Cybersécurité

Steve Kremer, Ludovic Mé, Didier Rémy, Vincent Roca

Cybersécurité

Défis actuels et axes de recherche à l'Inria

Steve Kremer, Ludovic Mé, Didier Rémy, et Vincent Roca

(Résumé de la version anglaise, mai 2019)

©Inria/illustration Clod

Préambule

Ce document est un résumé de la version anglaise du livre blanc d'Inria sur la cybersécurité[1], qui a été coordonné par Steve Kremer, Ludovic Mé, Didier Rémy et Vincent Roca. Ce document est un résumé de la version anglaise du livre blanc d'Inria sur la cybersécurité (1), qui a été coordonné par Steve Kremer, Ludovic Mé, Didier Rémy et Vincent Roca. La rédaction repose sur la contribution de nombreux scientifiques d'Inria et de ses partenaires. Le livre blanc s'adresse à un large public et a été écrit pour permettre différents niveaux de lecture. Son premier objectif est de présenter l'analyse d'Inria sur les défis en cybersécurité en matière de recherche. À cette fin, il inclut un aperçu général des sujets de recherche académique en cybersécurité et une cartographie des recherches existantes sur la cybersécurité chez Inria. Il comprend aussi des présentations techniques des différents domaines de la cybersécurité et une description détaillée du travail effectué au sein des équipes-projets communes entre Inria et ses partenaires académiques susceptibles d'intéresser des experts en cybersécurité ou toute personne cherchant des informations détaillées sur un sujet particulier. Il se termine par des recommandations générales. Dans ce résumé en français, nous nous concentrons sur les parties générales et stratégiques au détriment des parties techniques et de la cartographie détaillée des recherches menées.

1. La version complète en anglais est disponible aux adresses suivantes :
   — PDF : https://files.inria.fr/dircom/extranet/LB_cybersecurity_WEB.pdf
   — EPUB : https://files.inria.fr/dircom/extranet/livre_blanc_cybersecurite.epub
   — HTML : https://files.inria.fr/dircom/extranet/livre_blanc_cybersecurite/livre_blanc_cybersecurite.html

1 Introduction

La transformation numérique de notre société est en train de changer radicalement la manière dont les systèmes informatiques sont utilisés. Une grande partie de la population est connectée en permanence à Internet, utilisant un nombre important de services. Simultanément, nous sommes exposés en permanence à des attaques : nos données personnelles peuvent être volées, modifiées ou détruites. Nous courons également le risque de divulguer par erreur et de façon irréversible nos données personnelles sur Internet. Les entreprises, les états et leurs infrastructures critiques, qui sont aujourd'hui interconnectés, sont également vulnérables. Les dommages économiques et sociétaux des cyberattaques réussies peuvent être considérables. La cybersécurité est ainsi devenue une préoccupation générale pour tous, citoyens, professionnels et décideurs.

1.1 Quelques exemples instructifs

Malheureusement, les incidents de cybersécurité sont fréquents. Nous décrivons ici quelques exemples illustratifs, afin de mettre en évidence la grande diversité des attaques.

L'attaque ciblée de TV5 Monde : En 2015 la chaîne de télévision TV5 Monde a été victime d'un sabotage majeur qui a conduit aux modifications illégitimes des informations présentes sur son site Web et ses réseaux sociaux, puis à l'arrêt pendant deux jours de l'infrastructure réseau et donc de la diffusion des programmes. L'Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) a pu établir que l'attaque avait été soigneusement planifiée. Quelques mois plus tôt, les attaquants avaient pénétré une première fois sur le réseau interne en utilisant un login et son mot de passe volés. Ils ont ainsi progressivement recueilli des informations sur l'infrastructure réseau et les comptes existants, et exploité des services non configurés qui dépendaient encore de comptes et mots de passe par défaut. Les équipements réseau (routeurs et commutateurs) ont alors été corrompus afin de provoquer une panne et rendre tout redémarrage impossible.

Attaque Mirai et réseaux d'équipements domestiques zombies : Le virus Mirai a pour sa part permis de transformer des appareils domestiques vulnérables (telles que des caméras IP) en équipements zombies contrôlés à distance à l'insu de leur propriétaire. Ce réseau « botnet » a plus tard été utilisé pour lancer des attaques par déni de service. C'est ce qui s'est passé en octobre 2016 lorsque les serveurs de la société Dyn ont été pris pour cible et se sont rapidement écroulés sous la charge, conduisant au blocage pendant plusieurs heures d'une partie d'Internet.

Le rançongiciel WannaCry : En 2017, le virus WannaCry s'est très rapidement propagé dans le monde entier, infectant plus de 230 000 ordinateurs dans plus de 150 pays en une seule journée (source Wikipédia). Une fois un ordinateur infecté, ce virus se réplique sur de nouveaux ordinateurs cibles, puis chiffre les données de l'utilisateur avant de demander une rançon en échange de la clé de déchiffrement.

Vulnérabilités du vote électronique : Plusieurs pays européens ont organisé ces dernières années des élections politiques au moyen de systèmes de vote par Internet (par ex. les Français résidant à l'étranger ont pu voter par Internet pour les élections législatives de juin 2012). Il a été démontré qu'il était possible d'écrire des logiciels malveillants pouvant modifier la valeur d'un vote exprimé à l'insu de l'électeur. Lors des élections législatives estoniennes de 2011, une attaque similaire a été signalée et une expérience grandeur nature a été conduite sur des sujets parfaitement informés.

Ré-identification dans la base de données anonymisée d'AOL de requêtes sur le Web : En 2006, AOL a publié une base de données anonymisée contenant plus de 20 millions de requêtes de recherche sur le Web. Cependant, malgré l'anonymisation, certains utilisateurs ont aisément pu être ré-identifiés. L'anonymisation de bases de données est une tâche complexe qui comporte des pièges et qui exige de trouver un juste équilibre entre utilité et confidentialité.

Les vulnérabilités Spectre et Meltdown : Deux vulnérabilités matérielles, Spectre et Meltdown, publiées début 2018, exploitent l'optimisation d'exécution spéculative des processeurs modernes : même si l'exécution spéculative peut ensuite être écartée si l'hypothèse qui en contrôle l'exécution se révèle finalement fausse, un accès mémoire laisse une trace dans le cache. L'idée de ces deux attaques est de forcer un accès mémoire interdit par ce biais. Ces attaques sont particulièrement sévères car elles

exploitent des mécanismes qui sont au cœur de la conception des processeurs modernes et qui de ce fait ne peuvent pas être facilement modifiés.

Ampoules connectées et épilepsie : Il a été montré que certaines ampoules connectées étaient vulnérables à des attaques totalement inattendues, puisque les allumer à une fréquence bien choisie permet de déclencher des crises épileptiques. Cette attaque particulière exploite une combinaison de plusieurs défauts, notamment une absence d'authentification de l'équipement générant les ordres, mais aussi la possibilité de rassembler un grand nombre d'ordres dans une seule commande transmise.

Ces divers exemples illustrent la diversité des attaques possibles : les attaques couvrent tous les domaines, depuis les objets du quotidien jusqu'aux infrastructures ; certaines cherchent à affecter des cibles bien identifiées ou, à l'inverse, à toucher le plus large public possible ; certaines attaques exploitent des techniques élémentaires, d'autres à l'inverse utilisent des moyens techniques extrêmement sophistiqués.

1.2 Qu'est-ce que la cybersécurité ?

Wikipédia définit la cybersécurité comme « *la protection des systèmes informatiques contre le vol et l'endommagement de leur matériel, logiciels ou informations, ainsi que contre la perturbation ou le détournement des services qu'ils fournissent* ». Plus précisément, la cybersécurité consiste à assurer trois propriétés de l'information, des services et de l'infrastructure informatique : *confidentialité*, *intégrité*, et *disponibilité*. Ainsi, sécuriser un système d'information signifie empêcher une entité non autorisée d'accéder aux données informatiques, aux services informatiques ou à l'infrastructure informatique, de les modifier ou de les rendre indisponibles. Une autre propriété de plus en plus importante est la protection de la vie privée, qui peut être considérée comme la confidentialité du lien entre les personnes et les données. Notez que les termes sécurité et sûreté sont parfois mal utilisés. Tandis que la sûreté fait référence aux menaces accidentelles, la sécurité fait référence aux menaces intentionnelles. La sécurité et la sûreté demeurent des domaines très différents et bien identifiés qui reposent sur des hypothèses différentes, et les mécanismes de protection contre les menaces accidentelles et intentionnelles sont généralement complémentaires. Dans le livre blanc, nous nous limitons à la sécurité.

1.3 Organisation du livre blanc Inria sur la cybersécurité

Le chapitre 1 du livre blanc donne une introduction à la cybersécurité : il développe les exemples d'attaque résumés ci-dessus et les leçons qu'on peut en tirer. Il dresse en outre un rapide aperçu des domaines de recherche en cybersécurité.

Les chapitres 2 à 6 du livre blanc présentent chacun des sous-domaines de la cybersécurité plus en détails et décrivent les contributions des équipes Inria. Ces chapitres sont résumés ici dans la section 3, mais sans faire référence aux contributions de l'Inria.

Le chapitre 7 du livre blanc présente un bref aperçu des activités de cybersécurité à l'Inria et de leur positionnement dans le paysage académique français. Nous reprenons assez largement ces éléments dans ce document, dans la section 4.

En dépit des progrès importants réalisés récemment dans plusieurs domaines de la cybersécurité, d'importantes questions scientifiques restent ouvertes. Le livre blanc se termine donc par une liste de défis où l'Inria pourrait apporter des contributions majeures. Ceux-ci sont tous repris dans la section 5, mais décrits de façon beaucoup plus concise.

2 Aspects politiques, économiques et sociétaux de la cybersécurité

Au delà des sujets scientifiques qui seront discutés dans la prochaine section (§3), la cybersécurité pose aussi des questions d'ordre politique, économique et sociétale, dans lesquelles l'Inria s'implique également. En particulier, l'Inria mène des actions avec d'autres organismes français ou européens, qui peuvent apporter des éléments factuels lors de l'élaboration de réglementations comme le RGPD (§2.1). Les deux sujets les plus importants sont la protection de la vie privée et la sécurité des citoyens, des entreprises, ou des états. Nous décrivons plus en détail la tension entre ces deux aspects, un problème qui revient souvent (§2.2), ainsi que la question de la souveraineté (§2.3).

2.1 Actions européennes en cybersécurité

La stratégie de cybersécurité de l'Union européenne [2] a pour objectif d'améliorer la cyber-résilience et la réactivité de l'Europe tout en préservant pour chaque nation un niveau de souveraineté lui permettant de contrôler les principales composantes de sa stratégie de défense nationale.

L'Agence européenne chargée de la sécurité des réseaux et de l'information (ENISA) [3], créée en 2004, est un acteur important du paysage européen de la cybersécurité. Une extension significative de ses missions est actuellement à l'étude, visant à en faire l'interface privilégiée avec les états membres et aider à l'application des directives sur la cybersécurité.

Mi-2016, l'Union européenne a adopté la directive sur la sécurité des réseaux et des systèmes d'information (dite directive NIS). Cette directive se concentre à la fois sur les fournisseurs de services numériques (DSP) et les opérateurs de services essentiels (OES). Ceux-ci sont tenus responsables de la notification des incidents de sécurité, même si les services sont gérés par des sociétés non européennes ou si la gestion du système d'information est sous-traitée à des tiers. DSP et OES sont également tenus de fournir des informations qui permettent une évaluation approfondie de la sécurité de leurs systèmes d'information et de leurs politiques de sécurité. Enfin, les États membres sont tenus d'identifier les organismes chargés de la collecte et du traitement des incidents de sécurité, en plus d'une autorité nationale compétente (par exemple, l'ANSSI en France).

L'Union européenne contribue également à promouvoir le développement de produits et de services sécurisés dès la conception dans toute l'Europe. Pour atteindre cet objectif, elle propose de mettre en place un cadre européen de certification capable de délivrer des certifications de sécurité et des labélisations de produits et de services au niveau européen. Il s'agit là d'une tâche complexe, car même s'il existe un très large éventail de systèmes de certification de sécurité dans le monde, il n'existe pas de solution unifiée ou combinée.

L'Organisation européenne pour la cybersécurité (ECSO) [4] assure la liaison avec la Commission européenne pour la définition d'un cadre européen de certification en matière de sécurité. L'ECSO a publié un état de l'art des normes industrielles existantes en matière de cybersécurité pour différents domaines d'activité et travaille actuellement à l'élaboration d'un schéma générique englobant de nombreux systèmes de certification existants, évaluant le niveau de confiance fourni par chaque système et les mettant en correspondance avec un ensemble harmonisé de niveaux de sécurité [5].

La Commission européenne a également publié un projet officiel, intitulé *Commission Recommendation of 13.9.2017 on Coordinated Response to Large Scale Cybersecurity Incidents and Crises* [6], qui définit les objectifs et les modalités de la coopération entre les états membres et les institutions européennes pour répondre à des incidents ou des crises liés à la cybersécurité.

L'Union européenne est aussi à l'avant-garde en matière de protection de la vie privée et des données, avec le Règlement Général sur la Protection des Données (RGPD) et le règlement ePrivacy qui le complétera (§3.4.1).

2.2 Tension entre sécurité publique et vie privée

Avec l'augmentation de la menace terroriste, nous avons assisté, dans plusieurs pays, au déploiement de systèmes de surveillance de masse destinés à aider à combattre le terrorisme. En France, il s'agit de la loi relative au renseignement [7]. En particulier, cette loi impose le déploiement de boîtes noires au sein des fournisseurs d'accès Internet (FAI) français afin de collecter en temps réel les informations de connexion de certaines cibles préalablement identifiées et d'analyser les informations de connexion des abonnés FAI afin d'identifier des suspects potentiels via un processus automatique (dont les détails ne sont pas connus publiquement). La ré-identification de l'abonné nécessite une décision officielle du Premier ministre (ou d'un délégué).

Ces lois soulignent la tension entre la sécurité publique et la vie privée. Elles ont été critiquées en raison de leur coût économique et de leur inefficacité potentielle, en particulier face au *paradoxe du faux positif*. Le risque est que plus de faux positifs ne feront que surcharger les technologies, ce qui entraînera encore plus de travail pour les agents du renseignement, qui sont déjà surchargés.

L'expression « surveillance de masse » a été utilisée pour désigner les pratiques dans lesquelles des gouvernements ou des organisations gouvernementales assurent la surveillance et la collecte de données

2. https://ec.europa.eu/digital-single-market/en/cyber-security
3. Initialement dénommée *European Network and Information Security Agency*, voir https://www.enisa.europa.eu/
4. https://www.ecs-org.eu
5. Les documents produits par le groupe de travail "WG1" d'ECSO sont disponibles à l'adresse suivante : http://www.ecs-org.eu/working-groups/wg1-standardisation-certification-labelling-and-supply-chain-management
6. http://ec.europa.eu/transparency/regdoc/rep/3/2017/EN/C-2017-6100-F1-EN-MAIN-PART-1.PDF
7. Loi no 2015-912 du 24 juillet 2015 relative au renseignement, Legifrance.

à l'échelle nationale (voire à une échelle plus vaste). Cela s'oppose à la *surveillance ciblée* qui vise un individu particulier. En réaction à cette évolution de la surveillance (et en particulier aux révélations d'E. Snowden), l'IETF a considéré que *la surveillance omniprésente est une attaque* dans le RFC 7258 et que les protocoles de l'IETF devraient la limiter. Le chiffrement par défaut fait partie des initiatives de l'IETF pour cela.

2.3 Questions de souveraineté

Étant donné que la plupart des infrastructures essentielles sont maintenant contrôlées par des ordinateurs, souvent connectés à Internet, la protection de l'infrastructure exige la protection des systèmes et des réseaux informatiques. La cybersécurité est donc une question de souveraineté pour les États et l'UE. Par conséquent, les États et l'UE doivent être en mesure de comprendre les risques et les menaces. Cela exige les compétences scientifiques les plus élevées et ne peut être maintenu à long terme qu'en poursuivant des recherches avancées dans tous les domaines de la cybersécurité. Nous devons non seulement disposer des meilleurs experts, mais nous devons aussi les avoir en nombre suffisant pour couvrir les besoins croissants. En outre, nous avons également besoin d'experts à tous les niveaux pour être en mesure de mettre en œuvre correctement les politiques de sécurité.

Les États devraient également avoir le pouvoir d'agir. Cela exige un contrôle suffisant de l'infrastructure numérique et de l'ensemble de la chaîne de cybersécurité, car la sécurité de l'ensemble dépend de la sécurité du maillon le plus faible. Cela implique le contrôle des logiciels et du matériel utilisés dans les infrastructures critiques, afin qu'ils puissent être analysés et certifiés exempts de bogues et de portes dérobées, ainsi que le contrôle du stockage des données.

Le matériel est l'un des maillons faibles, la France et l'Europe n'ayant plus la capacité de concevoir et produire leur propre matériel. Par conséquent, une certaine forme de souveraineté a déjà été abandonnée. En effet, il est tout à fait possible que des dispositifs matériels soient équipés de portes dérobées ou de fonctions cachées qui permettent, par exemple, à un organisme gouvernemental ou à une entreprise d'espionner le trafic Internet ou d'empêcher le fonctionnement d'un service particulier.

En fait, la nature numérique et donc dématérialisée de la cybersécurité rend la souveraineté en matière de cybersécurité différente des autres formes de souveraineté, comme la défense. Alors que la seconde est l'apanage des États ou des organisations supranationales, la première peut être mise en œuvre à plus petite échelle. De nombreuses entités (citoyens, entreprises, associations, etc.) peuvent revendiquer une certaine souveraineté sur la sécurité de leurs propres données, systèmes informatiques et réseaux. Une conséquence de la numérisation est le transfert potentiel de certaines des souverainetés étatiques traditionnelles à d'autres entités : enregistrement foncier sur une blockchain, frappe de la monnaie avec des monnaies numériques, ou services d'identification des citoyens [8], etc. Ces différents niveaux de souveraineté ne s'excluent pas, mais se complètent, laissant la souveraineté de chaque type de données au niveau le plus approprié. Cette capacité de décentralisation de la cybersécurité ne devrait pas mettre en danger la souveraineté des États. Au contraire, c'est une chance qu'il faut exploiter, en laissant une certaine autonomie aux différentes entités dans certaines limites établies par des incitations, des réglementations et des lois.

3 Aspects techniques de la cybersécurité

La première étape de la cybersécurité consiste à identifier les menaces et à définir une politique de sécurité. Ces menaces peuvent cibler le matériel, le réseau, le système d'exploitation, les applications ou les utilisateurs eux-mêmes (§3.1). Des mécanismes de détection et de protection doivent être conçus pour se défendre contre ces menaces.

L'un des principaux mécanismes de protection est la cryptographie (§3.2) : les *primitives cryptographiques* peuvent assurer la confidentialité et l'intégrité des données. Elles doivent faire l'objet d'une analyse continue (on parle de cryptanalyse) pour assurer le plus haut niveau de sécurité. Cependant, les primitives cryptographiques ne suffisent pas à garantir la sécurité des communications et des services : cette tâche nécessite l'utilisation de ce qu'on appelle des *protocoles cryptographiques*, mettant en œuvre des interactions plus riches au-dessus des primitives. Les protocoles cryptographiques sont des systèmes distribués : s'assurer qu'ils atteignent leurs objectifs, même en présence d'un adversaire actif, nécessite l'utilisation de techniques de vérification formelle.

Bien que primitives et protocoles cryptographiques soient des éléments fondamentaux de la sécurité, des services de sécurité supplémentaires, tels que l'authentification et le contrôle d'accès, sont nécessaires pour appliquer une politique de sécurité (§3.3). Ces services de sécurité, généralement fournis par le

8. Par exemple "SecureIdentity" https://secureidentity.co.uk/

système d'exploitation ou les périphériques réseau, peuvent eux-mêmes être attaqués et parfois contournés. Par conséquent, les activités sur le système d'information doivent être supervisées afin de détecter toute violation de la politique de sécurité. Enfin, comme les attaques peuvent se propager extrêmement rapidement, le système doit pouvoir réagir automatiquement ou au moins se reconfigurer pour éviter de propager les attaques.

Bien que s'appuyant sur les primitives et protocoles cryptographiques, la protection de la vie privée (§3.4) fait intervenir des propriétés, des techniques et des méthodologies qui lui sont propres. De plus, l'étude de la vie privée exige souvent de prendre en compte des aspects juridiques, économiques et sociologiques.

Tous ces mécanismes de sécurité doivent être soigneusement intégrés dans les applications critiques (§3.5). Cela englobe les applications traditionnelles critiques du point de vue de la sûreté, qui en étant de plus en plus connectées sont aussi de plus en plus vulnérables aux attaques intentionnelles, mais aussi les nouvelles infrastructures fonctionnant dans le cloud ou connectées à une multitude d'objets (IoT).

3.1 Connaître, comprendre et modéliser les menaces

La connaissance fine de la manière dont un système peut être attaqué est bien entendu un avantage important lorsqu'on cherche à mettre en place des mécanismes de protections. A cet égard, les cryptologues ont montré la voie : les faiblesses des primitives cryptographiques sont systématiquement étudiées (crypanalyse) pour prendre une longueur d'avance sur des adversaires éventuels. Cette démarche est le fondement de la confiance que l'on peut avoir dans ces primitives. De manière similaire, l'étude des menaces et des attaques contre les systèmes d'information permet de concevoir des réponses adaptées et d'augmenter ainsi le niveau de sécurité global de ces systèmes et le niveau de confiance de leurs utilisateurs. Les menaces sont nombreuses, les attaques pouvant cibler le matériel, le réseau, le système ou les applications (très souvent à travers les actions d'un logiciel malveillant), mais aussi les utilisateurs eux-mêmes (ingénierie sociale, hameçonnage). L'attaquant peut quant à lui être interne ou externe à l'organisation attaquée.

Dans cette section, nous présentons les travaux menés pour étudier menaces et attaques : ils s'appuient sur des modèles d'attaque variés qui définissent la connaissance de l'attaquant sur le système à attaquer et les actions qu'il peut entreprendre sur celui-ci.

3.1.1 Attaques contre le matériel

Les attaques physiques contre le matériel sont une menace importante, car elles peuvent mettre à mal la sécurité d'un mécanisme qui a pourtant été mathématiquement prouvé sûr, parce qu'elles sortent du modèle dans le cadre duquel ce mécanisme à été étudié. Il y a deux grandes classes d'attaques matérielles : les attaques par observation se contentent de mesurer des paramètres physiques du matériel pendant son fonctionnement normal ; les attaques par perturbation cherchent à modifier les conditions physiques dans lesquels se déroule normalement l'exécution afin de provoquer un calcul incorrect et changer ainsi le comportement du programme. Ces deux types d'attaques nécessitent un accès physique au dispositif attaqué.

Plus récemment, le matériel a également été attaqué par logiciel. Cette forme d'attaque est d'autant plus dangereuse qu'elle ne nécessite pas forcément un accès physique au dispositif attaqué. Un scénario d'attaque que l'on peut envisager aujourd'hui sérieusement est une attaque matérielle déclenchée par une application JavaScript intégrée dans une page Web.

3.1.2 Attaques contre les services réseau

Les attaques contre les services réseau sont très nombreuses ; nous n'en donnerons que deux exemples, ciblant deux services essentiels d'Internet.

Le routage est un service de base qui permet de déterminer le chemin que peut prendre tout paquet réseau pour atteindre sa destination. Attaquer ce service peut permettre d'isoler une partie d'Internet ou de rediriger tout le trafic d'un pays à travers un point de surveillance.

Le service DNS (Domain Name System) permet pour sa part de traduire les noms humainement compréhensibles (au sein d'une URL, par exemple) en adresses IP. Une attaque contre ce service peut permettre de rediriger un utilisateur vers un faux site Web, avec comme risque de voler les informations d'identification et d'authentification qui peuvent être demandées à l'utilisateur. Une version sécurisée de ce service, appelée DNSSEC, est disponible, mais son déploiement prend malheureusement du temps.

3.1.3 Le facteur humain

Comme dans beaucoup d'autres domaines, il existe un adage bien connu en sécurité informatique qui veut que la principale menace pesant sur un système se trouve entre la chaise et le clavier. C'est peut-être exagéré, mais il faut reconnaître que les utilisateurs sont trop souvent la source de problèmes de sécurité. Ils peuvent être la cible d'attaques comme l'hameçonnage (*phishing*). Ils peuvent aussi contourner les mécanismes de protection disponibles, consciemment, en raison de la complexité (réelle ou perçue) excessive de leur utilisation, ou inconsciemment, en évaluant mal les risques réels en raison d'un niveau d'éducation et de formation insuffisant.

3.1.4 Modélisation des menaces et attaques par arbre d'attaque

Un arbre d'attaque est une représentation graphique décrivant les actions que l'attaquant peut entreprendre pour atteindre son objectif et leur enchaînement possible : chaque feuille de l'arbre exprime une étape que l'attaquant doit effectuer pour mener à bien son attaque ; chaque nœud non terminal de l'arbre contient une étiquette (et, ou, séquence) indiquant comment ses enfants sont connectés. Les arbres d'attaque sont largement utilisés lors de l'étape d'analyse des risques : les menaces ainsi modélisées, une politique de sécurité peut-être élaborée pour les parer, puis mise en œuvre en utilisant les mécanismes de sécurité les plus appropriés.

3.2 Cryptographie : primitives, schémas, et protocoles

La cryptographie joue un rôle essentiel et constitue la base de la cybersécurité. Dans cette section, nous couvrons de multiples aspects de la cryptographie, de la conception des primitives aux protocoles plus complexes qui fournissent des garanties de haut niveau pour la sécurité des communications et des transactions.

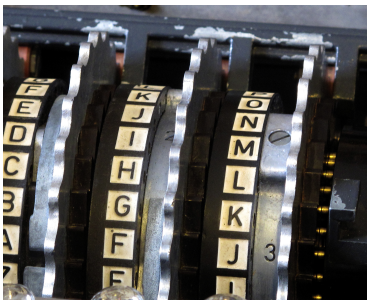

FIGURE 1 – Gros plan des rotors de la machine Enigma - TedColes via Wikipedia, CCO

La cryptographie vise à fournir des techniques et des outils pour sécuriser les communications, même en présence d'un adversaire. Historiquement, le but principal de la cryptographie était d'assurer la confidentialité des messages par le chiffrement, c'est-à-dire que l'information reste cachée pour les personnes non autorisées. Les premières méthodes de chiffrement étaient généralement assez naïves, comme le chiffrement de César qui consiste à décaler chaque lettre d'une constante (par exemple, remplacer « A » par « D », « B » par « E », etc.), ce qui est facilement cassé en utilisant une analyse de fréquence d'occurrence des lettres. Une avancée substantielle s'est produite pendant la Seconde Guerre mondiale avec la machine à rotors Enigma utilisée par l'Allemagne. Casser le chiffrement Enigma a nécessité des efforts et des ressources considérables. Aujourd'hui, la cryptographie repose sur des bases mathématiques solides et vise à garantir bien plus que la simple confidentialité : elle fournit des outils pour protéger l'intégrité et l'authenticité des messages (en évitant par exemple que le montant d'une transaction financière soit modifiée), assurer la non-répudiation (l'expéditeur ne peut nier être l'auteur du message) et l'anonymat. De plus, ces outils peuvent être combinés pour atteindre des objectifs plus complexes. A l'inverse, le but de la cryptanalyse est de « casser » les techniques cryptographiques, ce qui est en fait utilisé pour vérifier leur robustesse.

Cette section comporte trois parties. La première est consacrée aux primitives cryptographiques, qui sont les composants de base permettant de chiffrer ou de signer un message. Puis nous continuons avec les schémas cryptographiques, qui en s'appuyant généralement sur ces primitives, fournissent des services de sécurité plus élaborés, par exemple en garantissant l'intégrité et l'authenticité des messages de taille arbitraire. Enfin les protocoles cryptographiques s'appuient sur ces schémas pour atteindre des objectifs de sécurité plus complexes, par exemple pour établir un canal de communication sécurisé.

3.2.1 Les primitives cryptographiques

Les primitives cryptographiques, telles que les fonctions de chiffrement et les signatures numériques, constituent les briques de base. On distingue deux familles : les primitives symétriques et asymétriques. La cryptographie symétrique suppose que les parties en communication partagent une clé secrète. Ce type de cryptographie est plus efficace que la cryptographie asymétrique, mais nécessite l'échange préliminaire de la clé secrète. La cryptographie asymétrique, ou à clé publique, ne nécessite pas cet échange, car la clé publique (utilisée pour le chiffrement ou pour la vérification d'une signature) n'a pas besoin de rester secrète. L'utilisation de ces deux types de primitives est complémentaire. Une approche typique

est d'utiliser la cryptographie asymétrique pour échanger une clé privée utilisée ensuite pour chiffrer (de façon symétrique) les communications ultérieures. Ainsi, les deux types de cryptographie sont requis dans la plupart des mises en œuvre.

Aujourd'hui, nous disposons de primitives symétriques et asymétriques matures. Néanmoins, il y a encore un besoin de recherche à la fois en cryptanalyse et en conception. Le but du cryptanalyste est de trouver les faiblesses et d'évaluer la résistance des constructions existantes. D'une part, ce travail permet d'évaluer finement la difficulté de résoudre les problèmes sous-jacents supposés difficiles, en tenant compte à la fois de l'évolution de la puissance de calcul et de l'amélioration des algorithmes. D'autre part, ce travail explore de nouvelles méthodes d'attaque, telles que les attaques reposant sur l'existence d'un ordinateur quantique ou des attaques qui exploitent des canaux auxiliaires. La conception de nouvelles primitives peut être rendue nécessaire par des percées en cryptanalyse (auxquelles nous devons nous attendre) ou par de nouveaux besoins industriels, comme par exemple la cryptographie à bas coût pour des équipements à faible consommation d'énergie.

3.2.2 Les schémas cryptographiques

Alors que les primitives cryptographiques sont les éléments de base, les schémas cryptographiques atteignent des propriétés plus fortes avec des modes d'opération spécifiques. Certaines applications, telles que l'externalisation du calcul, peuvent nécessiter des fonctionnalités plus avancées que le chiffrement classique. Par exemple, les schémas de chiffrement homomorphe et fonctionnel permettent de travailler sur des données chiffrées et des preuves cryptographiques (« preuves de connaissance ») peuvent être utilisées pour obtenir des preuves que le calcul a été effectué correctement. Une tendance récente apparue avec la complexité croissante des schémas cryptographiques et de leurs preuves de sécurité consiste à développer des outils permettant de vérifier les preuves de sécurité sur machine et atteindre ainsi un niveau de confiance très élevé dans la sécurité de certaines constructions.

3.2.3 Les protocoles et services cryptographiques

La sécurité des communications et des transactions est de nos jours assurée par des protocoles cryptographiques, tels que TLS. La sécurité des primitives et des schémas cryptographiques sous-jacents n'est cependant pas suffisante pour garantir des objectifs globaux en matière de sécurité, tels que la confidentialité, l'authenticité ou l'anonymat. Même un examen minutieux de ces protocoles par des experts, ne peut garantir l'absence de vulnérabilités : des preuves de sécurité rigoureuses, éventuellement assistées par ordinateur, allant de la spécification à l'implémentation, sont devenues indispensables pour garantir un très haut niveau de confiance.

On peut distinguer trois approches dans ce domaine. La première utilise des preuves par réduction pour montrer que le fait de casser la sécurité d'un protocole impliquerait de casser les primitives cryptographiques sous-jacentes. Ce sont des preuves mathématiques, généralement manuscrites, bien qu'une nouvelle tendance consiste à utiliser des assistants de preuve pour les vérifier mécaniquement. La deuxième direction utilise des outils de vérification automatique pour analyser le protocole et trouver des vulnérabilités dans la logique du protocole (par exemple des attaques comme celle de l'homme du milieu). Ces outils sont capables d'analyser des protocoles complexes mais idéalisent la cryptographie sous-jacente. Enfin, la troisième approche vise à produire des implémentations vérifiées. Elle permet des implémentations vérifiées de bout en bout, mais s'appuie sur des systèmes de type expressifs pour des langages de programmation dédiés et requiert une grande expertise. Un succès majeur dans ce domaine est une implémentation complètement vérifiée de TLS.

3.3 Services et mécanismes de sécurité

Pour utiliser un système informatique, un utilisateur doit tout d'abord s'identifier (identification) puis prouver qu'il est vraiment celui ou celle qu'il prétend être (authentification). Cette identité prouvée est ensuite utilisée par le système pour accorder ou restreindre les accès de cet utilisateur aux ressources du système (contrôle d'accès) ou pour éviter qu'une information spécifique ne circule vers une destination donnée (contrôle de flux). Par ailleurs, lorsqu'un utilisateur a besoin d'exécuter un programme sur une machine non fiable, un mécanisme matériel peut être utilisé pour garantir l'isolation et l'intégrité du logiciel (informatique de confiance).

Tous les services de sécurité mentionnés ci-dessus sont préventifs. Ils peuvent malheureusement être parfois contournés. Il est donc nécessaire de leur adjoindre des mécanismes de supervision permettant de vérifier que les actions réalisées sur le système n'enfreignent pas sa politique de sécurité (détection d'intrusion et corrélation d'alerte).

Bien sûr, une telle violation de politique peut s'être produite à l'insu de l'utilisateur qui peut avoir été attaqué par un logiciel malveillant agissant alors sous l'identité usurpée de l'utilisateur. L'analyse et la détection des logiciels malveillants sont donc aussi des services de sécurité à offrir. Idéalement, si une intrusion ou un logiciel malveillant est détecté, le système doit réagir, au moins pour se reconfigurer afin d'éviter la propagation de l'attaque ainsi que des attaques similaires.

Dans le reste de cette section, nous présentons plus en détail chacun des services de sécurité introduits ci-dessus.

3.3.1 Identification et authentification

L'identification, pour une entité donnée (par exemple un utilisateur, un service, ou un dispositif) est l'acte d'affirmer son identité. L'authentification, pour cette même entité, est l'acte de prouver qu'elle est vraiment celle qu'elle prétend être. L'authentification est réalisée en présentant un *authentifiant* qui appartient généralement à l'une des catégories suivantes : une connaissance (comme par exemple un mot de passe ou un code pin), une possession (comme par exemple une carte d'accès), une caractéristique physique (par exemple une empreinte digitale). Pour augmenter la force de la preuve, ces authentifiants peuvent aussi être combinés dans ce que l'on appelle l'authentification multifacteurs. Malgré leurs nombreux inconvénients, les mots de passe restent le moyen d'authentification le plus courant.

Une autre forme d'identification est liée au lien qu'on peut chercher à créer entre une donnée et son propriétaire. Par exemple, le tatouage numérique consiste à insérer des messages cachés (un filigrane) dans les données. Une bonne technique de tatouage doit créer un lien solide entre la donnée et le filigrane, de telle sorte que la déformation éventuelle des données (par exemple par traitement d'une image) n'efface pas le filigrane. Cette technique est utilisé dans la gestion des droits d'auteur et la protection contre la copie, en se limitant principalement au contenu multimédia. Le spectre d'usage est cependant plus large et le tatouage peut aussi être utilisé pour protéger du code source, des bases de données, etc.

3.3.2 Contrôle d'accès et contrôle de flux

La mise en œuvre de la sécurité passe par la définition précise des entités qui peuvent avoir accès à telle ou telle information. Classiquement, les droits d'accès (en lecture ou écriture, par exemple) à l'information sont accordés si et seulement si une condition est remplie (par exemple, l'utilisateur qui demande l'accès est correctement identifié et authentifié). Une séquence de lecture-écriture engendre un flux d'informations qui peut également être contrôlé dans certains cas.

3.3.3 Informatique de confiance

L'informatique de confiance s'appuie sur des garanties fournies par un matériel sécurisé. Un tel matériel permet d'assurer l'intégrité de la plate-forme sur laquelle il est installé et par conséquence la sécurité du lancement du système d'exploitation. Il est également possible d'instaurer une confiance au niveau des applications en exécutant du code dans des environnements isolés et fiables, appelés enclaves. Ces enclaves offrent en outre la possibilité d'attester les résultats produits par un code donné.

3.3.4 Détection d'intrusion et corrélation d'alerte

Presque tous les systèmes présentent des défauts qu'un attaquant peut exploiter pour contourner les mécanismes de sécurité préventive existants. Par conséquent, la supervision du système est d'une importance cruciale pour identifier toute violation de la politique de sécurité, c'est-à-dire toute intrusion.

L'approche principale et largement déployée de nos jours pour détecter ces violations consiste à définir les symptômes des actions malveillantes liées à la violation et à rechercher l'apparition de ces symptômes dans les données dont on dispose (trafic réseau, fichier du système d'exploitation, journaux applicatifs, etc.). Une alternative à cette approche consiste à définir les activités normales du système supervisé pour mesurer ensuite les écarts potentiels par rapport à cette référence. Dans les deux cas, le défi consiste à détecter toutes les intrusions, mais seulement les intrusions. Dans la pratique, cependant, les mécanismes de détection sont loin d'être parfaits, ce qui entraîne de nombreux faux positifs (fausses alertes) ou faux négatifs (attaques non détectées).

Comme la détection d'intrusion conduit souvent à un grand nombre de faux positifs, une étape supplémentaire, la corrélation des alertes, est souvent nécessaire. Elle consiste à appliquer au flux d'alertes une série de transformations pour améliorer progressivement leur contenu (par exemple sur l'origine de ces attaques, sur la vulnérabilité qui a été exploitée, etc.). Une étape importante consiste à vérifier que les conséquences de l'attaque dénoncée sont bien présentes dans le système, sans quoi l'alerte correspondante est probablement un faux positif.

3.3.5 Analyse et détection des logiciels malveillants

Les logiciels malveillants (virus, vers, rançongiciels, logiciels espions, logiciels publicitaires, chevaux de Troie, enregistreurs de frappe, rootkits, etc.) constituent une menace majeure pour nos systèmes d'information (systèmes d'exploitation, applications et données), en particulier du côté client (PC ou smartphones). Ils doivent donc être étudiés et analysés pour en permettre ensuite la détection.

Le but de l'analyse des logiciels malveillants est d'obtenir une compréhension complète d'un code présumé malveillant. Il s'agit d'identifier les cibles (par exemple un utilisateur particulier, ou une machine exécutant un système d'exploitation particulier), le but de l'attaque (par exemple fuite d'informations ou chiffrement des données), les techniques utilisées pour contourner les mécanismes de sécurité, celles utilisées pour éviter sa propre détection. Avant de conduire cette étude, l'analyste doit en premier lieu contourner les protections anti-analyse, comme les techniques d'obscurcissement du code, mises en place par le créateur du logiciel malveillant.

La détection des logiciels malveillants s'appuie généralement sur l'analyse de toute l'information reçue par un appareil (une machine, un téléphone, un pare-feu, etc.) ou sur l'analyse complète des fichiers de cet appareil. Le moteur de détection compare ces données à une base de référence contenant les caractéristiques des logiciels malveillants connus. Le défi consiste à maintenir cette base à jour, les auteurs de logiciels malveillants générant constamment de nouvelles variantes du même code malveillant afin d'éviter cette détection. Des projets de recherche proposent donc des techniques de détection s'appuyant sur le comportement concret du logiciel malveillant, qui reste souvent identique entre ses différentes variantes.

3.3.6 Réaction aux attaques détectées

Après la phase de détection des intrusions et des logiciels malveillants, l'étape suivante consiste à répondre (éventuellement quasi-automatiquement) aux attaques détectées par des actions appropriées : modification de la politique de sécurité, nouvelles configurations des mécanismes de sécurité existants, mise en œuvre de nouveaux mécanismes de sécurité, déploiement de correctifs, etc. Bien entendu, il est important d'éviter que les contre-mesures n'aient elles-mêmes des conséquences similaires, voire pires, que celles de l'attaque.

3.4 Protection de la vie privée et des données personnelles

La protection de la vie privée est souvent définie comme la capacité des individus à contrôler leurs données personnelles et à décider ce qu'ils veulent révéler, à qui et dans quelles conditions. En France, la loi « Informatique et Libertés » de 1978 définit les données à caractère personnel comme des informations pouvant être directement ou indirectement liées à une personne, par le responsable de traitement, ou par tout tiers, par tout type de moyen. Cette notion et les obligations qui en découlent constituent la pierre angulaire de la réglementation française et européenne.

Le travail sur la notion de vie privée recouvre plusieurs dimensions : des dimensions légales, techniques, mais aussi économiques, pour comprendre l'écosystème sous-jacent qui détermine souvent les pratiques de collecte, culturelles, car on constate des approches différentes suivant les racines culturelles des utilisateurs, et sociologiques, l'utilisateur final se déclarant souvent concerné par la vie privée tout en se comportant de manière opposée.

Cette section aborde essentiellement les aspects légaux et techniques, et s'achève avec une analyse des fuites de données personnelles dans différents systèmes, que ces fuites soient le résultat d'actions délibérées des utilisateurs ou inversement aient lieu à leur insu.

3.4.1 Législation

Pour tenir compte des évolutions intervenues au cours de la dernière décennie, l'Union européenne a adopté le Règlement Général sur la Protection des Données (RGPD), en vigueur depuis mai 2018. Le principal changement est l'accent mis sur la responsabilité juridique des responsables de traitements et leurs sous-traitants le cas échéant. Ces acteurs doivent effectuer des études d'impact sur la protection des données, mettre en œuvre des mesures de protection dès la conception, et être en position de rendre des comptes sur les traitements effectués. Les droits des personnes sont également renforcés par une meilleure information et un meilleur contrôle sur ses données.

Toutefois, le RGPD fournit très peu d'indications sur la mise en œuvre efficace de ces concepts. Aussi, un travail interdisciplinaire reste nécessaire pour réduire cet écart entre les instruments juridiques et techniques, par exemple en définissant des méthodes rigoureuses d'analyse des risques, de protection

de la vie privée dès la conception, ou en proposant des techniques pour renforcer la transparence et le contrôle utilisateur.

3.4.2 Outils pratiques

Passer des principes de haut niveau du RGPD à des produits et services respectueux de la vie privée demande des outils spécifiques pour faire une évaluation d'impact, aider à la protection de la vie privée dès la conception, ou anonymiser des bases de données. En particulier, l'anonymisation, qui exige un difficile compromis entre respect de la vie privée et utilité, est une tâche complexe pour laquelle la confidentialité différentielle s'est avérée être un concept clé pour fournir des garanties prouvables. L'avènement du stockage dans un cloud personnel est également un outil essentiel pour redonner du contrôle aux utilisateurs.

3.4.3 A propos des systèmes existants

Notre monde connecté est à l'origine de nombreuses fuites de données, et des domaines aujourd'hui hors d'atteinte pourraient sous peu être eux-aussi concernés. Certaines fuites sont soit délibérées soit consenties par l'utilisateur. C'est le cas des réseaux sociaux, pour lesquels un partage important d'informations peut avoir des conséquences nombreuses pour la vie privée de l'utilisateur. Ainsi le profilage rendu possible par ce partage a permis l'émergence d'entreprises spécialisées dans la manipulation des utilisateurs, notamment lors d'élections ou de consultations publiques.

La géolocalisation est une information fréquemment collectée par de multiples moyens et à de multiples occasions, qui a la particularité de permettre beaucoup d'inférences, allant des habitudes de vie jusqu'à des informations sensibles comme la fréquentation d'un lieu de culte. Différentes parades ont été conçues pour en limiter les risques, parmi lesquelles la notion de géo-indiscernabilité qui s'appuie sur la confidentialité différentielle.

L'utilisateur peut parfois être invité à fournir des éléments biométriques, comme ses empreintes digitales, afin de permettre son identification et authentification par des systèmes de contrôle d'accès ou avec des documents d'identité sécurisés. Un soucis est que ces données sont à la fois hautement discriminantes et stables (non modifiables), ce qui crée des risques majeurs en matière de sécurité et de protection de la vie privée.

Mais très souvent les fuites de données personnelles surviennent sans le consentement utilisateur. C'est le cas lorsque l'on navigue sur le Web puisque chaque site visité peut déclencher de multiples échanges de données à l'insu de l'utilisateur et au profit de sociétés tierces spécialisées dans le profilage. L'information peut être utilisée à des fins de publicité ciblée, mais aussi pour discriminer les utilisateurs (par exemple, avec un tarif personnalisé). Les protections proposées sont à la fois des réglementations mais aussi des mécanismes côté client, comme les outils de blocage de publicités ou de traçage dont l'efficacité reste variable. Ce domaine est en constante évolution, autant du coté des outils de traçage que des outils de protection.

L'avènement des smartphones et de l'Internet des Objet ont contribué à ce que les fuites de données atteignent un volume et une précision sans précédent, sans que l'utilisateur en ait toujours conscience. Les objectifs de la recherche sont ici d'analyser ces systèmes, de donner une information transparente sur les comportements cachés, de mettre en évidence les bonnes et mauvaises pratiques, de proposer des techniques pour améliorer la transparence et le contrôle utilisateur, voire d'encourager certains acteurs à changer de pratiques.

Internet, en tant que moteur de notre monde connecté, est aussi, intrinsèquement, à la source de fuites de données. Les réseaux d'accès sans fil que nous utilisons au quotidien en sont un exemple. Ainsi des centres commerciaux ont mis en place des systèmes de traçage afin d'analyser leur fréquentation. Ces systèmes sont basés sur l'analyse des trames Wi-Fi envoyées par un smartphone à la recherche d'un point d'accès connu. Des recherches sont nécessaires pour analyser ces technologies et proposer des versions préservant la vie privée dans la mesure du possible.

3.5 Infrastructures, systèmes et applications critiques

3.5.1 Les infrastructures critiques

Les infrastructures de communication sont principalement conçues pour offrir un service, souvent avec la facilité d'utilisation et l'efficacité comme objectif principal. Cependant, la sécurité est également cruciale, car ces infrastructures peuvent être utilisées pour stocker et manipuler des données sensibles. Une perte de disponibilité (ou simplement d'efficacité) ou de l'intégrité des données peuvent également

avoir un fort impact économique. Nous donnons ici trois exemples d'infrastructures critiques : le Cloud, les SDN, et la Blockchain.

Le cloud : les problèmes de sécurité et de protection de la vie privée sont exacerbés dans le contexte de la distribution de l'information et du Cloud. La sécurité reste un obstacle à une adoption plus large des services Cloud par les organisations pour leurs besoins en matière de services critiques. La protection de la vie privée est quant à elle cruciale pour les utilisateurs du Cloud à l'ère de l'IoT, qui collecte toujours plus de données personnelles. La sécurisation des hyperviseurs et des systèmes d'exploitation, ainsi que la vérification des propriétés de sécurité sur ces systèmes sont des sujets de recherche importants et d'actualité.

Les SDN : la transition vers les « Software Defined Networks » (SDN) et la virtualisation des fonctionnalités est une évolution majeure dans le domaine des réseaux. L'objectif est d'en améliorer la flexibilité tout en réduisant les coûts. Cependant, ces évolutions centralisent le contrôle du réseau, fournissant un point unique de défaillance qui peut être ciblé par l'attaquant. Les SDN permettent un meilleur couplage entre le réseau et les applications, ce qui facilite considérablement le développement applicatif, mais crée également une nouvelle surface d'attaque.

La blockchain : en proposant un registre partagé immuable, les blockchains permettent de nombreuses applications. La sécurité fournie et le niveau de confiance atteint restent cependant à analyser finement par les chercheurs spécialisés en cryptographie et en systèmes distribués. En outre, comme pour tout système reposant sur Internet, les blockchains peuvent subir des attaques réseau de bas niveau. Enfin, l'anonymat offert peut entrer en conflit avec les exigences de sécurité des organismes publics et légaux.

3.5.2 Les systèmes critiques

Les systèmes critiques sont des systèmes qui exigent un niveau de fiabilité très élevé, car toute défaillance pourrait avoir des conséquences extrêmement néfastes, telles que la mise en danger de vies humaines, de graves dommages à une infrastructure, ou d'importantes pertes économiques.

Ainsi, la plupart des systèmes qui sont critiques en matière de sûreté le sont en fait également sur le plan de la sécurité et doivent résister aux cyberattaques. Ces systèmes sont souvent « cyber-physiques » (on parle de CPS), des éléments informatiques collaborant pour le contrôle et la commande d'entités physiques. Composés de nombreux sous-systèmes interconnectés, fonctionnant avec différents protocoles de communication à différentes échelles, ils offrent une large surface d'attaque. Certains CPS sont des cibles de prédilection parce qu'une attaque réussie pourrait mener à une fermeture d'une infrastructure critique d'un État, avec un impact économique et politique énorme. Par conséquent, la cybersécurité est aujourd'hui souvent une question essentielle pour les CPS.

Bien que la cybersécurité des CPS puisse réutiliser les approches, méthodes et techniques traditionnelles de sécurisation des systèmes et des réseaux, de nouvelles approches doivent également être développées pour faire face à la dynamique de ces systèmes. Il est intéressant de noter que ces nouvelles approches relèvent souvent de la sécurité réactive, où la supervision et la réaction à des situations anormales sont importantes. L'apprentissage statistique est ici susceptible d'occuper un rôle majeur et introduira de nouveaux problèmes de sécurité.

L'Internet des objets et les systèmes industriels sont aujourd'hui les deux systèmes critiques qui attirent le plus d'attention.

L'Internet des Objets (IoT) : la révolution de l'IoT est en train de changer la façon dont nous interagissons avec les dispositifs physiques. Cette évolution s'accompagne évidemment d'enjeux pour la sécurité et la protection de la vie privée. Cependant, l'IoT n'est pas qu'une nouvelle forme de systèmes distribués, et leurs particularités, comme par exemple les ressources limitées de ces objets, apportent de nouveaux défis, les solutions de sécurité classiques n'étant pas toujours applicables. De nombreuses directions de recherche doivent être explorées : la mise à jour sécurisée des firmware, les systèmes d'exploitation sécurisés, la cryptographie à bas coût et assistée par le matériel, la sécurité des technologies sans fil, des politiques de sécurité dédiées à l'IoT, des méthodes de détection d'intrusions adaptées, la programmation sécurisée et les compilateurs pour les applications IoT, ainsi que des protocoles d'authentification dédiés.

Les systèmes industriels : l'une des principales difficultés avec les systèmes industriels est qu'ils n'ont pas fait l'objet d'une réflexion sur la sécurité dès la conception. Ces systèmes n'ont généralement pas été conçus pour être connectés à Internet et les protocoles utilisés ne sont pas sécurisés. Les pare-feux à usage

général et les dispositifs de détection d'intrusion ne gèrent pas ces protocoles, dont les spécifications sont rarement publiques. Enfin, les périphériques sont construits avec des processeurs top lents pour utiliser des protocoles et primitives cryptographiques usuels, et ils reposent sur des mécanismes ad hoc. Les défis sont donc nombreux.

3.5.3 Exemples de domaines critiques

Nous illustrons ici les problèmes de sécurité au travers de trois cas d'usage.

La médecine : la médecine est considérablement transformée par la révolution numérique et est de ce fait de plus en plus exposée aux cyber menaces. L'une de ces menaces est la fuite de données médicales sensibles. Comme ces données médicales sont également cruciales pour la recherche, il faut en permettre l'utilisation tout en protégeant la vie privée des personnes. Le compromis entre utilité et protection revêt ici une importance particulière. Les appareils médicaux, tels que les robots chirurgicaux, sont quant à eux de plus en plus connectés. De même, les implants médicaux offrent des connexions sans fil pour éviter des interventions chirurgicales. Ces connexions augmentent la surface d'attaque de la même manière que pour l'IoT et les CPS.

Les véhicules autonomes : les véhicules autonomes sont de véritables systèmes d'information qui échangent avec leur environnement (autres voitures et dispositifs extérieurs). Le système à bord doit être protégé, tout comme les échanges. Il faut assurer l'intégrité, l'authenticité et aussi dans certains cas la confidentialité des échanges. La cryptographie a, ici aussi, un rôle essentiel. Par ailleurs, le système informatique interne de la voiture contient des parties critiques (liées au pilotage, par exemple) et des parties non critique (liées au divertissement à bord, par exemple). Ces parties doivent être rigoureusement isolées afin que des attaques ne se propagent pas aux parties les plus critiques.

L'apprentissage statistique : l'apprentissage statistique (≪ Machine Learning ≫) est utilisé dans de nombreuses applications. Son impact le plus fort reste sûrement pour les applications de reconnaissance d'image. Les techniques d'apprentissage souffrent principalement de deux menaces qu'il s'agit de considérer sérieusement. La première, appelée apprentissage antagoniste (≪ adversarial learning ≫), consiste à ajouter du bruit soigneusement conçu (à peine visible à l'œil nu) à une image afin d'obtenir une mauvaise classification. La deuxième est liée à la protection de la vie privée et consiste à extraire d'un système des informations sur les données d'entraînement.

4 Les forces Académiques à l'Inria et en France

Nous donnons ici un bref aperçu des activités de cybersécurité à l'Inria et de leur positionnement dans le paysage académique français.

Organisation de la recherche à l'Inria : la recherche à l'Inria est organisée en petites équipes partageant un projet de recherche commun, qui sont souvent des équipes conjointes avec d'autres institutions académiques telles que le CNRS, des universités, des écoles d'ingénieurs spécialisées ou d'autres instituts de recherche (INRA, INSERM, etc.). En moyenne, environ la moitié des personnels de recherche des équipes de l'Inria sont employés par des institutions partenaires, ce qui donne à l'Inria un effet de levier. Dans ce document, la recherche à l'Inria signifie toujours la recherche dans les équipes Inria incluant les partenaires. La taille moyenne d'une équipe est de 18 personnes (chercheurs, enseignants, jeunes chercheurs en thèse ou en postdoc, et ingénieurs de recherche) mais avec une grande variation, allant de 3 à 5 personnes pour les plus petites équipes jusqu'à 45 à 50 personnes pour les plus grandes.

Les forces en cybersécurité : la recherche en cybersécurité a été l'une des priorités de l'Inria au cours des quinze dernières années [9]. La cybersécurité couvre maintenant 7% de l'activité de l'Inria, avec environ 30 équipes travaillant dans ce domaine [10], dont deux tiers ont la cybersécurité comme thème unique ou principal de recherche. Au total, cela représente l'équivalent d'environ 200 postes à temps plein, ce qui correspond à un quart des forces académiques françaises en cybersécurité. Les autres forces académiques sont principalement les scientifiques du CNRS et les enseignants chercheurs des universités et des écoles

9. La sécurité a été largement discutée dans le plan stratégique de l'Inria pour la période 2003-2007.
10. Toutes les équipes de l'Inria travaillant dans le domaine de la cybersécurité sont en annexe de la version anglaise du livre blanc.

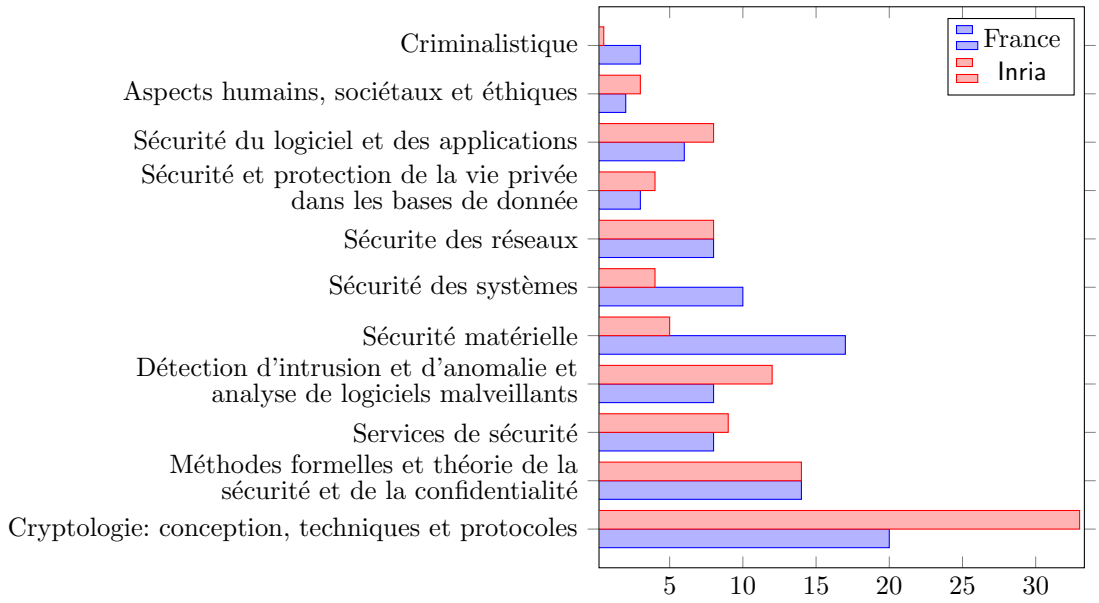

FIGURE 2 – Principaux thèmes de la cybersécurité en % des forces académiques françaises[11]

d'ingénieurs hébergés dans les UMR du CNRS mais en dehors des équipes Inria, les enseignants chercheurs de l'Institut Mines Telecom (IMT), et les scientifiques du CEA. Les forces académiques françaises dans le domaine de la cybersécurité ont pratiquement doublé au cours de la dernière décennie. Cette croissance se poursuit, mais à un rythme plus modéré. Bien qu'il y ait eu de nouvelles embauches, cette forte croissance est surtout due à des chercheurs et des membres du corps professoral d'autres domaines de l'informatique qui se sont tournés vers la cybersécurité.

Les domaines de recherche : la Figure 2 décrit la proportion des activités de recherche entre les principaux domaines de la cybersécurité à l'Inria, en rouge, et dans l'ensemble des entités académiques françaises, en bleu. La comparaison est instructive. La cryptologie est une force de l'Inria, avec un tiers de ses effectifs. Ceci résulte de la longue implication de l'Inria dans la théorie des nombres, le calcul formel, la cryptanalyse et la théorie des codes. En effet, l'Inria joue un rôle clé au niveau mondial dans la conception de nouvelles primitives et protocoles cryptographiques et dans la cryptanalyse des primitives cryptographiques.

Le deuxième domaine de recherche le plus important à l'Inria est celui des méthodes formelles appliquées à la sécurité et à la vie privée, ce qui est dû en grande partie au transfert de compétences issues des méthodes formelles et au fait que l'Inria, et la France en général, sont très bien positionnés dans le domaine des méthodes formelles.

A l'inverse, la sécurité des matériels et des systèmes est sous-représentée à l'Inria, mais ce domaine est encore bien couvert ailleurs en France, au CEA, au CNRS et à l'IMT. Les recherches en criminalistique, quasiment absentes à l'Inria, existent un peu en France [12].

Répartition géographique : les principaux acteurs français de la cybersécurité en dehors de l'Île-de-France sont en Bretagne (Brest et Rennes) et dans la région Rhône-Alpes (Lyon et Grenoble). Viennent ensuite Nancy et la Côte d'Azur (Nice et Sophia Antipolis). Les forces de l'Inria se situent surtout en Île-de-France, Nancy et Rennes. Toutefois, alors que Rennes est un centre important en cybersécurité, la présence de l'Inria y reste relativement faible.

Forces non académiques : la recherche en cybersécurité est parfois aussi menée dans des entreprises privées, bien qu'elle soit alors souvent plus appliquée. Des organismes institutionnels, tels que la DGA [13], qui fait partie du ministère français des armées, et l'ANSSI [14], l'agence française de cybersécurité, qui

11. Source : cartographie des forces académiques françaises en cybersécurité faite par le groupe de travail sur la cybersécurité de l'alliance Allistene ; voir https://www.allistene.fr/files/2018/03/VF_cartographie_2017-06-13.pdf.
12. Les activités criminalistiques en France sont probablement sous-estimées, car les données de la Figure 2 n'incluent pas la recherche en sciences humaines.
13. Direction Générale de l'Armement
14. Agence Nationale de la Sécurité des Systèmes d'Information

fait partie du SGDSN [15] placé sous la tutelle du Premier ministre, ont également une expertise de haut niveau en cybersécurité, dont certaines activités de recherche, et jouent un rôle clé dans la définition et la conduite des politiques françaises en cybersécurité. Il existe aussi quelques laboratoires dépendant du ministère de l'Intérieur, comme le CREOGN [16]. Des recherches sur les enjeux et la réglementation en matière de protection de la vie privée sont menées au sein du laboratoire d'innovation LINC [17] de la CNIL [18]. Les équipes Inria collaborent régulièrement sur des projets spécifiques avec la plupart de ces organismes.

Animation de la communauté : un certain nombre d'associations jouent un rôle important dans l'animation de la communauté dans le domaine de la cybersécurité. Le CNRS a créé en 2016 un pré-GDR [19] *Sécurité Informatique* [20] sur la cybersécurité dont l'objectif est d'animer la communauté académique française, notamment par l'organisation d'ateliers et d'écoles d'été.

L'alliance Allistene [21] a créé un groupe de travail sur la cybersécurité [22] où l'Inria et les principaux acteurs académiques sont représentés pour échanger des informations, construire une vision commune, mener des études telles que la cartographie des forces académiques décrite ci-dessus, et coordonner des actions comme la participation des membres Allistene au FIC [23].

Enfin, l'Inria fait partie du groupe de travail consacré à la recherche et à l'innovation du CoFIS (Comité de la Filière industrielle de sécurité) dont le rôle principal est de favoriser l'industrie française de la sécurité en proposant des actions ciblées pour accroître la compétitivité et la sécurité aux niveaux national et européen. Inria est aussi membre de l'association professionnelle ACN [24] dont le rôle est de fédérer et représenter les principaux acteurs industriels de la cybersécurité. HEXATRUST est une autre association importante composée de 29 PME de la cybersécurité où l'Inria et la recherche académique ne sont cependant pas représentés.

5 Défis et recommandations

Tout le monde est aujourd'hui concerné par la cybersécurité : les États, les industries et les citoyens. La cybersécurité est un sujet important qui soulève de nombreuses questions économiques, sociétales, politiques ou géopolitiques en matière de sécurité et le restera très probablement au cours des prochaines décennies. Nous concluons ce document avec la liste des défis scientifiques et quelques quelques recommandations générales relatives à l'organisation de la recherche en cybersécurité.

5.1 Défis scientifiques

Voici les principaux défis scientifiques que nous avons identifiés [25]. Ce ne sont pas les seuls, mais nous considérons ceux-ci comme particulièrement importants et nous recommandons qu'ils soient explorés de façon prioritaire.

Cryptographie post-quantique : la construction d'un ordinateur quantique est largement considérée comme faisable dans les prochaines décennies et la plupart des techniques de cryptographie utilisées aujourd'hui pourraient être facilement cassée par un tel ordinateur. C'est pourquoi il est important de penser dès maintenant à la cryptographie post-quantique, c'est-à-dire résistant à une cryptanalyse qui utiliserait un ordinateur quantique, car les informations chiffrées aujourd'hui pourront être encore sensibles lorsque les ordinateurs quantiques apparaîtront.

Calcul sur des données chiffrées : le besoin de calculer sur des données chiffrées s'est fait sentir avec l'apparition du Cloud et de l'externalisation des calculs. Ce problème peut être résolu à l'aide de techniques appelées chiffrement homomorphe et chiffrement fonctionnel. En 2009, une percée théorique

15. Secrétariat Général de la Défense et de la Sécurité Nationale
16. Centre de recherche de l'École des officiers de la gendarmerie nationale
17. Laboratoire d'Innovation Numérique de la CNIL
18. Commission Nationale de l'Informatique et des Libertés
19. Un GDR (Groupement De Recherche) est une structure dépendant du CNRS pour animer la communauté scientifique française dans un domaine de recherche particulier.
20. http://gdr-securite.irisa.fr/index.html
21. http://www.allistene.fr/
22. https://www.allistene.fr/organisation-allistene/groupes/groupe-cybersecurite/
23. Forum International de la Cybersécurité
24. Alliance pour la Confiance Numérique
25. Chaque défi peut être mis en contexte en se rapportant à la section correspondante du livre blanc.

a été réalisée avec le premier schéma de chiffrement totalement homomorphe, mais ce schéma est resté impraticable en raison de son inefficacité. Beaucoup de progrès ont été réalisés depuis, mais des recherches supplémentaires restent nécessaires et toute avancée significative devrait avoir rapidement un impact économique.

Protocoles cryptographiques formellement vérifiés de bout en bout : la sécurité des protocoles cryptographiques est extrêmement difficile à assurer et l'utilisation de méthodes rigoureuses et formelles est une nécessité. Les preuves de sécurité assistées par ordinateur doivent inclure tous les aspects de la spécification jusqu'à la mise en œuvre. Des travaux récents, en particulier autour de TLS 1.3, ont montré que cette approche est désormais réalisable. Cependant, cela nécessite toujours une co-conception minutieuse de la preuve et du code qui ne peut être effectuée que par des experts. L'exploitation de ces preuves pour obtenir un code plus général et des propriétés de sécurité plus complexes, par exemple des propriétés d'anonymat, reste un énorme défi.

Sécurité de l'IoT : la sécurité de l'IoT est un défi majeur. Les attaques sont encore relativement faciles (de nombreux dispositifs n'ont pas été conçus pour la sécurité), invasives (par exemple, dans nos vies), et ont un impact potentiel majeur en raison du grand nombre de dispositifs disponibles, ce qui augmente la surface d'attaque et rend les attaques par déni de service distribué (« DDoS ») beaucoup plus faciles. Les axes de recherche sont nombreux : citons par exemple la mise à jour sécurisée du logiciel embarqué, le besoin de primitives cryptographiques à bas coût adaptées à des ressources limitées, l'analyse de la sécurité des nouvelles technologies de réseaux étendus sans fil de faible puissance, la détection des intrusions ou des dispositifs fonctionnant anormalement, ou le besoin de cadres et de protocoles pour faciliter le développement des dispositifs IoT où la sécurité est prise en compte dès la conception.

Protection de la vie privée des citoyens : notre monde connecté a permis une croissance sans précédent des pratiques de collecte de données personnelles, y compris par leur aspect intrusif, que ce soit en surfant sur le Web, en utilisant son smartphone, dans son habitat intelligent, et bientôt en conduisant un véhicule connecté. Le manque de transparence, le fait que de nombreux services et appareils se comportent comme des boîtes noires et le manque de contrôle des utilisateurs sont des problèmes majeurs. Comment exprimer un consentement en l'absence d'information ou d'interface utilisateur ? L'identification des pratiques, souvent cachées, de collecte de données personnelles est entravée par le nombre, la complexité et la diversité des applications et technologies sous-jacentes. Des activités de recherche transversales ambitieuses sont nécessaires pour apporter la transparence, mettre en évidence les bonnes et les mauvaises pratiques et permettre aux régulateurs d'appliquer les lois sur la protection des données.

Données ouvertes et anonymisation : les initiatives de données ouvertes (« Open Data ») peuvent parfois impliquer la publication de bases contenant des données personnelles (voire sensibles). Afin d'assurer la protection de la vie privée des personnes, les données doivent être anonymisées. L'anonymisation robuste, qui résiste efficacement aux attaques de désanonymisation, est un sujet de recherche très actif. Si la confidentialité différentielle est devenue un outil scientifique clé pour obtenir des garanties d'anonymisation prouvables, la recherche d'un meilleur compromis entre utilité et protection de la vie privée reste un sujet ouvert.

Attaques logicielles ciblées sur le matériel : des attaques récentes d'un type nouveau, telles que Rowhammer, Spectre ou Meltdown, ont montré que les attaques mises en œuvre dans les logiciels peuvent exploiter des optimisations cruciales comme l'exécution spéculative, utilisées pendant des années pour améliorer la performance du matériel. Elles sont d'autant plus dangereuses qu'elles permettent des attaques matérielles à distance, contrairement aux attaques classiques par canal auxiliaire qui nécessitent une proximité physique. Des connaissances pointues en matériel, firmware et systèmes d'exploitation sont nécessaires pour mieux comprendre comment ces attaques, pour l'instant de type « preuve de concept », pourront être déployées à plus grande échelle et pour proposer des contre-mesures efficaces qui n'affectent pas trop les performances.

Sécurité et facilité d'utilisation : pour éviter que les utilisateurs contournent les mécanismes de sécurité qui les gênent pour utiliser un service, les services de sécurité doivent être aussi simples à utiliser que possible. Des recherches interdisciplinaires avec des experts en sciences cognitives sont nécessaires pour proposer des interfaces et des mécanismes de sécurité adaptés aux utilisateurs non experts qui garantissent que l'utilisateur est bien conscient des conséquences de ses actions et lui évitent de commettre des erreurs de nature à compromettre la sécurité.

Détection d'intrusion pour réseaux chiffrés : de nos jours, la détection d'intrusion est essentiellement réalisée au niveau du réseau. Si, comme prévu dans un avenir proche, le trafic était systématiquement chiffré, ce qui serait bien sûr une bonne pratique pour la sécurité et la confidentialité, l'analyse des paquets réseau deviendrait de facto inopérante, hormis l'analyse de certaines entêtes. Il faut donc concevoir de nouveaux mécanismes de supervision des systèmes d'information et de production d'alertes, au niveau de l'application, du middleware, du système d'exploitation et même du firmware ou du matériel.

Comprendre la protection de la vie privée et élaborer des outils pratiques : la compréhension des règlements en matière de protection de la vie privée est à la base de toute activité dans ce domaine qui a récemment connu des évolutions majeures avec l'arrivée du règlement européen sur la protection des données (RGPD) d'une part, et de nouvelles opportunités de collecte de données personnelles d'autre part. Le RGPD promeut plusieurs concepts et objectifs sans toujours fournir d'indications sur la mise en œuvre effective de ces nouvelles dispositions réglementaires. Il introduit en particulier le droit à la transférabilité des données, qui permet à l'utilisateur de récupérer ses données dans un format lisible par un humain et portable par machine. Ce droit ouvre de nouveaux champs de recherche autour de la gestion individualisée et du contrôle de ses données personnelles.

Systèmes industriels sécurisés : les systèmes industriels reposent de plus en plus sur des mécanismes logiciels qui peuvent être attaqués. Leur sécurité est donc devenue un enjeu majeur, d'autant plus que les conséquences d'une attaque contre de tels systèmes peuvent être dramatiques. Les spécificités des systèmes industriels nécessitent de revoir les mécanismes traditionnels de sécurité pour les adapter à ce nouveau contexte : le contrôle des systèmes en temps réel étant souvent requis, la sécurité doit aussi être applicable en temps réel ; les protocoles de communication utilisés dans ce contexte et conçus sans soucis de la cybersécurité ne pouvant pas être modifiés du jour au lendemain, ils devront être intégrés de façon transitoire dans des protocoles sécurisés ; la sécurité réactive y prend une place prépondérante en raison de la difficulté de modifier les dispositifs industriels, et donc d'y ajouter a posteriori des mécanismes de sécurité préventive.

5.2 Recommandations générales

Nous terminons par quelques recommandations générales.

La société devrait profiter davantage de l'expertise scientifique académique : pour la plupart des questions de cybersécurité, il existe des universitaires possédant une expertise technique très spécifique qui peuvent fournir des conseils scientifiques utiles et qui le font parfois déjà. Cependant, les scientifiques sont souvent sous-représentés dans les comités consultatifs nationaux ou industriels par rapport aux membres industriels. En outre, certaines positions ou décisions prises par les décideurs, à différents niveaux, montrent une insuffisance de connaissances et d'avis scientifiques.

Favoriser le transfert d'expertise entre la cybersécurité et d'autres domaines informatiques : le besoin d'expertise en cybersécurité est frappant dans la plupart de ses domaines d'application : systèmes industriels, systèmes médicaux, robotique et, plus crucial encore, l'Internet des objets. Malheureusement, la cybersécurité n'y est pas encore suffisamment identifiée comme une priorité et n'est donc pas prise en compte dès le stade de la conception des applications, violant ainsi un principe fondamental. Inversement, la recherche en cybersécurité nécessite de plus en plus de compétences pointues dans d'autres domaines comme les méthodes formelles ou l'apprentissage statistique.

Promouvoir la sécurité également en tant que science expérimentale : certains domaines, comme la sécurité des systèmes et des réseaux où la recherche est plus expérimentale et technologique, semblent souffrir d'un manque de prestige dans le monde académique, du moins en France. Cette image défavorable est accentuée par une difficulté à obtenir des données expérimentales issues du monde réel, pourtant indispensables pour tester et comparer les nouveaux mécanismes proposés. Il faut donc aider les chercheurs à générer ou accéder à ces données.

Éducation : l'éducation est une composante essentielle de la sécurité et des efforts importants de diffusion des connaissances et de sensibilisation devraient être faits pour tous les publics : enseignants, acteurs industriels et spécialistes, citoyens ordinaires, y compris les jeunes enfants. La médiation scientifique doit cibler la société en général. La formation doctorale est bien sûr un outil pédagogique naturel, mais les masters professionnels ou les cursus de fin d'études d'écoles d'ingénieurs devraient tous enseigner les bases

de la cybersécurité en allant au-delà de l'état de l'art dans l'industrie. Les MOOC pourraient aussi être un vecteur efficace de dissémination et d'éducation, ainsi qu'un outil de communication majeur, grâce à une participation active de plusieurs milliers de personnes par session.

La cyber-résilience dès la conception : de nombreux rapports récents sur l'analyse des cyber menaces alertent sur le fait que les États devront faire face à des cyber attaques massives. Sommes-nous prêts à résister à une « cyber-tornade » ? La résilience des infrastructures critiques et, en particulier, des opérateurs d'importance vitale (OVI) est un véritable enjeu. Une grande partie de ce défi est en fait d'ordre organisationnel et technique, plutôt qu'académique, et l'ANSSI a déjà déployé de nombreux efforts pour accroître notre résilience. Néanmoins, la recherche peut être utile : le principe de la sécurité dès la conception s'applique également à la cyber-résilience, qui ne doit pas être considérée après coup, mais prise en compte dès la phase initiale de conception des systèmes, des réseaux et des infrastructures numériques. Par ailleurs, de nombreux sous-domaines de la cybersécurité sont indirectement impliqués dans ce défi : sécurité réactive et détection de logiciels malveillants ; sécurité préventive, et surtout les méthodes formelles qui devraient avoir un impact significatif à long terme sur la résilience des systèmes industriels.

5.3 Remarques finales

Ces dernières années, un effort national important a été fait pour renforcer la sécurité de l'État et des systèmes d'information des opérateurs d'importance vitale. En particulier, l'ANSSI et la DGA—pour ne citer que deux exemples—ont recruté de nombreux experts. Néanmoins, même si l'importance de la recherche et de la formation en matière de cybersécurité est de plus en plus reconnue, très peu d'efforts ont été faits pour recruter davantage d'enseignants et de chercheurs. Dans le domaine de la cyber-résilience, les ressources humaines sont essentielles : nous appelons donc à la poursuite des efforts entrepris par certaines institutions nationales, et que ces efforts s'étendent aux universités et aux écoles ainsi qu'aux établissements de recherche. Tout comme les pays qui font des efforts importants pour améliorer leurs performances en matière de cybersécurité, il est clé que la France continue de renforcer ses forces académiques et, plus globalement, tout son écosystème d'innovation dans ce domaine.