1.2.4, le besoin d'une quantification continue du risque est également un point clé dans cette démarche de prévention. Cette section a donc pour objectif de discuter de l'applicabilité du modèle dans ce contexte de prévention (et donc de diminution du risque), tout en réfléchissant à la rentabilité de cette démarche pour l'assureur.

En effet, le modèle développé et les méthodologies mises en place peuvent faciliter l'accompagnement des assurés par l'assureur. Tout d'abord, l'assureur peut suivre l'évolution du risque encouru par l'assuré au fil du temps, comme discuté dans la section 3.2.2 dans le cas de l'apparition d'une faille critique. Lorsqu'une vulnérabilité est identifiée, l'assureur pourrait avertir les assurés à risque et, en s'appuyant sur les bases de données CVE, partager des méthodes de remédiation, lorsque celles-ci existent. Ce processus, pouvant être automatisé, aurait un coût minimal pour l'assureur tout en apportant une valeur de conseil appréciable à ses assurés. Par exemple, en prenant la valeur de $\mathbb{E}(F_i) = 0,45$, l'assureur pourrait réduire le risque global de son portefeuille de

$$(447093 - 416513) \times \mathbb{E}(F_i) = 13761 \text{ €}.$$

En outre, l'assureur disposerait également d'une estimation de l'impact financier des changements induits par l'apparition d'une nouvelle faille. Cela lui permettrait de prioriser les assurés les plus exposés, comme les entreprises *ENT62Ind* et *ENT79Com* dans le tableau 3.7, dont le risque augmente drastiquement avec l'apparition de la nouvelle faille. Il pourrait alors offrir un accompagnement personnalisé à ces assurés afin de réduire rapidement leur exposition au risque.

Ce type de service, en plus d'être attractif pour les PME qui bénéficieraient d'un soutien accru, pourrait également se révéler rentable pour l'assureur. Celui-ci pourrait optimiser l'allocation d'un budget en fonction de l'augmentation du risque provoquée par une nouvelle faille et maximiser ainsi la réduction de risque sur l'ensemble du portefeuille grâce à des actions de mitigation ciblées (comme évoqué dans le cas d'une faille critique dans la section 3.2.2). Si la réduction du risque obtenue est supérieure au montant investi, l'assureur réalise une opération strictement gagnante. Il diminue le

risque de son portefeuille, améliore la stabilité de ses assurés et réalise mathématiquement un bénéfice sur l'opération.

L'annexe A.3 discute plus en détail de ce concept en posant et fournissant une application de ce problème d'optimisation (avec un exemple sur le portefeuille).

3.2.4 Étude d'une perte systémique sur le portefeuille

Pour conclure cette section sur l'application du modèle et la quantification dynamique dans le cadre du portefeuille de PME, cette sous-section s'intéressera à un autre domaine d'application du modèle.

Dans la sous-section 1.2.3.2, la composante systémique du risque cyber avait été abordée. Cette dimension est particulièrement pertinente dans le contexte actuel, où de nombreuses entreprises dépendent de services mutualisés comme le cloud. En effet, un nombre croissant de PME externalisent leurs infrastructures numériques vers des services de cloud computing. Selon FranceNum, 2024, en 2024, 61% des PME dépendent d'un serveur cloud externe pour l'installation de leurs solutions numériques, avec 39% de ces PME en dépendant entièrement.

Ce phénomène de concentration est amplifié par le fait que ces services sont dominés par un nombre restreint d'acteurs majeurs, tels que Microsoft, Amazon et IBM. Par exemple, pour les infrastructures IaaS (Infrastructure as a Service), Amazon détenait 47,8% des parts de marché en 2018, selon les données de Jones, 2023. Cette concentration peut entraîner un risque accru en cas de défaillance ou d'attaque ciblée contre l'un de ces fournisseurs, impactant simultanément plusieurs entreprises d'un même portefeuille. Cela augmenterait de manière significative le risque systémique pour l'assureur.

Il est possible d'utiliser le modèle développé (et en particulier le graphe d'impact) pour simuler un scénario de perte de fonctionnement des serveurs cloud utilisés par les structures du portefeuille. Une telle simulation permettrait à l'assureur d'anticiper les effets potentiels d'une défaillance commune à plusieurs entreprises et d'ajuster sa stratégie de gestion des risques et de tarification en conséquence.

Par exemple, l'assureur peut simuler une attaque majeure ciblant Amazon, entraînant une indisponibilité totale de ses services. Toutes les PME dépendantes d'Amazon subiraient une perte complète d'accès à leurs serveurs, paralysant ainsi une partie de leurs activités. En appliquant le **graphe d'impact** à ces entreprises, il serait possible d'évaluer la portée des perturbations causées par cette perte de disponibilité sur le portefeuille global.

Pour mettre en œuvre ce scénario de **stress test** sur notre portefeuille, les données évoquées précédemment sont utilisées. Nous sélectionnons 61% des PME du portefeuille, représentant la proportion de celles qui dépendent d'un serveur cloud externe. Parmi ces entreprises, 47,8% utilisent des serveurs hébergés par Amazon. En appliquant cette simulation, la perte **journalière** pour l'ensemble des entreprises touchées serait d'environ **9 500 €** en moyenne (en appliquant la perte d'opérabilité sur l'ensemble du portefeuille et en multipliant le résultat obtenu par la proportion choisie).

Avec les informations que l'assureur peut obtenir grâce au graphe d'impact appliqué à des stress tests comme celui-ci, il pourrait prendre des mesures pour mitiger le risque systémique en amont. Par exemple, l'assureur pourrait opter pour une répartition équilibrée de l'usage des fournisseurs de services cloud parmi les entreprises de son portefeuille. Cela permettrait d'éviter une concentration excessive des risques autour d'un seul prestataire majeur, comme Amazon. Une telle stratégie permettrait de diversifier les sources de risque et ainsi de protéger le portefeuille contre des pertes massives en cas d'incident affectant un fournisseur cloud dominant.

Cette section a permis de développer le cadre applicatif du modèle construit. Elle a mis en perspective les capacités du modèle à fournir une quantification dynamique de la perte en cas d'attaque

3.2.1.1, illustré la sensibilité du modèle aux différentes données d'entrée, telles que le graphe d'attaque ou le graphe d'impact 3.2.1.2, et présenté divers cas d'application qui pourraient aider l'assureur à offrir un service plus attractif pour les PME. En particulier, l'implication de l'assureur dans la diminution du risque a été discutée 3.2.3.

3.3 Pour aller plus loin

Au fil des différentes sections, certaines limites et observations ont été soulevées. Qu'elles concernent la méthodologie, les données ou encore le modèle lui-même, elles constituent une source de perspectives d'amélioration dans l'application des modèles graphiques dans un contexte cyberassurantiel. Cette section compile l'ensemble de ces limites et discute de certaines pistes permettant de lever certains verrous à l'avenir.

3.3.1 Autour du modèle

Si l'architecture du modèle a été détaillée dans la partie 2.3, de nombreux concepts restent à approfondir dans des travaux futurs. Cette sous-section récapitule les différentes limites développées tout au long de ce mémoire et propose des pistes de développement.

3.3.1.1 Estimer $\mathbb{E}(F)$

L'étude de la fréquence d'attaque F n'a pas été développée dans ce travail, car elle aurait nécessité un mémoire à part entière. Néanmoins, la place centrale de cet élément dans la modélisation rend son étude critique pour le bon fonctionnement du modèle. Dans la partie 2.3, le choix a été fait de considérer la fréquence d'attaque et le coût de l'attaque (exprimé par le modèle développé dans ce mémoire) comme indépendants. Il est ainsi possible de multiplier le coût moyen obtenu par l'application du modèle sur une entreprise i par l'espérance de cette fréquence, notée $\mathbb{E}(F_i)$. Ce qui distingue cette méthodologie d'un schéma "coût x fréquence" usuel est que l'intérêt ne porte pas sur un **sinistre**, mais bien sur une **attaque**. Ce déplacement de l'attention dans la modélisation est ce qui a permis de développer l'approche graphique proposée dans ce mémoire.

L'estimation de $\mathbb{E}(F)$ (ou de la loi de F en général) peut alors être effectuée en appliquant un modèle adapté au cyber-risque. En particulier, comme le montre la figure 1.6 du chapitre 1, cette fréquence semble dépendre du secteur de l'entreprise. De plus, le caractère systémique de cette fréquence (c'est-à-dire sa dépendance à des facteurs extérieurs variant dans le temps) a été discuté dans la section 1.2.3.2. Plusieurs sources indiquent une double influence des vulnérabilités, tant sur l'impact qu'elles ont sur l'entreprise (comme cela a été discuté tout au long de cette partie) que sur la volonté d'attaque de l'attaquant. Si des failles critiques existent à un instant t, l'attaquant sera davantage incité à exploiter cette vulnérabilité, une influence détaillée dans Hillairet and Lopez, 2022.

L'estimation de cette fréquence doit passer par l'utilisation de modèles capables de prendre en compte ces spécificités. L'utilisation de processus tels que les processus de Hawkes est alors envisageable (Boumezoued et al., 2023). Une modélisation par machine learning est aussi possible, comme le propose Laux et al., 2023, qui présente un score prédictif de risque d'attaque pour les entreprises.

L'estimation de la fréquence des attaques présente certains avantages par rapport à celle des sinistres, notamment en ce qui concerne la collecte de données pour la calibration de la loi de F. En effet, toutes les attaques ne se traduisent pas nécessairement par des dommages pour l'entreprise (comme cela est illustré par les coûts nuls dans notre modèle). Il est donc parfois plus facile et moins

sensible d'obtenir ce type d'informations, rendant l'assureur moins dépendant des déclarations de ses assurés. Ces données sont, par exemple, disponibles de manière très détaillée auprès d'entreprises qui fournissent des solutions antivirus ou des pare-feu. Ces entreprises disposent en temps réel de bases de données sur les tentatives d'attaque, comme le montre clairement le site de Fortinet. Il est également possible de simuler des entreprises fictives (des *honeypots*) pour observer leur attractivité pour les attaquants et ainsi obtenir plus d'informations sur la fréquence des attaques (Crowdstrike, 2022).

3.3.1.2 Choisir un ρ adapté

La sensibilité du modèle au taux de remédiation ρ a été discutée dans la sous-section 3.2.1.1. Ce taux est particulièrement important pour les résultats du modèle et sa calibration doit faire l'objet d'une étude approfondie.

En particulier, il est possible que cette variable varie en fonction du type d'entreprise, de sa résilience cyber et de sa taille. Cette valeur pourrait alors nécessiter une modélisation spécifique.

3.3.1.3 Adapter $p(e_i)$

Dans la partie 2.3, la variable aléatoire $p(e_i)$ a été définie comme représentant la probabilité d'exploitation d'une vulnérabilité. Dans ce mémoire, il a été choisi d'adopter une loi normale pour faciliter les calculs. La loi normale présente l'avantage de permettre une interprétation simple de la probabilité, en tant que paramètre de risque, sa forme symétrique évitant de donner une influence excessive aux petites ou grandes valeurs, comme indiqué dans Tatar et al., 2020.

Cependant, d'autres distributions pourraient être envisagées. Un choix pertinent serait la loi Beta, définie sur l'intervalle $[0, 1]$, qui serait ainsi naturellement adaptée à la modélisation d'une probabilité. Pour évaluer et comparer les différentes distributions candidates, des tests sur des données d'entraînement pourraient être réalisés. Cette procédure devrait être effectuée une fois que l'ensemble des autres paramètres du modèle aura été clairement défini.

3.3.2 Autour des données

Certaines limites concernant la récupération des données ont été évoquées tout au long de ce mémoire. Cette sous-section récapitule les différents points abordés.

3.3.2.1 Le graphe d'impact

Lors de la création du portefeuille fictif, des procédures réalistes pour la récupération du graphe d'impact ont été discutées (sous-section 3.1.4.3). Il a été noté que ce type de graphe restait un sujet particulièrement théorique et que peu d'applications avaient été réalisées, encore moins dans un contexte de cyberassurance.

Une recherche approfondie est donc nécessaire dans ce domaine afin de rendre l'obtention du graphe d'impact parfaitement applicable et robuste (Bahşi et al., 2018). Cette recherche pourrait inclure des entretiens à plus grande échelle avec des entreprises pour créer un questionnaire structuré. Elle pourrait également passer par le développement de méthodes automatiques de récupération de l'architecture, comme évoqué dans Bahşi et al., 2018.

En général, il serait bénéfique que ces méthodes soient davantage étudiées et connues pour permettre le développement de solutions comme le modèle de ce mémoire.

3.3.3 Des ajouts méthodologiques

Certains points méthodologiques pourraient être étudiés dans de futurs travaux afin de compléter et d'adapter le modèle à des cas plus complexes.

3.3.3.1 Ajout de vulnérabilités humaines

Dans la section 2.2.2.1, il a été mentionné la possibilité d'adapter les vulnérabilités humaines pour les traiter de manière similaire aux vulnérabilités techniques, afin de les intégrer dans le modèle d'évaluation du risque. Ce concept a été partiellement exploré par Tatar et al., 2020, qui a proposé la création de métriques basées sur des scores similaires au CVSS. Cependant, une analyse plus approfondie serait nécessaire pour que ces métriques puissent être utilisées efficacement dans le modèle.

Pour adapter ce concept, il faudrait établir une liste de métriques évaluant la fragilité du facteur humain dans l'entreprise. Par exemple, National Institute of Standards and Technology, 2018 présente un ensemble de métriques dans le cadre d'un framework visant à évaluer la résilience cyber d'une entreprise non seulement du point de vue technique, mais aussi organisationnel. Ces métriques pourraient être adaptées afin d'obtenir des scores similaires au CVSS et intégrées dans un futur graphe d'attaque représentant des vulnérabilités humaines, telles que des attaques ciblant directement les employés. Il serait alors nécessaire de repenser le graphe d'attaque pour inclure ces vulnérabilités humaines, bien que le modèle actuel resterait capable d'évaluer cette nouvelle dimension sans modifications profondes.

Il est important de souligner que ces vulnérabilités humaines n'ont pas été introduites dans le chapitre 3 en raison des hypothèses significatives qu'elles impliqueraient concernant le portefeuille fictif et les métriques utilisées. De plus, en raison de la méthodologie de création des graphes d'attaque établie pour le portefeuille fictif, il aurait fallu attribuer les vulnérabilités humaines directement aux ordinateurs, ce qui aurait transformé ces derniers en nœuds centraux plutôt que terminaux, rendant ainsi le graphe cyclique, complexifiant davantage le processus. Il convient néanmoins de noter que Wang et al., 2008 propose une méthode permettant de transformer un graphe d'attaque avec cycles en un graphe acyclique tout en conservant des propriétés similaires en termes d'attaque.

3.3.3.2 Des modèles complémentaires

Il a été vu dans la partie 1.1.2.3 que l'attaquant suit un processus d'attaque bien particulier. Lors de la création du modèle, cette méthodologie a été prise en compte.

Néanmoins, certaines modifications pourraient rendre le modèle encore plus adaptatif. Il serait, par exemple, possible de prendre en compte la réponse active du défenseur (ici l'assuré) lors de l'attaque. En effet, au lieu d'être considéré comme passif, le défenseur pourrait jouer un rôle actif dans la protection de son entreprise et dans la détection de l'attaque avant qu'elle ne soit menée à son terme. Ce rôle actif pourrait être modélisé par des méthodes mettant en scène deux acteurs au lieu d'un seul. Le modèle développé dans ce mémoire pourrait donc être complété par ce type de modélisation.

Ce type de modélisation à deux acteurs est d'ailleurs utilisé aujourd'hui dans la recherche d'un optimum d'investissement en cybersécurité pour les entreprises. C'est, par exemple, le cas de l'article de Awiszus et al., 2023, qui utilise la théorie des jeux pour analyser les interactions stratégiques entre des acteurs interconnectés. Des méthodes bayésiennes peuvent également être développées à ces fins et sont au centre de la recherche dans ce domaine (Wang and Neil, 2021). Elles pourraient être plus facilement applicables dans le cadre de la modélisation proposée. Une présentation de ce type de méthode est disponible en annexe A.1. Ces différents types de modèles offrent une vision **macro** du processus d'attaque, qu'il pourrait être intéressant de combiner avec la vision **micro** présentée.

Conclusion

Le risque cyber est aujourd'hui plus que jamais central dans la société. Un nombre croissant d'individus, d'entreprises et d'institutions dépendent du numérique pour faire fonctionner leurs activités et sont de plus en plus vulnérables aux cyberattaques. L'assurance cyber peut alors permettre aux entreprises de mieux gérer ce risque en offrant une protection financière contre les pertes liées aux incidents de cybersécurité, tels que les violations de données ou les interruptions de service.

Cette assurance en plein développement se heurte néanmoins aux spécificités du risque cyber. La forte évolutivité de ce risque appelle une évaluation dynamique, capable de s'adapter rapidement, alors que le manque de données de sinistre restreint l'analyse et rend difficile l'évaluation précise des menaces auxquelles les entreprises sont confrontées. De plus, l'assurance cyber, bien implantée chez les grandes entreprises, l'est beaucoup moins chez les plus petites structures qui sont néanmoins souvent les plus vulnérables. Différents acteurs appellent à une assurance plus élargie, capable de fournir des services de prévention pour augmenter l'attractivité de l'offre, en particulier pour les PME/ETI, tout en diminuant le risque porté par les assurés.

Le secteur de la cyberassurance n'est néanmoins pas le seul à s'intéresser à l'évaluation du risque cyber. Le domaine de la cybersécurité apporte également des connaissances techniques très poussées, notamment sur les méthodes d'attaque, l'analyse des vulnérabilités et la modélisation des réseaux et failles.

C'est dans ce contexte que nous avons étudié l'utilisation de données de cybersécurité afin d'améliorer la quantification du risque en cyberassurance. L'objectif était alors de fournir des pistes pour une quantification dépendant le moins possible des données de sinistre et pouvant permettre une évaluation dynamique du risque.

Pour ce faire, nous avons exploré les modèles graphiques, en particulier les graphes bayésiens. Nous avons analysé leur application dans un contexte cyber à travers les graphes bayésiens d'attaques. Parallèlement, nous avons examiné leur application dans le domaine de l'assurance en transformant les pertes sur le réseau en pertes économiques pour l'entreprise grâce aux graphes d'impact. Ces travaux nous ont permis de développer un modèle spécifique pour la perte d'exploitation cyber. Ce modèle permet d'évaluer le risque d'une entreprise en fonction de sa structure interne. Nous avons également apporté des ajustements pour rendre le modèle stochastique, afin d'obtenir des distributions de risques et nous avons défini un cadre précis pour cette évaluation.

Afin d'illustrer notre modélisation, nous avons mis l'accent sur les PME. Nous avons créé un portefeuille fictif et discuté de la récupération des différentes données nécessaires dans un cadre réaliste. La méthodologie de la création des graphes d'attaques a été détaillée et, grâce à des entretiens avec des professionnels, la création du graphe d'impact a pu être affinée.

Nous avons ensuite appliqué le modèle construit au portefeuille. Cela nous a permis d'étudier les résultats de ce modèle. Nous avons étudié la sensibilité du modèle aux données d'entrée, notamment les graphes d'attaques et d'impact. Cela a permis de vérifier la cohérence entre la mesure du risque

et le profil de l'entreprise, de montrer l'influence du secteur d'activité et de la structure du graphe d'attaques sur le risque porté par l'assuré, et de discuter de la possibilité de quantifier ce risque sans historique de coûts.

L'aspect dynamique du modèle a ensuite été illustré par l'étude de l'évolution de la quantification du portefeuille après l'apparition d'une faille critique. Cette étude nous a permis de développer une méthodologie de prévention en utilisant les résultats et le dynamisme du modèle. Cette méthodologie peut non seulement permettre de diminuer le risque pour l'assuré, mais pourrait s'avérer rentable pour l'assureur. Ainsi, ce type de modèle pourrait non seulement offrir une évaluation plus adaptée au risque cyber, mais aussi un accompagnement personnalisé des assurés, ce qui constituerait un atout pour conquérir le marché des PME.

La modélisation demande néanmoins d'être approfondie par d'autres travaux. Par exemple, la fréquence des attaques, non abordée dans ce mémoire, pourrait être estimée à l'aide de modèles statistiques et des données collectées via des *honeypots*. De plus, certains paramètres comme ρ mériteraient une étude plus poussée, voire une modélisation spécifique, compte tenu de leur importance dans le modèle. Les graphes d'impact, encore largement théoriques, nécessitent également des recherches supplémentaires pour être mieux appliqués dans la pratique.

Bibliography

Alay-Eddine, M. (Dec. 2022). EPSS : qu'est-ce que l'Exploit Prediction Scoring System ? URL: https://cyberwatch.fr/veille/epss-quest-ce-que-lexploit-prediction-scoring-system/.

AMRAE (2023). LUCY - Light Upon CYber insurance.

AMRAE (2024). LUCY - LUmière sur la CYberassurance.

ANSSI (2022a). La directive NIS. URL: https://cyber.gouv.fr/la-directive-nis.

ANSSI (July 2022b). Tendances - Les cybermenaces. URL: https://cyber.gouv.fr/tendances-les-cybermenaces.

ANSSI (2023a). La directive NIS 2. URL: https://cyber.gouv.fr/la-directive-nis-2.

ANSSI (2023b). PANORAMA DE LACYBERMENACE. Tech. rep. ANSSI. URL: https://cyber.gouv.fr/publications/panorama-de-la-cybermenace-2023.

AP news (2018). Uber agrees to $148M settlement with states over data breach. URL: https://apnews.com/article/a63faf22d9c94e2680f74bb355b486a7.

Awiszus, K., Bell, Y., L¨uttringhaus, J., Svindland, G., Voß, A., and Weber, S. (2023). Building Resilience in Cybersecurity–An Artificial Lab Approach.

Awiszus, K., Knispel, T., Penner, I., Svindland, G., Voß, A., and Webera, S. (2022). Modeling and Pricing Cyber Insurance – Idiosyncratic, Systematic, and Systemic Risks. *ArXiv*.

Axonius (n.d.). Assets in Cybersecurity. URL: https://www.axonius.com/blog/what-is-an-asset.

Bahşi, H., Udokwu, C. J., Tatar, U., and Norta, A. (2018). Impact Assessment of Cyber Actions on Missions or Business Processes: A Systematic Literature Review. Tech. rep. Department of Engineering Management & Systems Engineering, Old Dominion University, USA.

Baker, K. (May 2024). 12 MOST COMMON TYPES OF CYBERATTACKS. *Crowdstrike*.

Bessy-Roland, Y. (2019). Modélisation stochastique individuelle de sinistres cyber. MA thesis. Euria.

Boumezoued, A., Cherkaoui, Y., and Hillairet, C. (2023). Cyber risk modeling using a two-phase Hawkesprocess with external excitation.

Boydron, M.-H. (2018). WANNACRY SON HISTOIRE. *Cyber Cover*.

Caillebotte, E. (2024). Chiffres Microsoft : les statistiques à connaître en 2024. *BDM*.

Cambridge Centre for Risk Studies (2016). Managing cyber Insurance Accumulation Risk. *Centre For Risk Studies*.

Cesin (2024). Baromètre de la cybersécurité des entreprises - 9e edition. *Opinionway*.

Chockalingam, S., Pieters, W., Teixeira, A., and van Gelder, P. (Nov. 2017). Bayesian Network Models in Cyber Security: A Systematic Review. *Secure IT Systems*.

CNIL (2016). Le règlement général sur la protection des données - RGPD. URL: https://www.cnil.fr/fr/reglement-europeen-protection-donnees.

CNIL (2019). La loi Informatique et Libertés. URL: https://www.cnil.fr/fr/la-loi-informatique-et-libertes.

CNIL (2022). Stratégie européenne pour la donnée : la CNIL et ses homologues se prononcent sur le Data Governance Act et le Data Act. URL: https://www.cnil.fr/fr/strategie-europeenne-pour-la-donnee-la-cnil-et-ses-homologues-se-prononcent-sur-le-data-governance.

Coalition (2023). Cyber Threat Index 2023. Tech. rep. Coalition.

Cohen, D. (2023). Architecture réseau : Définition, typologies et sécurité. *DataScientest*.

Cooper, G. (1990). The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks. *Artificial Intelligence*. Ed. by Ar.

CrowdStrike (2021). CrowdStrike GlobalSecurity Attitude Survey.

Crowdstrike (2022). Qu'est-ce-qu'un Honeypot en Cybersécurité ? URL: https://www.crowdstrike.fr/cybersecurity-101/honeypots-in-cybersecurity-explained/.

CVE Program (2024). CVE Numbering Authority (CNA) Operational Rules. URL: https://www.cve.org/ResourcesSupport/AllResources/CNARules.

CyberInstitut (n.d.). LPM et Cybersécurité : Implications et Stratégies. URL: https://cyberinstitut.fr/lpm-et-cybersecurite-implications-et-strategies/.

CyberMalveillance (2022). Chiffres et tendances des cybermenaces : Cybermalveillance.gouv.fr dévoile son rapport d activité 2021. *CyberMalveillance*.

Cyentia Institute (Feb. 2023). The Evolving CVE Landscape. Tech. rep. F5 Labs. URL: https://www.f5.com/labs/articles/threat-intelligence/the-evolving-cve-landscape.

Dauxois, J.-Y. (2014). Cours de Probabilités. *Cours de Probabilités*.

Dreyer, P., Jones, T., Klima, K., Oberholtzer, J., Strong, A., Welburn, J. W., and Winkelman, Z. (2018). Estimating the Global Costof Cyber Risk - Methodology and Examples. Tech. rep. RAND Corporation.

EGE (2024). Comment créer une architecture réseau informatique. URL: https://www.ege.fr/infoguerre/comment-creer-une-architecture-reseau-informatique.

Euler, L. (1736). Solutio problematis ad geometriam situs pertinentis. *Mémoires de l'Académie des sciences de Berlin*.

FAIR Institute (n.d.). What is FAIR? URL: https://www.fairinstitute.org/what-is-fair.

FAIR Institute and EY (2024). Cybersecurity Risk Report.

FIRST (2023). Common Vulnerability Scoring System version 4.0: Specification Document. URL: https://www.first.org/cvss/v4.0/specification-document.

France Assureurs (Jan. 2023). 6e cartopgraphie prospective des risques. URL: https://www.franceassureurs.fr/espace-presse/les-communiques-de-presse/6e-cartographie-prospective/.

FranceNum (2024). Baromètre France Num 2024 : perception et usages du numérique par les TPE et PME. *FranceNum*.

FS Community (2022). Choisir un switch adapté pour petites entreprises. URL: https://community.fs.com/fr/article/how-to-choose-a-suitable-small-business-switch.html.

Garvey, P. R. (2009). An Analytical Framework and Model Formulation for Measuring An Analytical Framework and Model Formulation for Measuring Risk in Engineering Enterprise Systems: A Capability Portfolio Perspective. PhD thesis. Old Dominion University.

Giannotti, L. (2023). The biggest cyberattacks of 2023. *TechMonitor*.

Gomersall, M. (2024). 12 Key Network Types and their Assets. *vc4*.

Greenberg, A. (2018). The Untold Story of NotPetya, the Most Devastating Cyberattack in History. *Wired*.

Hillairet, C. and Lopez, O. (Jan. 2022). Cyber-assurance : enjeux, modélisations et leviers de mutualisation. *Opinions & Débats*.

INSEE (2016). Comprendre les résultats et les ratios comptables des entreprises réunionnaises.

INSEE (2024). Nomenclature d'activités française. URL: https://www.insee.fr/fr/information/2406147.

itaia (2023). Mettre en place un réseau informatique en entreprise. URL: https://www.itaia.fr/reseau-informatique-entreprise/.

Jacobs, J., Romanosky, S., Suciu, O., Edwards, B., and Sarabi, A. (2023). Enhancing Vulnerability Prioritization: Data-Driven Exploit Predictions with Community-Driven Insights. eprint: arXiv:2302.14172v2. URL: https://arxiv.org/abs/2302.14172.

Jakobson, G. (2011). Mission Cyber Security Situation Assessment Using Impact Dependency Graphs. *14th International Conference on Information Fusion*.

Jeamaon, A. and Khemapatapan, C. (2020). Cybersecurity Risk Assessment for Insurancein Thailand using Bayesian Network Model. *The 7th International Conference on Digital Arts, Media and Technology (DAMT)*.

Jones, E. (2023). Part de Marché du Cloud : Un Regard sur L'écosystème du Cloud en 2024. *Kinsta*.

Kanakogi, K., HironoriWashizaki, Fukazawa, Y., Ogata, S., Okubo, T., Kato, T., Kanuka, H., Hazeyama, A., and Yoshioka, N. (2021). Tracing CVE Vulnerability Information to CAPEC AttackPatterns Using Natural Language Processing Techniques. *MDPI*.

Kremp, E. and Sklénard, G. (2019). Productivité du travail et du capital : une mesure renouvelée au niveau de l'entreprise. *INSEE*.

Lau, P., Wang, L., Liu, Z., Wei, W., and Ten, C.-W. (2021). A Coalitional Cyber-Insurance Design ConsideringPower System Reliability and Cyber Vulnerability. *IEEE TRANSACTIONS ON POWER SYSTEMS*.

Laux, J., Anderson, J., Boni, M. A., Knapp, J., Bell, O., and Kao, A. (2023). Enhancing Risk Differentiation. *CyberCube*.

Legifrance (Dec. 2022). Journal officiel "Lois et Décrets" - n 0294. URL: https://www.legifrance.gouv.fr/jorf/jo/2022/12/20/0294.

Mensah, P. (2019). Generation and Dynamic Update of Attack Graphs in Cloud Providers Infrastructures. PhD thesis. CentraleSupélec. URL: https://theses.hal.science/tel-02416305.

Ministère de l'économie (2022). Assurance du risque cyber : publication du rapport de la direction générale du Trésor. URL: https://presse.economie.gouv.fr/07-09-2022-assurance-du-risque-cyber-publication-du-rapport-de-la-direction-generale-du-tresor/.

Ministère de l'économie (2023). ADOPTION DU DATA ACT AU CONSEIL DE L'UNION EUROPÉENNE. URL: https://www.entreprises.gouv.fr/fr/actualites/adoption-du-data-act-au-conseil-de-l-union-europeenne.

Ministère de l'économie des Finances et de la Souveraineté Industrielle et Numérique (2022). Risques cyber : des pistes pour la protection des entreprises. URL: https://www.economie.gouv.fr/risques-cyber-pistes-protection-entreprises.

MITRE (2019). About CAPEC. URL: https://capec.mitre.org/about/index.html.

Munich Re (2023). Cyber insurance: Risks and trends 2023. URL: https://www.munichre.com/landingpage/en/cyber-insurance-risks-and-trends-2023.html.

Murphy, K. P. (2002). Dynamic Bayesian Networks:Representation, Inference and Learning. PhD thesis. UNIVERSITY OF CALIFORNIA, BERKELEY.

National Institute of Standards and Technology (2018). Framework for Improving Critical Infrastructure Cybersecurity. *National Institute of Standards and Technology*.

Naïm, P., Wuillemin, P.-H., Leray, P., Pourret, O., and Becker, A. (2011). Réseaux bayésiens. Ed. by Editions Eyrolles. Editions Eyrolles.

OpenAI (2024). ChatGPT. Utilisé uniquement à des fins de reformulation. URL: https://chat.openai.com.

Pearl, J. (1985). Bayesian Networks : A model of self-activated memory for evidential reasoning. *Seventh Annual Conference of the Cognitive Science Society*.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems. Ed. by Kaufmann, M. Morgan Kaufmann.

Pearl, J. (2013). Causality. Ed. by Cambridge Univrsity Press. Cambridge Univrsity Press.

Peyrat, T. (2022). Risque cyber, un modèle épidémiologique sur réseaux pour le risque d'accumulation du cyber silencieux. MA thesis. Dauphine.

Phillips, C. and Swiler, L. P. (1998). A Graph-Based System for Network-Vulnerability Analysis. *NSPW '98*.

Ponts de Königsberg et cycle eulérien (n.d.). URL: https://www.bibmath.net/dico/index.php?action=affiche\&quoi=./p/pont.html.

Poolsappasit, N., Dewri, R., and Ray, I. (2012). Dynamic Security Risk Management Using Bayesian Attack Graphs. *IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING*.

Prevost, S. (2023). Firewall d'entreprise : back to basics. *stormshield*.

Pwc (2023). La réglementation DORA sur la résilience opérationnelle numérique. URL: https://www.pwc.fr/fr/expertises/gestion-des-risques/maitrise-des-risques-technologiques/dora-exigences-reglementaires-europeennes.html.

PyAgrum (n.d.). Inference. URL: https://pyagrum.readthedocs.io/en/1.15.1/BNInference.html.

Red Hat (Nov. 2021). Une CVE, qu'est-ce que c'est ? URL: https://www.redhat.com/fr/topics/security/what-is-cve.

Reddit (2019). Comment puis-je déterminer le nombre de routeurs dont un réseau d'entreprise « a besoin » ? URL: https://www.reddit.com/r/networking/comments/dmth5q/how_do_i_determine_how_many_routers_a_company/?tl=fr&rdt=54220.

RedHat (2023). CVE and CVSS explained — Security Detail. URL: https://www.youtube.com/watch?v=oSyEGkX6sX0.

Rieß-Marchive, V. (2022). Cyberattaques : en quoi consiste le déplacement latéral ? *LeMagIT*.

Rigo, M. (2009). Théorie des graphes. URL: http://www.discmath.ulg.ac.be/cours/main_graphes.pdf.

Rios, I., Rios, J., and Banks, D. (2009). Adversarial Risk Analysis. *Journal of the American Statistical Association* 104.486, pp. 841–954.

RiskLens (2021). The Six Types of Loss in Cyber Incidents. URL: https://www.risklens.com/resource-center/blog/the-six-types-of-loss-in-cyber-incidents.

Sachs, H., Stiebitz, M., and Wilson, R. J. (1988). An Historical Note: Euler's Konigsberg Letters. *Journal of Graph Theory*.

Sanguino, L. A. B. and Uetz, R. (2017). Software Vulnerability Analysis Using CPE and CVE. Tech. rep. Fraunhofer FKIE, Bonn, Germany.

Singhal, A. and Ou, X. (2011). Security Risk Analysis of Enterprise Networks Using Probabilistic Attack Graphs. *NIST Interagency Report 7788*.

Spacey, J. (2023). 41 examples of IT Assets. *Simplicable*.

Statista (2024a). Annual amount of monetary damage caused by reported cybercrime in the United States from 2001 to 2023. URL: https://www.statista.com/statistics/267132/total-damage-caused-by-by-cybercrime-in-the-us/.

Statista (2024b). Market share held by the leading computer (desktop/tablet/console) operating systems worldwide from January 2012 to February 2024. URL: https://www.statista.com/statistics/268237/global-market-share-held-by-operating-systems-since-2009/.

Suarez, J. N. and Salcedo, A. (2017). ID3 and k-means Based Methodology for Internet of Things Device Classification. *Conference: 2017 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE)*.

Tatar, U., Keskin, O., Bahsi, H., and Pinto, C. A. (2020). Quantification of Cyber Risk for Actuaries - An Economic-Functional Approach. *Society of Actuaries*.

Verizon (2023). DBIR2023 Data BreachInvestigations Report.

Wang, J. and Neil, M. (2021). A Bayesian-network-based cybersecurity adversarial risk analysis framework with numerical examples. URL: https://arxiv.org/abs/2106.00471.

Wang, J., Neil, M., and Fenton, N. (2020). A Bayesian Network Approach for Cybersecurity Risk Assessment Implementing and Extending the FAIR Model. *Computers & Security*.

Wang, L., Islam, T., Long, T., Singhal, A., and Jajodia, S. (2008). An Attack Graph-Based Probabilistic Security Metric. *IFIP Annual Conference on Data and Applications Security and Privacy*.

Wang, S. S. (2019). Integrated framework for information security investment andcyber insurance. *Pacific-Basin Finance Journal*.

wikipedia (n.d.). Colonial Pipeline ransomware attack. URL: https://en.wikipedia.org/wiki/Colonial_Pipeline_ransomware_attack.

Wikipedia (2024). Common Vulnerability Scoring System. URL: https://en.wikipedia.org/wiki/Common_Vulnerability_Scoring_System.

World Economic Forum (2024a). 2023 was a big year for cybercrime – here's how we can make our systems safer. URL: https://www.weforum.org/agenda/2024/01/cybersecurity-cybercrime-system-safety/.

World Economic Forum (2024b). Global CybersecurityOutlook 2024.

Xing, L., Guo, M., Liu, X., Wang, C., Wang, L., and Zhang, Y. (2017). An improved Bayesian network method for reconstructing gene regulatory network based on candidate auto selection. *BMC Genomics*.

Zhu, W. and Nguyen, N. L. C. (2022). Structure Learning for Hybrid Bayesian Networks. Tech. rep. Data61,ARC Centre for Data Analytics for Resources and Environments, The University of Sydney. URL: https://arxiv.org/pdf/2206.01356.

Appendix A

Compléments relatifs aux éléments présentés dans le mémoire

A.1 Modélisation à plusieurs agents dans la quantification cyber

A.1.1 Modèle D-A-D, une vision séquentielle du phénomène d'attaque-défense

Comme nous l'avons vu au cours du chapitre 1, une attaque cyber suit un déroulement bien précis (1.8). À la différence d'autres risques, le risque cyber (hors perte due à l'erreur) est un *jeu* entre deux acteurs : **l'attaquant**, dont l'objectif est d'atteindre ses fins, et **le défenseur**, dont l'objectif est de minimiser son risque (ses pertes). De plus, l'attaque est un processus séquentiel (MITRE ATT&CK), qui se déroule en plusieurs étapes où le défenseur et l'attaquant doivent faire des choix en fonction de leurs connaissances actuelles.

Compte tenu de cette problématique, une modélisation utilisant la **théorie des jeux** est une option judicieuse, puisqu'elle permet de modéliser l'interaction entre différents acteurs et possiblement de trouver l'optimum d'investissement pour l'assuré. C'est d'ailleurs l'approche adoptée par plusieurs travaux, comme (Awiszus et al., 2023), qui met en place le *Security Investment Game* pour analyser les interactions stratégiques entre des acteurs interconnectés. Néanmoins, comme mentionné dans (Wang and Neil, 2021), l'utilisation de la théorie des jeux nécessite la recherche d'un équilibre de Nash, et lorsque le modèle devient réaliste et complexe, il devient alors difficile de calculer une solution.

Une alternative est l'utilisation de l'**Analyse de Risque Adversarial** (ARA) (Rios et al., 2009), qui intègre également les comportements stratégiques des adversaires dans l'évaluation des risques. L'objectif de (Wang and Neil, 2021) est d'adapter ce type de cadre à la modélisation d'un réseau bayésien.

Le modèle ARA *Defense-Attack* La modélisation ARA repose sur l'utilisation d'un Diagramme d'Influence (DI), un type particulier de réseau bayésien hybride liant les décisions des deux parties, des variables aléatoires et des utilités.
Le document se place d'abord dans le cadre séquentiel Défense-Attaque (D-A), où le défenseur prend des décisions, puis l'attaquant réagit en conséquence. Ce type de modélisation permet de prendre en compte l'*hygiène cyber* et les particularités de l'entreprise, qui influencent (comme vu dans le chapitre 1) les choix de l'attaquant et, par conséquent, la perte finale.

Nous pouvons observer la visualisation du cadre D-A sur la figure A.1. Les différents éléments sont les suivants :

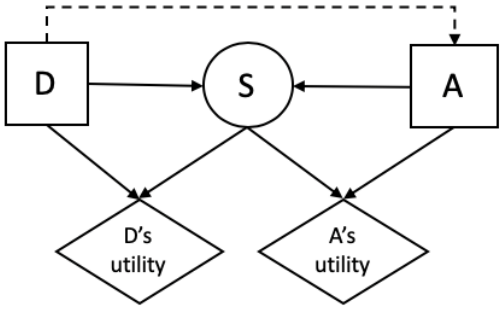


Figure A.1: Diagramme d'Influence pour un jeu D-A (Wang and Neil, 2021)

- D représente la variable aléatoire du choix du défenseur au temps $t = 0$,
- S représente la variable aléatoire du succès de l'attaque au temps $t = 0$,
- A représente la variable aléatoire du choix de l'attaquant au temps $t = 0+\delta$ (avec δ une variation infinitésimale permettant à l'attaquant de récupérer l'information),
- Les flèches pleines représentent des liens causaux (usuels dans les réseaux bayésiens),
- La flèche en pointillés représente une causalité "informationnelle" (possiblement partielle), permettant d'indiquer l'ordre temporel du graphe (ici, D prend sa décision avant A, donc A dispose en théorie de plus d'informations),
- Les trapèzes représentent les utilités de chaque partie.

La méthode ARA fournit un cadre pour identifier la stratégie de défense optimale. Tout d'abord, le problème est envisagé du point de vue de l'attaquant, dont l'objectif est de maximiser son utilité en fonction du choix du défenseur. Si nous posons, $\forall (d, a) \in D \times A$, alors

$$\Psi_A(d, a) = \mathbb{P}(S = 0|d, a)u_A(d, a, S = 0) + \mathbb{P}(S = 1|d, a)u_A(d, a, S = 1).$$

Alors, son choix optimal d'attaque en fonction de d est

$$a^*(d) = \arg\max_{a \in A} \Psi_A(d, a), \quad \forall d \in D.$$

Ensuite, du point de vue du défenseur, dont l'objectif est de maximiser son utilité en connaissant le choix optimal de l'attaquant pour chaque d. De la même manière, si nous notons

$$\Psi_D(d, a) = \mathbb{P}(S = 0|d, a)u_D(d, a, S = 0) + \mathbb{P}(S = 1|d, a)u_D(d, a, S = 1).$$

Alors, la deuxième partie de notre problème d'optimisation devient

$$d^* = \arg\max_{d \in D} \Psi_D(d, a^*(d)).$$

Pour résoudre ce problème de manière informatique, le document propose une représentation en réseau bayésien.

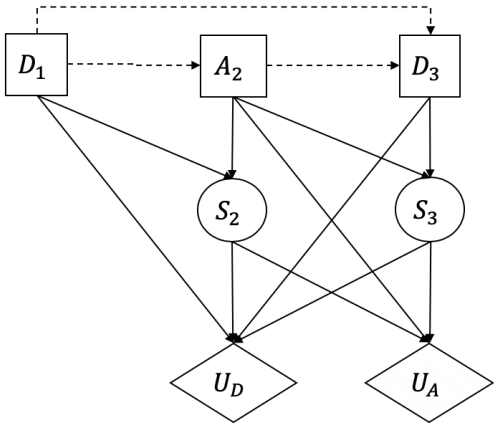


Figure A.2: Diagramme d'Influence pour un jeu D-A-D (Wang and Neil, 2021)

Le cadre séquentiel D-A-D L'article (Wang and Neil, 2021) étend cette modélisation à un processus en plusieurs étapes d'attaque et de défense. Le défenseur prend une décision à $t = 1$, l'attaquant répond à $t = 2$, et le défenseur prend une ultime décision à $t = 3$.
Cela se rapproche encore plus d'un processus d'attaque réel, où, après avoir observé l'attaque, le défenseur peut prendre des décisions de remédiation pour diminuer l'impact des prochaines attaques.

Comme nous le voyons sur la figure A.2, le processus est maintenant en plusieurs étapes et dispose de plusieurs succès. Le défenseur, en $t = 3$, peut prendre sa décision en fonction du choix de l'attaquant et de son propre choix précédent. Il est parfaitement envisageable d'étendre ce jeu à un horizon temporel plus long (Wang and Neil, 2021). Cette modélisation étendue peut être assimilée à un réseau bayésien dynamique.

A.2 Questionnaire : Création du graphe d'impact

Après avoir fait une introduction sur l'ensemble des nœuds d'un graphe d'impact et le fonctionnement des nœuds constituants

1. **Quels sont les nœuds que vous identifiez pour la couche business ? Quels sont les différents axes/missions générant du profit économique pour votre entreprise ?**

   Exemple de réponse : *Construire des véhicules, réparer des véhicules, ...*

2. **Comment lier les nœuds business entre eux selon la logique C, I et D ? Il est possible que certains nœuds soient indépendants et qu'aucun lien de dépendance n'existe.**

   Exemple de réponse : *Si la construction de véhicules diminue, le service après-vente aussi (Disponibilité → Disponibilité), ...*

   Nota : *Les logiques OR et AND sont totalement possibles ici, comme pour tous les liens demandés par la suite.*

3. **Si vous deviez diviser le chiffre d'affaires produit par votre entreprise, quelle proportion attribueriez-vous à chaque nœud business ?**

   Exemple de réponse : *Environ 85% pour la production de véhicules, 5% pour le service après-vente et 10% pour la réparation de véhicules, ...*

4. **Quels sont les services clés de votre entreprise ?**

   Exemple de réponse : *Service de maintenance industrielle, service de gestion de la production, service de contrôle qualité, ...*

5. **Comment lier les nœuds de la couche business avec ceux de la couche service selon une logique C, I et D ?**

   Exemple de réponse : *La disponibilité du service de gestion de la production influe sur la disponibilité de la création de véhicules.*

6. **Pouvez-vous établir des liens entre les différents services selon une logique C, I et D ?**

   Exemple de réponse : *La confidentialité du service de gestion de la production dépend de l'intégrité du service de contrôle qualité, ...*

7. **Quels sont, selon vous, les actifs les plus importants dans votre entreprise ? En particulier, si vous utilisez des serveurs, combien en utilisez-vous et pour quelles raisons ?**

   Exemple de réponse : *Les ordinateurs personnels, car ils permettent aux employés de travailler sur les problématiques, les serveurs, ...*

8. **Après avoir identifié les actifs dans la couche correspondante, pouvez-vous établir des liens entre eux (au niveau des nœuds constituants) ?**

   Exemple de réponse : *La disponibilité du serveur de données dépend de la disponibilité du routeur, ...*

9. **Établissez des liens entre la couche d'actifs précédemment créée et la couche de service.**

   Exemple de réponse : *La disponibilité du service de contrôle qualité dépend de la disponibilité du serveur de données.*

Après avoir introduit la quantification FDNA et les valeurs de α et β, et parcouru le premier questionnaire ci-dessus :

1. **Pour chaque lien identifié, pouvez-vous fournir une approximation des alphas (et bêta si existants) ?**

   Exemple de réponse : *Entre serveurs et ordinateurs personnels, sur le lien C→C :* $\alpha \approx 0.5$ *et* $\beta \approx 0.25$*, ...*

A.3 Optimisation de la prévention après apparition d'une faille

Dans cette annexe, les propos de la sous-section 3.2.3 sur l'implication de l'assureur dans la diminution du risque seront illustrés par un exemple. L'objectif est de montrer que, sous certaines hypothèses et contraintes, la connaissance des différentes augmentations de risque entre deux instants peut être utilisée pour optimiser l'investissement de l'assureur dans chacun de ses assurés et ainsi diminuer le risque de façon optimale.

Pour illustrer nos propos, le portefeuille fictif sera repris et l'apparition d'une faille critique sur les *firewalls* réutilisée (étudiée dans la partie 3.2.2.1). L'objectif est alors de montrer que **dans ce contexte**, et en supposant d'autres hypothèses sur la situation de l'assureur, le risque peut être diminué de manière optimale.

A.3.1 Le problème

Cette section a pour objectif de poser le problème, ses contraintes et ses hypothèses de la manière la plus complète possible.

Expression la plus globale Un assureur détient un portefeuille dont il connaît le risque au temps $t = 0$. Ce risque lui est acceptable et il est considéré qu'il n'y a pas de possibilité de prévention envers ses assurés pour le diminuer. Au temps $t = 1$, une faille critique est détectée sur différents *firewalls* (exactement la même situation que celle décrite dans la section 3.2.2.1, avec les mêmes proportions). Grâce au modèle, il est capable d'évaluer l'augmentation du risque pour chaque entreprise de son portefeuille. Les résultats de cette évaluation sont représentés par les tests du modèle de la partie 3.2.2.1. Il est également supposé que cette vulnérabilité est *patchable* (c'est-à-dire qu'il existe une procédure pour la corriger).

L'assureur souhaite alors diminuer le risque de son portefeuille de la manière la plus optimale possible. Plusieurs hypothèses sont alors nécessaires :

1. Il dispose d'un budget maximal de x €.

2. Le coût (en €) supplémentaire dû à la faille sera noté C_i pour l'assuré i et celui en pourcentage du chiffre d'affaires sera noté c_i.

3. Il existe une variable aléatoire R_i suivant une loi de Bernoulli, avec comme succès la réparation de la faille par l'entreprise i et comme échec l'absence de réparation. La probabilité de succès est notée p_i, qui est donc le paramètre de cette loi.

4. p_i dépend du budget x_i qui lui est alloué par l'assureur. Il existe donc une fonction $f_i : [0; +\infty] \to [0; 1]$ reliant x_i à p_i. Pour simplifier, la valeur pour chaque entreprise sera notée $p_i(x_i)$.

5. Le budget total B pour l'assureur doit vérifier la contrainte $B = \sum_i x_i \leq x$.

Il est ainsi possible de définir la variable aléatoire exprimant la réduction du risque (en €) à la fin de cette campagne de prévention. Cette variable, que nous nommerons M, s'exprime comme

$$M = \sum_i R_i \cdot C_i.$$

L'objectif est alors de maximiser $\mathbb{E}(M)$, or

$$\mathbb{E}(M) = \sum_i \mathbb{E}(R_i) \cdot C_i = \sum_i p_i(x_i) \cdot C_i.$$

A.3.2 Optimisation sous connaissance de f_i

Dans cette partie, il sera supposé que f_i pour chaque assuré est connu. L'objectif est ainsi de trouver le vecteur

$$\hat{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

qui maximise $\mathbb{E}(M)$ sous les contraintes $0 \leq \sum_i x_i \leq x$ et $x_i \geq 0$ pour tout i. Pour répondre au problème, il sera également supposé que les f_i sont des fonctions concaves et différentiables sur $[0; +\infty]$. Ainsi, la fonction objectif $\mathbb{E}(M)(.)$ est aussi concave et différentiable sur cet intervalle.

Le problème peut alors être formulé comme un problème d'optimisation sous contraintes pouvant être résolu à l'aide des conditions de Karush-Kuhn-Tucker (KKT). En notant $F(.) = \mathbb{E}(M)(.)$ pour plus de lisibilité, et puisque $F(.)$ est concave, alors $-F(.)$ est convexe. Les contraintes peuvent être exprimées sous la forme

$$\begin{cases} g_0(\mathbf{x}) = -x + \sum_i x_i \leq 0 \\ \forall i \in \{1, ..., n\}, \quad g_i(\mathbf{x}) = -x_i \leq 0 \end{cases}.$$

Ainsi, le Lagrangien peut être défini comme

$$\mathcal{L}(\mathbf{x}, \mu) = -F(\mathbf{x}) + \mu \mathbf{g}^T(\mathbf{x}).$$

En particulier, le gradient de g_i est indépendant des f_i et facilement calculable. En effet, $\nabla g_0 = (1, ..., 1)^T$ et $\forall i \in \{1, ..., n\}, \quad \nabla g_i = (0, ..., 0, 1, 0, ..., 0)^T$.

A.3.2.1 Exemple de connaissance de f_i

Cette sous-section fournira un exemple concret d'application des idées développées ci-dessus. Il sera supposé que f_i, la fonction reliant l'investissement de l'assureur pour une entreprise et la probabilité de remédiation de la faille par l'entreprise, est connue.

Il sera supposé que, sans l'aide de l'assureur, la probabilité de remédiation est égale à u_i (c'est-à-dire que lorsque $x_i = 0$, alors $p_i = u_i$). De plus, on suppose que $p_i \to 1$ lorsque $x_i \to +\infty$. Nous introduisons également l'hypothèse selon laquelle, pour un investissement m_i de la part de l'assureur, $p_i = 0.80$ (autrement dit, il existe une constante m_i telle que $f(m_i) = 0.80$).

Pour cet exemple, nous prendrons les fonctions f_i sous la forme :

$$f_i(x) = 1 - \frac{1}{m_i \cdot x + \frac{1}{1-u_i}}.$$

Les propriétés évoquées précédemment sont vérifiées pour ce type de fonctions.

A.3.2.2 Avec une unique valeur de f

Dans cette section, les résultats seront analysés avec $f_i = f$, c'est-à-dire que chaque entreprise partage la même fonction de transformation de l'investissement en probabilité de remédiation. De

plus, nous prendrons $u = 10\%$ et $m = 2000\,€$. La fonction f est donc de la forme

$$f(x) = 1 - \frac{1}{\frac{\frac{1}{0.2} - \frac{1}{0.9}}{2000} \cdot x + \frac{1}{0.9}}.$$

Ainsi, la dérivée de cette fonction est donnée par

$$f'(x) = -\frac{\frac{1}{0.2} - \frac{1}{0.9}}{2000} \cdot \frac{1}{\left(\frac{\frac{1}{0.2} - \frac{1}{0.9}}{2000} \cdot x + \frac{1}{0.9}\right)^2}.$$

Pour illustrer cet exemple, il est possible de reprendre le cas donné dans la section 3.2.2.1 et d'optimiser $\mathbb{E}(M)$ pour ce scénario.

Rentabilité de l'assureur Après avoir repris le tableau des augmentations de risque consécutives à l'apparition de la faille sur les *firewalls*, il est possible d'appliquer les conditions de Karush-Kuhn-Tucker (KKT) présentées précédemment, en utilisant le choix de f dans un programme Python.

L'objectif est ici de faire varier l'investissement de l'assureur x (de 0 € à 10 000 €, avec un pas de 100), d'appliquer l'optimisation pour chaque x, et ainsi d'identifier une valeur optimale. Cela permettra également d'observer l'évolution du résultat de la campagne de prévention "optimale" en fonction de la somme investie.

La figure A.3 présente le résultat obtenu en allouant le budget de manière optimale en fonction de l'investissement (en supposant ici encore $\mathbb{E}(F) = 0.45$). Pour rappel, l'augmentation de risque initiale était de 13 761 €.

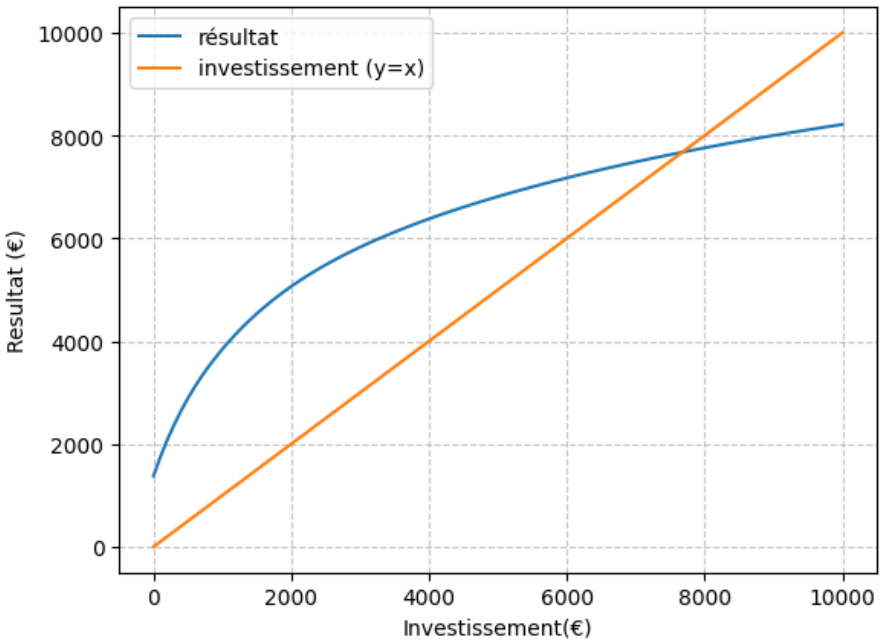


Figure A.3: Résultat de la campagne de prévention optimale en fonction de l'investissement

Plusieurs informations intéressantes ressortent de ces courbes :

1. Le résultat de la campagne "optimale" est concave. Plus l'investissement augmente, moins le gain marginal est significatif. Cela est dû au choix de f, qui est elle aussi concave. Ce phénomène

est réaliste, car un petit investissement apporte souvent plus de résultats qu'un ajout de cette même valeur à un investissement déjà conséquent.

2. Un faible investissement de l'assureur est toujours bénéfique, même s'il existe un résultat sans campagne, en raison des entreprises qui auraient résolu le problème de manière autonome (estimées ici à 10%).

3. L'investissement de l'assureur est rentable jusqu'à un seuil de 7 700 €. Au-delà de cette valeur, les bénéfices de la campagne ne couvrent plus le coût engagé.

4. La valeur de x maximisant le bénéfice de l'assureur est de 1 900 €. Pour cette valeur, le bénéfice attendu pour l'assureur est de 3 069 €. En d'autres termes, en moyenne, pour cette faille, l'assureur réaliserait un bénéfice net de 3 069 € en investissant 1 900 € dans la prévention.

5. Ce type de résultat s'apparente à une étude de fonction d'utilité.

Comme discuté dans la section 3.2.2.1, l'assureur peut effectivement être bénéficiaire dans le cadre d'une campagne de prévention. Cet exemple illustre bien que l'argent investi en prévention est récupéré en moyenne grâce à la réduction des sinistres et de leur gravité.

Résultats à la maille assurée Nous reprendrons ici la valeur de $x = 1\,900$ € et regarderons quels sont les assurés qui bénéficient de l'aide de l'assureur dans ce cadre optimisé.

Seuls 50% des entreprises du portefeuille étaient affectées par une augmentation de risque du fait de la faille. Parmi ces 50, seuls 4 recevraient de l'aide de l'assureur. Il est à noter que ce nombre augmente avec l'investissement ; ainsi, pour $x = 3\,000$ €, le nombre passe à 7. La liste de ces 4 entreprises ainsi que la somme reçue par chacune est visible dans le tableau (A.4).

Nom	Augmentation	Investissement de l'assureur
ENT27Ind	2 544 €	583.48 €
ENT79Com	2 262.3 €	517.66 €
ENT80Ind	2 615.53 €	599.60 €
ENT97Ind	1 132.80 €	199.24 €

Figure A.4: Entreprises choisies par l'investissement optimal, augmentation du risque et somme investie par l'assureur

Ces quatre entreprises sont aussi les quatre entreprises avec la plus grande augmentation de risque en € après l'arrivée de la faille. Il est également à noter que la cinquième entreprise avec la plus grande augmentation n'a qu'une augmentation de 400€, ce qui fait une grande différence avec la quatrième. Cette méthode privilégie donc bien les assurés avec le plus grand risque. Elle semble préférer (avec ce choix de f) donner au maximum à celles ayant le plus besoin.

Ce type de méthode permet donc de sélectionner les assurés avec le plus grand besoin et de savoir quelle somme est à attribuer à qui pour avoir une rentabilité optimale du côté de l'assureur.

Cette méthode laisse pourtant certaines entreprises sur le côté. Si nous reprenons le tableau (3.7), qui présentait les entreprises dont l'apparition de la faille a créé la plus grande augmentation en pourcentage, il est possible de remarquer que certaines entreprises présentes ne le sont plus dans le tableau (A.4). En d'autres termes, certaines entreprises verraient leur risque (et donc leur prime) augmenter fortement mais ne bénéficieraient d'aucune action de l'assureur. Cela pourrait ne pas être bénéfique pour l'assureur, qui pourrait sembler les laisser sur le côté (alors que ces entreprises se

sentent en danger et/ou ne comprennent pas l'augmentation de la prime). L'assureur pourrait alors choisir d'utiliser d'autres fonctions à maximiser (et non plus $\mathbb{E}(M)$). Il pourrait par exemple remplacer les C_i par les pourcentages d'augmentation du risque mentionnés ci-dessus. Ainsi, les entreprises les plus touchées seraient privilégiées. Cela pourrait néanmoins diminuer la rentabilité de l'assureur. Une métrique d'optimisation intermédiaire pourrait alors être choisie. Ces choix constituent un pilotage de la part de l'assureur.