La cybercriminalité comportementale : Historique et régulation

Philippe Baumard

LA CYBERCRIMINALITÉ COMPORTEMENTALE : HISTORIQUE ET REGULATION¤

Philippe Baumard*

RÉSUMÉ

La régulation du cyberespace constitue un effort international où se croisent intérêts publics, privés et souverains. En analysant l'évolution historique des menaces cybernétiques et des doctrines étatiques y répondant, ainsi que le cas particulier de la régulation des contre-mesures, cet article soulève le problème de l'inadéquation du cadre de régulation face aux ruptures technologiques actuelles et à venir dans le domaine de la cyberdéfense. L'article souligne en particulier la nécessité d'ancrer la réflexion sur la régulation dans le corps des sciences cognitives et comportementales, et dans le domaine du machine-à-machine.

Mots-clefs : Attaques informatiques ; Contre-mesures ; Cybercriminalité ; Doctrines ; Régulation.

ABSTRACT

The regulation of cyberspace has grown into an international effort where public, private and sovereign interests often collide. By analyzing the history of cyber-

---

¤ Cet article s'inspire de Baumard P. "The behavioral paradigm shift in fighting cybercrime: Counter-measures, innovation and regulation issues", *Internatiinal Journal on Criminology*, 2(1), 2014, pp.11-22, ainsi que P. Baumard, « La régulation des contre-mesures contre les cyber-attaques », *Archives de philosophie du droit*, vol.56, 2013, pp.177-195. Il est actualisé au 1er octobre 2014.

* Professeur des universités, Conservatoire national des arts et métiers (CNAM).

regulation of counter-measures, this article investigates the difficulty of obtaining a global agreement on the regulation of cyber-warfare. A review of the motives for disagreement between parties suggests that the current regulation framework is not adapted to the current technological change in the cyber-security domain. The article suggests a paradigm shift in handling and anchoring cyber-regulation into a new realm of behavioral and cognitive sciences, and their application to machine learning and cyber-defense.

Keywords : Cyber-attacks ; Counter-measures ; Cyber criminality ; Doctrines ; Regulation.

INTRODUCTION

Cet article interroge l'évolution technologique de la cybercriminalité depuis son émergence dans les années 1970 jusqu'à ses développements récents en 2014. À partir de cette évolution, nous tirons des conclusions pour les doctrines, les stratégies d'innovation, et la régulation de la cybercriminalité dans le contexte de changements techniques radicaux conférant aux attaquants une capacité à se soustraire à l'identification, à l'attribution et au contrôle géographique et souverain des origines des attaques. L'évolution technologique, plus rapide que celle des cadres de régulation ou des doctrines nationales, pose le problème d'une refonte urgente et nécessaire des cadres de coopération internationaux. En rappelant la nature et l'histoire des activités de *hacking* et de cybersécurité, nous essaierons de mesurer dans un premier temps l'adéquation des doctrines et stratégies nationales vis-à-vis de ces ruptures techniques, puis dans un second temps, de cerner les tenants et les aboutissants du cadre de régulation international négocié, mais non encore ratifié, depuis les accords de Budapest en 2001.

La cybercriminalité se définit comme l'utilisation de capacités numériques, électroniques ou logicielles pour dévoyer, détourner, détruire, ou illégalement exploiter des systèmes d'informations publics ou privés. L'histoire technique de la cybercriminalité est celle, sans surprise, d'un dialogue permanent entre l'épée et le bouclier, entre l'attaque et sa contre-mesure. Les « contre-mesures » sont des réponses que l'on oppose à une action ou à un événement de façon à les interdire, les prévenir ou enrayer leur prolifération à la source. Les contre-mesures « d'interdiction » peuvent se contenter de mettre fin, *hic* et *nunc*, à une opération malveillante. C'est ce que font les logiciels de sécurité identifiant un code malicieux (« virus »), l'isolant, le plaçant en quarantaine, pour éventuellement le supprimer. Les contre-mesures dites de « prévention » vont enregistrer et caractériser ce comportement malicieux (par sa signature, sa reconnaissance comportementale) et s'assurer qu'il soit stoppé dès sa détection. Finalement, les contre-mesures dites « actives » ou « contre-offensives » vont prolonger cette interdiction temporaire par une recherche active de sa source d'émission afin de procéder à sa neutralisation.

Les contre-mesures d'interdiction, ou « défensives », se situent au point de réception des attaques. Il s'agit principalement de mesures de sécurité informatique visant à contrôler l'accès, l'usage, la commande de systèmes informatiques à partir de mesures techniques de contrôle des identités numériques (certifications, certificats, prévention d'intrusions, parades anti-virus, etc.). Elles n'impliquent pas de « réponses actives », et peuvent être délimitées au « point d'impact des attaques ». La contre-mesure active, de son côté, implique l'identification de l'attaquant (l'attribution). Dans un contexte technologique permettant la dispersion des attaques, ainsi que

la dissimulation de leurs origines, ces contre-mesures actives requièrent souvent des enquêtes d'attribution qui dépasse le cadre géographique national. Dès lors, elles peuvent mettre en jeu la souveraineté territoriale des partenaires et posent la question de la nécessité d'un droit international numérique permettant des exceptions et des adaptations pour gérer l'ubiquité de ces nouvelles menaces. Une analyse historique de l'évolution technique des attaques nous permettra dans un premier temps de mesurer la distance entre le fait technique et la réponse du droit international, pour discuter, dans un deuxième temps, des évolutions nécessaires à la négociation d'un cadre de régulation débuté en 2001 à Budapest, et n'ayant pas à ce jour trouvé d'accord stable et durable entre les nations signataires.

I. DE L'ACTE SPONTANÉ À L'ACTION PRÉMÉDITÉE: HISTOIRE D'UNE ÉVOLUTION COMPORTEMENTALE

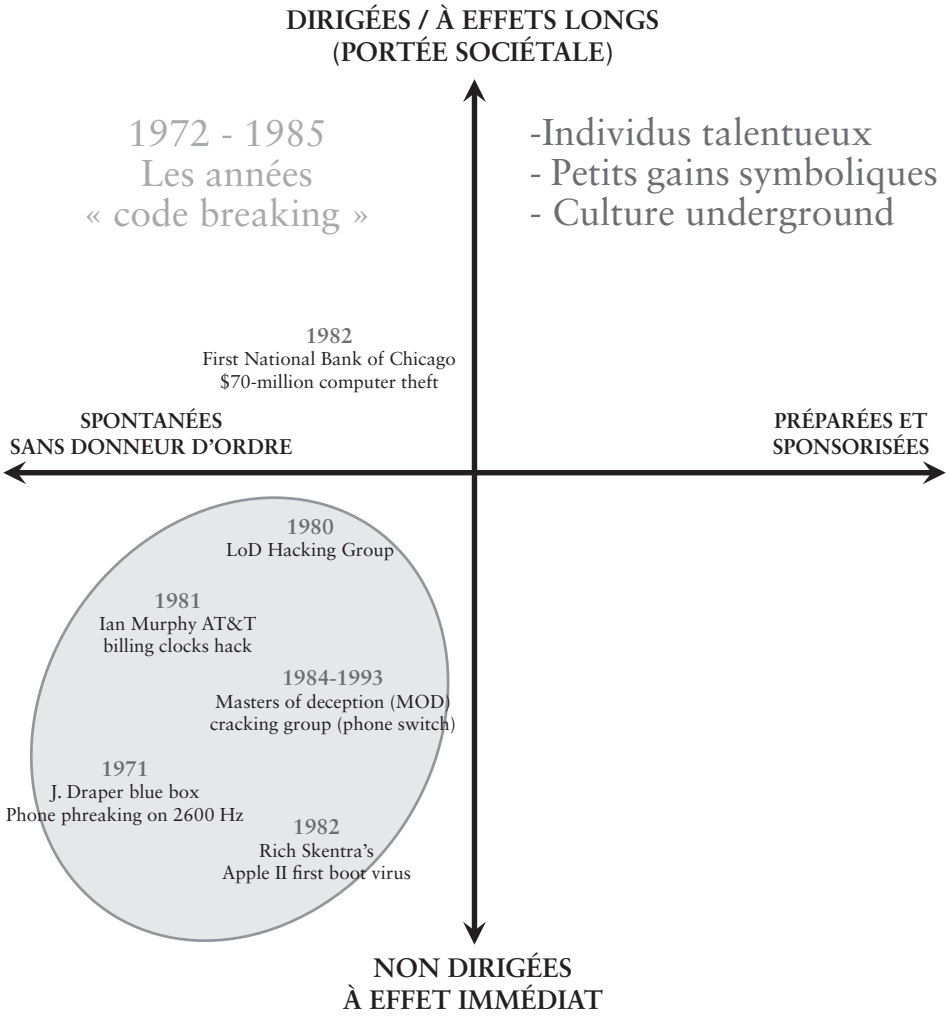
Les origines de la cybercriminalité sont concomitantes avec les efforts de pionniers de la technologie qui exploraient les possibilités techniques d'innovations naissantes. La logique d'exploration et d'appropriation autonome est toujours, à ce jour, une motivation de la création de *hacks*. John Draper était l'un de ces enthousiastes qui ont aidé à populariser les activités de *phreaking*, consistant à la génération de tonalités fréquence, plus tard connu comme la *Blue box*, reproduisant la fréquence de 2 600 hertz de l'infrastructure téléphonique de longue distance d'AT&T dans les années 70.

Ces premières attaques historiques étaient spontanées, motivées par l'exploration technique, non dirigées (sans cible spécifique à l'esprit) et immédiates dans leurs effets. Avec la croissance de l'informatique personnelle, ces pionniers du cracking se sont réunis en associations spontanées, épousant les discours du temps sur la liberté individuelle, la résistance à l'autorité, et jouant des détours offerts par ces technologies émergentes. *Phreaking* et *hacking* devinrent des pratiques partagées qui cimentèrent des amitiés durables entre développeurs, pionniers de l'industrie (Wozniak, Jobs, etc.) et des enthousiastes de la technologie aux motivations parfois politiques. La frontière entre cette culture underground émergente (*yippies*, *hackers*) et une sous-culture criminelle était floue et instable, avec très peu d'autorégulation, et comprenant aussi bien des *teenagers* que des développeurs avancés et des explorateurs autodidactes. Nous appellerons cette période « les années des casseurs de code » où des individus talentueux sont principalement motivés par des gains symboliques, un sentiment d'appartenance et la construction d'une identité.

Au milieu des années 80, les bulletins techniques des groupes de hackers commencent à diffuser des méthodes d'intrusion, parfois tangibles, et fondées sur

du code (comme le premier numéro de *Legion of Doom LOD/H* du 1er janvier 1987)[1]. LOD et MOD (*Masters of Deception*) eurent ainsi une influence décisive dans la transformation de ces mouvements pionniers en communautés organisées de cracking, s'éloignant de leur culture originelle (voir la figure 1).

FIGURE 1
Les années pionnières : le paradigme du « code breaking »



1 *The LOD/H Technical Journal*, vol.1(1), 1987, disponible en ligne : <http://www.textfiles.com/magazines/LOD/lod-1> [Dernière visite, le 01/10/2014].

La guerre froide et la bataille underground pour la libération de Berlin-Est ont également joué un rôle déterminant dans l'évolution de la culture du *hacking* dans les années 80. Aux États-Unis, l'épisode Clifford Stoll (un astronome du LBL qui découvrit accidentellement une intrusion informatique menée depuis l'Allemagne de l'Ouest dans son laboratoire) fut le premier cas qui mit en évidence l'importance d'une coordination entre agences et les difficultés de l'attribution pour des attaques internationales[2]. Ce cas fut aussi celui des premiers symptômes (1986) de menaces persistantes avancées, mettant en avant la complexité et le caractère sophistiqué des campagnes d'intrusion[3].

Le début de l'année 1990 est ainsi concomitant à l'émergence d'une sous-culture criminelle cybernétique. Dans les années 1980, les événements de cracking conduisant à des attaques de large échelle étaient rares. Les deux exceptions notables sont, en 1986, Pak Brain, connu comme le premier virus, et, en 1982, le piratage de la First National Bank of Chicago (70 millions de dollars). La *Great Hacker War* (conflit entre les groupes *Masters of Deception* et *Legion of Doom*, circa 1991-1992) est un autre exemple, - aujourd'hui disputé comme une simple exagération de confrontations triviales - du caractère interpersonnel de ces premières « brouilles » cybernétiques[4]. Il faudra attendre l'opération Sundevil, en 1990, pour voir apparaître la première intervention de taille nationale (sur quinze villes américaines), mais qui ne conduisit qu'à seulement trois arrestations pour des faits mineurs[5]. Les faits incriminés concernaient l'interception de communication privée, la fraude à la carte bancaire ou à la carte d'appels téléphoniques[6]. Les publications comme 2600 et l'émergence du cyberespace accélèrent la démocratisation du *cracking*, du *phreaking* et des techniques de *hacking*, les rendant plus versatiles à une destination d'usage au-delà du simple exploit technologique. Le contrôle distant devient une motivation récurrente, engendrant une démocratisation des chevaux de Troie, qui fut sans doute influente dans la délocalisation de communautés cyber criminelles, tout autant que dans la généralisation d'attaques sociétales (voir la figure 2).

2 C. Stoll, *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*, New-York, Doubleday, 1989.

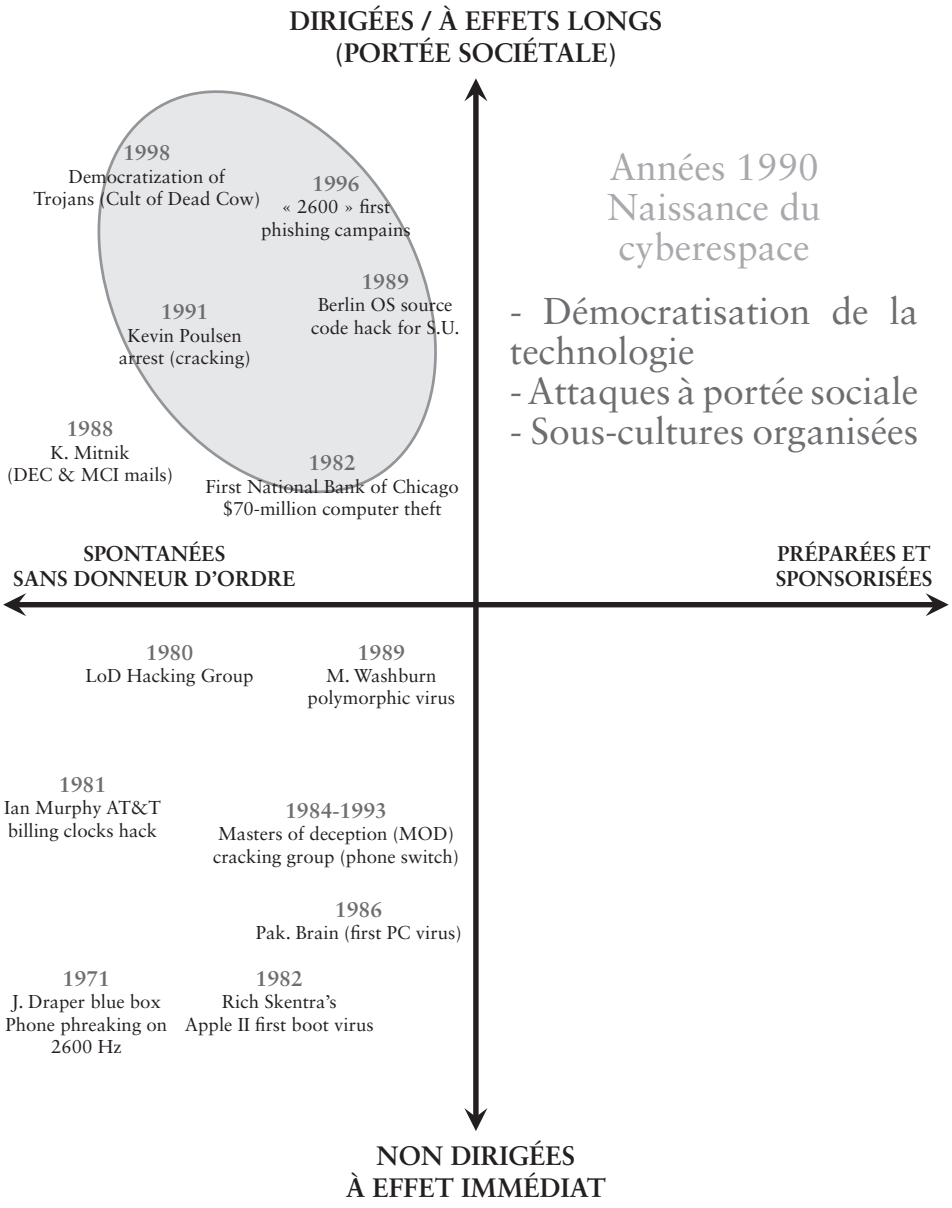
3 Pour plus de détails, v. C. Stoll, « Stalking the wily hacker », *Communication of the ACM*, vol.31(5), 1988, pp.484-500, disponible en ligne : <http://pdf.textfiles.com/academics/wilyhacker.pdf> [Dernière visite, le 01/10/2014].

4 MOD.book.FOUR [End of '90-1991], disponible en ligne : <http://www.textfiles.com/hacking/modbook4.txt> [Dernière visite, le 01/10/2014].

5 A.L. Clapes, *Softwars: the legal battles for control of the global software industry*, Westport, Quorum Books, 1993.

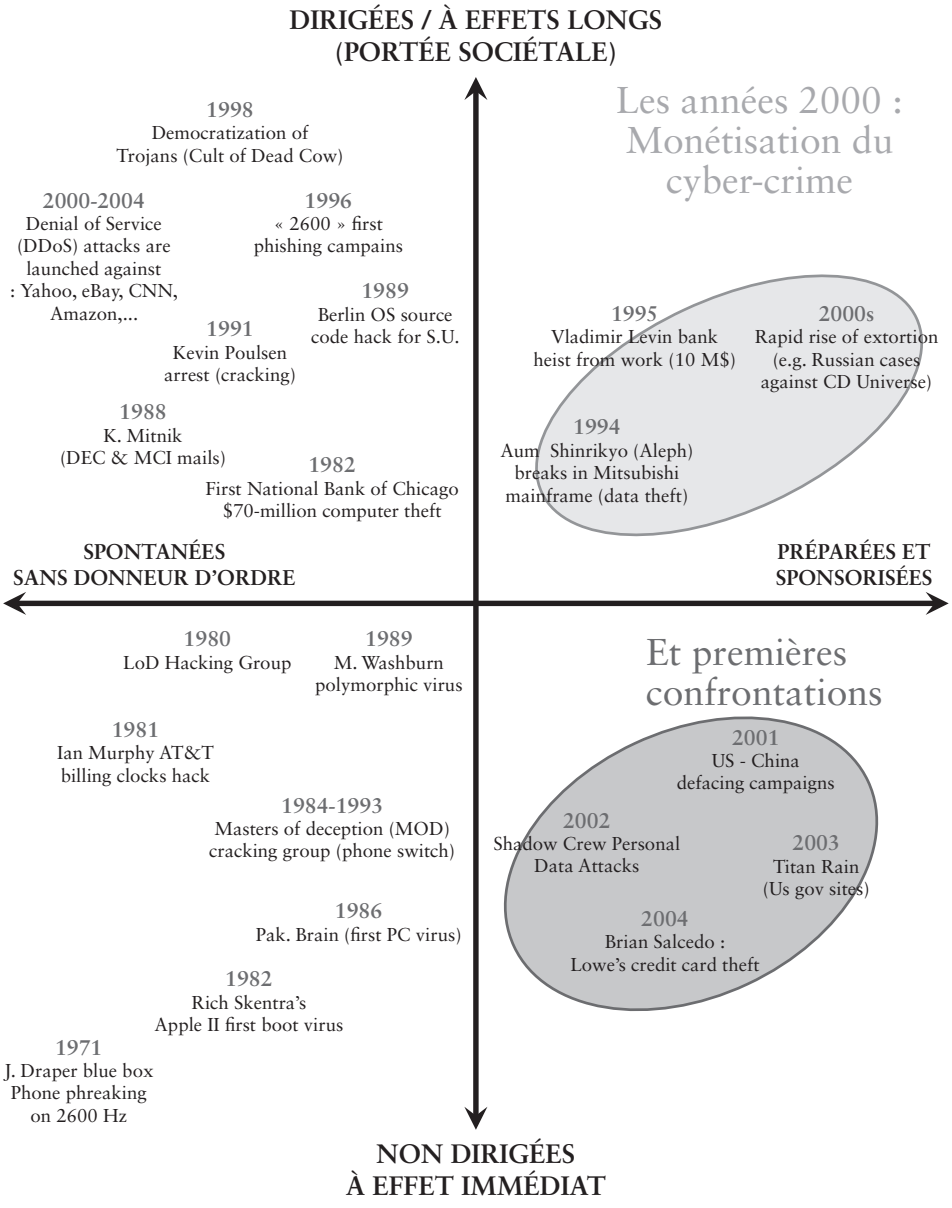
6 B. Sterling, « Part Three: Law and Order », in *The Hacker Crackdown: Law and Disorder on the Electronic Frontier*, New York, Bantam Books, 1994.

FIGURE 2
Les années 1990 : la démocratisation de la cybercriminalité



Les années 2000 signent la fin des attaques archétypales à point d'attaque unique. La croissance du commerce électronique permet d'envisager la multiplication et la systématisation des gains par des attaques coordonnées, et cette monétisation du cybercrime est une promesse de profitabilité pour le crime organisé. La numérisation des industries culturelles (MP3s) crée également un appel d'air pour une popularisation du *cracking*. Le profil des *hackers* prend deux directions. D'un côté, des pirates amateurs (*script kiddies*, consommateurs domestiques) commencent à utiliser des outils mis en ligne sans posséder une connaissance technique avancée (P2P, « CD » d'outils de *cracking*). De l'autre, la production de *malware* fait l'objet d'un marché noir profitable. La corruption des DNS, les dénis de service, les campagnes de « défaçage » et l'espionnage industriel font l'objet d'une monétisation rapide. Les années de 2000 à 2002 sont parmi les plus actives dans la production de *malware*, avec la création de virus tels que ILOVEYOU, Klez.h., Code Red, etc. Le groupe Anonymous est créé en 2003 comme un espace de coordination lâchement couplé d'intérêts très variés, allant de l'activisme militant, le partage de techniques de cracking ou le partage d'images sur la plateforme 4chan. Des raids massifs et spontanés (les *4chan raids*) popularisent une perception du *hacking* comme un mélange de militantisme politique, de blagues potaches, de raids satiristes et vengeurs, bien que l'action politique coordonnée ne soit pas encore à l'agenda de ces premières années (2003 à 2006). Cette ambiguïté causale sur les motivations et la destination de l'acte de rébellion numérique crée un « brouillard de guerre » opportun pour la conduite d'opérations sponsorisées et préparées. *Titan Rain* (2003-2006) est un exemple de ces premières cyber-guerres impliquant des attaques brutales peu sophistiquées, mais articulées dans des campagnes avancées (voir la figure 3).

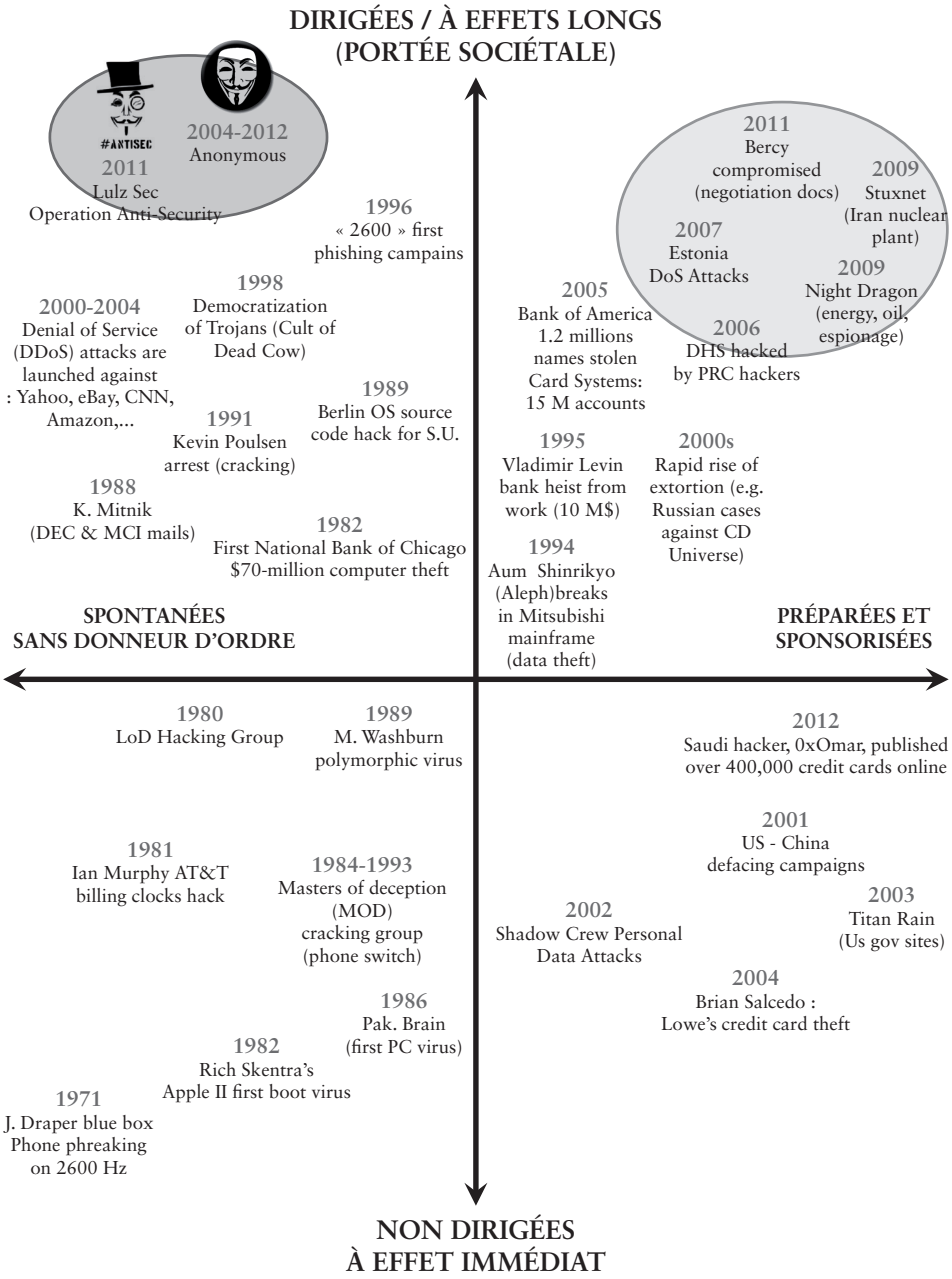
FIGURE 3
Les années 2000 : la monétisation du cyber-crime et les premières grandes confrontations



Les années de 2005 à 2013 marquent un double tournant, dans une certaine mesure, un schisme, entre d'un côté des campagnes ciblées étatiques et la naissance d'une grande criminalité organisée, et de l'autre des campagnes spontanées et virales menées par des groupes sociétaux aux articulations temporaires et incertaines. Cette période est caractérisée par des attaques de grande échelle contre des intérêts stratégiques (Kerem125 contre les Nations Unies, la campagne chinoise APT1, l'Estonie, Stuxnet, Aurora, etc. – Fig. 4).

Les technologies utilisées dans ces campagnes de grande échelle ne diffèrent pas des technologies pionnières utilisées dans les années 1980 et 90. 125 lignes de code sont toujours efficaces en 2013 pour conduire l'exploitation d'une vulnérabilité, même lorsque les lignes de défense consomment des efforts démesurés de programmation. Comme la plupart des innovations du début du XXI^{e} siècle, la performance de ces campagnes repose sur l'accessibilité et la diffusion d'apprentissages combinatoires, c'est-à-dire la capacité d'apprendre plus rapidement et plus systématiquement que peuvent le faire les cibles grâce à une meilleure intelligence comportementale.

FIGURE 4
Les années 2005-2013 : au-delà de l'objectif technique, l'émergence des campagnes d'attaques avancées



La formation de deux groupes distincts (le *hacking* de masse spontané *vs* le *hacking* avancé à maîtrise d'ouvrage) est d'ailleurs conforme aux deux stratégies qui permettent d'obtenir une supériorité d'apprentissage comportemental. Les grands ensembles humains spontanés (Anonymous, LulzSec) bénéficient de la coordination d'un apprentissage astucieux distribué, c'est-à-dire de la coordination numérique de l'apprentissage isolé d'*hackers*, lui conférant ainsi une ubiquité collective. Les campagnes d'attaques avancées d'État ou de la GCO, de leur côté, bénéficient des avancées technologiques permettant d'embarquer de l'apprentissage machine dans les campagnes d'attaques (ex. : APT, Stuxnet, FLAME).

La plupart des systèmes défensifs, face à ces ruptures, restent fondés sur la reconnaissance de signatures de codes malicieux, ou sur l'analyse normative des « bons comportements » (systèmes de détection à base de connaissances). Aussi bien l'apprentissage collectif des groupes spontanés de *hacking*, que les technologies avancées d'apprentissage machine, sont aujourd'hui capables d'un apprentissage supérieur à celui des systèmes de détection fondés sur des signatures. La nature de la rupture paradigmatique actuelle est, en ce sens, très similaire à celle que connut le renseignement dans les années 1990. Nous faisons face à une disruption stratégique où les défenseurs consolident leurs infrastructures d'informations, tandis que les attaquants mènent une « guerre de la connaissance »[7]. Une connaissance supérieure, par combinaison astucieuse, peut être obtenue à partir d'une information tronquée et partielle. Une information supérieure peut rarement défaire une véritable connaissance, même lorsqu'elle est pauvrement articulée.

En juin 2010, un code malicieux fut introduit dans le logiciel d'un composant *hardware* de la firme allemande Siemens destiné à intégrer le système de contrôle et d'acquisition de données (SCADA) d'un site d'enrichissement d'uranium à Natanz, en Iran. Ce code malicieux exploitait quatre vulnérabilités de Windows WinCC, dont on ignorait encore l'existence (vulnérabilité dite « *zero days* » ou « zéro jours »). Ce code malicieux permettait d'établir une communication externe, et de déclencher une instruction visant à paralyser, puis saboter, l'installation nucléaire dont il avait fait cible. Bien qu'une telle attaque ne constituait pas une première, elle possédait quelques caractéristiques qui ne sont pas étrangères au réveil des négociations internationales sur la régulation de la cyberdéfense, en général, et des contre-mesures, en particulier[8]. Stuxnet est ce qu'on appelle une campagne

7 P. Baumard, « From Information Warfare to Knowledge Warfare », in W. Schwartau (ed.), *Information Warfare: Chaos on the Electronic Superhighway*, New York, Thunder's Mouth Press, 1994, pp.611-626.

8 Pour une documentation technique, v. N. Falliere, L.O. Murchu and E. Chien, *W32.Stuxnet Dossier*, Symantec, 2011, disponible en ligne : <http://www.h4ckr.us/library/Documents/ICS_Events/Stuxnet%20Dossier%20(Symantec)%20v1.4.pdf> [Dernière visite, le 01/10/2014].

de menace avancée persistante, ou *Advanced persistent threat* (APT). De telles cyber-attaques sont ainsi nommées parce qu'elles reproduisent le comportement d'une attaque complexe, intelligente, avec des capacités de raisonnement et de déclenchement de commandes autonomes ou pilotées à distance. Cette famille de menaces n'est pas caractérisée par une technologie particulière, ni par une génération ou une typologie spécifique de technologies. Ses caractéristiques sont la programmation comportementale autonome (leur capacité à adapter leurs attaques), leur adaptabilité et leur persistance (après une période de reconnaissance de leurs cibles, elles ont pour objectif de compromettre un système en y résidant de manière anonyme, ou en trompant la vigilance des systèmes de détection en augmentant leurs privilèges). Leur caractérisation s'établit plus sur la subtilité de leurs stratégies de raisonnement, que sur leur force brutale, jusqu'à éventuellement mobiliser Sherlock Holmes pour en décrire les stratégies d'attaque ![9]

Stuxnet avait pour objectif de reprogrammer, à des fins de sabotage, les contrôleurs logiques (PLC) de la centrale nucléaire. Il était notamment capable d'autoréplication (en déclenchant sa propre installation à partir de supports externes amovibles), de s'exécuter à distance à travers un partage en réseau, de se mettre à jour par un réseau de *peer-to-peer*, de prendre le contrôle du centre de commandes, de cacher ses propres traces binaires, d'identifier les produits de sécurité résidant sur le réseau ainsi que de modifier puis de cacher les codes sabotés directement sur les contrôleurs industriels (PLC) de Siemens[10].

Cette dernière caractéristique, l'autoréplication associée à une élévation astucieuse de privilège, est une des causes de la formidable prolifération du Stuxnet, qui infecta près de 100 000 machines entre le jour de sa découverte (le 25 janvier 2010, si l'on ne prend pas en considération les versions dormantes de 2009) et le 29 septembre 2010, dont 14 sites industriels iraniens, et plus de 60 000 hôtes en Iran seulement[11]. Par son ampleur, la sophistication de sa construction qui peut effectivement signer l'œuvre d'un État, ou de la coopération entre plusieurs États, Stuxnet est ainsi identifié comme l'événement débutant la première cyberguerre globale[12]. Stuxnet est un révélateur de la possibilité d'une guerre, mais surtout

9 Pour une analyse détaillée de stratégies d'attaques utilisant des scénarios inspirés du héros fictif de Sir Conan Doyle, v. J. Ari and T.-F. Yen., « Sherlock Holmes and The Case of the Advanced Persistent Threat », in *Proceedings of the 5th USENIX conference on Large-Scale Exploits and Emergent Threats* (LEET'12), Berkeley, USENIX Association, 2012.

10 Falliere, Murchu and Chien, *supra* note 8, p.1.

11 *Ibid*, p.5.

12 J.P. Farwell and R. Rohozinski, « Stuxnet and the Future of Cyber War », *Survival: Global Politics and Strategy*, vol.53(1), 2011, pp.23-40. Pour un récapitulatif, v. J.R. Lindsay, « Stuxnet and the Limits of Cyber Warfare », *Security Studies*, vol.22(3), 2013, pp.365-404.

un très efficace démonstrateur des failles « systémiques » de l'industrie et des gouvernements : les systèmes de détection fondés sur l'identification des signatures de codes malveillants (repérage de codes malicieux dans un trafic de données) étaient inopérants face aux attaques intelligentes de ce type. De plus, le degré d'automation de Stuxnet (qui fut repéré par une entreprise de sécurité bélarusse car les machines n'arrêtaient pas de se relancer et de se mettre à jour !) faisaient voler en éclat le sentiment de quiétude quant à une « sécurité maîtrisée ». Schouwenberg, qui fut à la tête de l'équipe de Kapersky ayant contribué à défaire Stuxnet, était sincèrement impressionné par l'élégance de sa programmation, qui combinait l'exploitation croisée de quatre vulnérabilités « zéro jours »[13] avec des déclencheurs logiques particulièrement astucieux. La cyberguerre devenait une réalité palpable, et le cyberespace un domaine physique[14] avec des conséquences sociétales et stratégiques facilement mesurables. Et comme dans *Le rivage des Syrtes* de Julien Gracq, rien ne donne plus envie aux hommes de faire la guerre qu'un faible signal lumineux sur une rive éloignée.

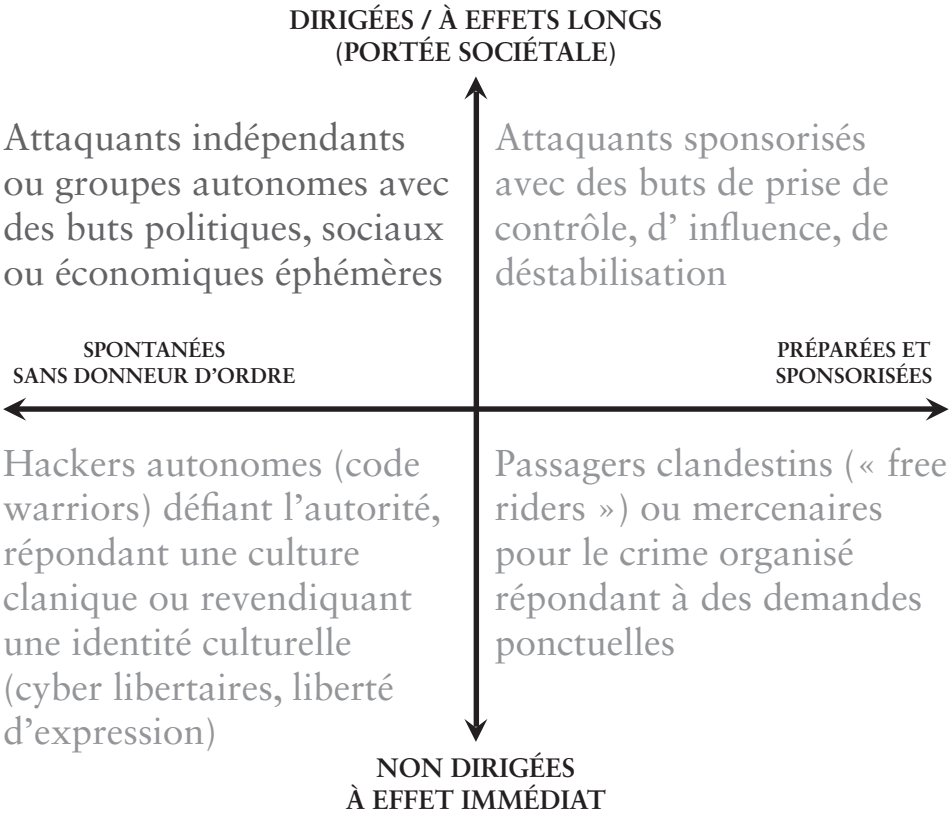
Ce nouveau paradigme, fondé sur l'intelligence comportementale, est synonyme d'une escalade du nombre d'attaques « *zero days* ». Un apprentissage pervasif et disponible à coût marginal permet de créer des variantes d'exploitation de vulnérabilités (« exploits ») à un rythme plus soutenu que celui de la production de signatures de codes malicieux, et cela même dans un délai de 24 heures. La ré-encapsulation et la recombinaison d'exploitations de failles non découvertes (« *zero days* ») est rendue possible par l'état d'avancement des techniques d'apprentissage causatifs (Bayesien, AI) ou lorsqu'elles ne sont pas disponibles, par la coordination de groupes de *hacking* spontanés menant des expérimentations de recombinaison. Dans un tel paradigme, se focaliser sur une stratégie de défense *ex post*, s'appuyant sur la détection de vulnérabilités déjà connues et compilées, est simplement désastreux.

13 D. Kushner, « The Real Story of Stuxnet », *IEEE Spectrum*, vol.50(3), 2013, pp.48-53.

14 La doctrine du cyberespace comme « domaine physique » a été initiée aux États-Unis par Daniel Kuehl en 1997 : I*nformation as an environment may be a difficult concept to grasp, but there is no arguing that there is a physical environment to which information is uniquely related: cyberspace. Cyberspace is that place where computers, communications systems, and those devices that operate via radiated energy in the electromagnetic spectrum meet and interact*. D. Kuehl, « Defining Information Power, » *Strategic Forum*, n°115, 1997, 1, p.3.

II. LA MISE À L'ÉPREUVE DES DOCTRINES CONTEMPORAINES FACE AU CHANGEMENT TECHNIQUE

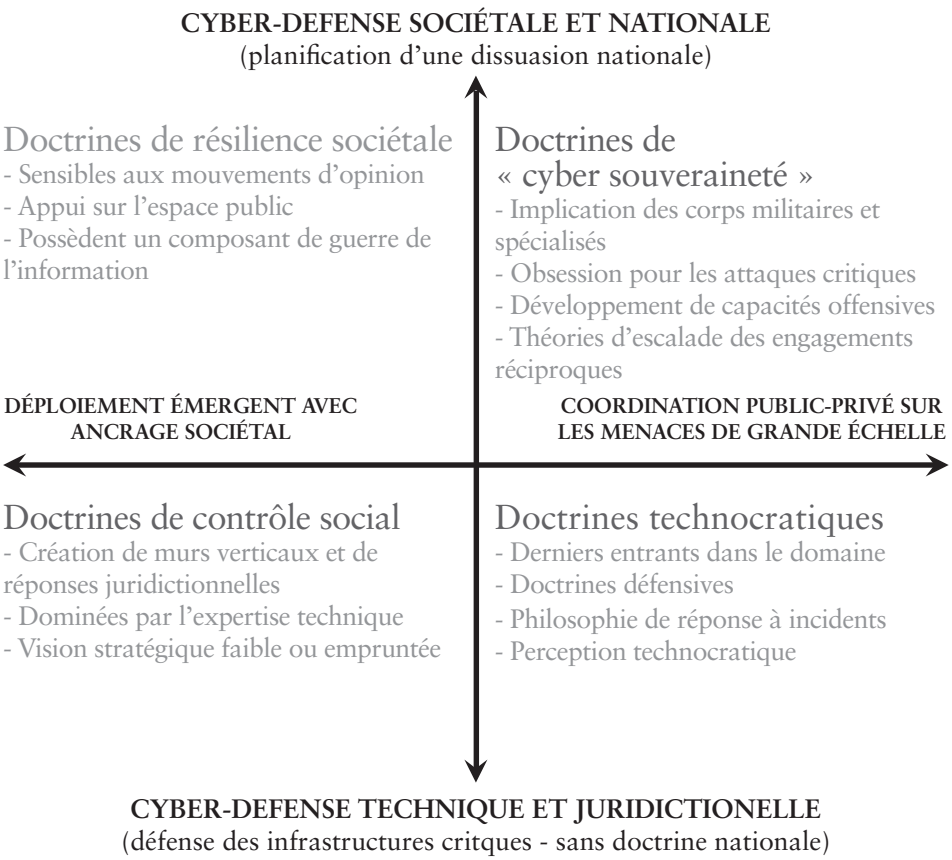
Dans cette seconde partie, nous essayons d'évaluer la robustesse des stratégies nationales de prévention et de contre-mesures contre la cybercriminalité face à l'évolution technologique des attaques. Pour ce faire, nous avons analysé, à partir de documents publics, 38 doctrines et systèmes nationaux de mise en œuvre de cyber-défense, de résilience informationnelle et de cyber-sécurité. Nous utiliserons le même cadre d'analyse, développé précédemment, pour analyser l'histoire des menaces cybercriminelles, en utilisant leur destination (« ciblées et de longue portée » c. « immédiates et non dirigées »), et leur degré de préparation (« spontanées » c. « préparées et sponsorisées »). Nous avons ainsi identifié quatre classes de cybercriminalité : les « guerriers du code » (I), les « passagers clandestins » (II), les « collectifs autonomes » (III) et les « attaquants sponsorisés » (IV).



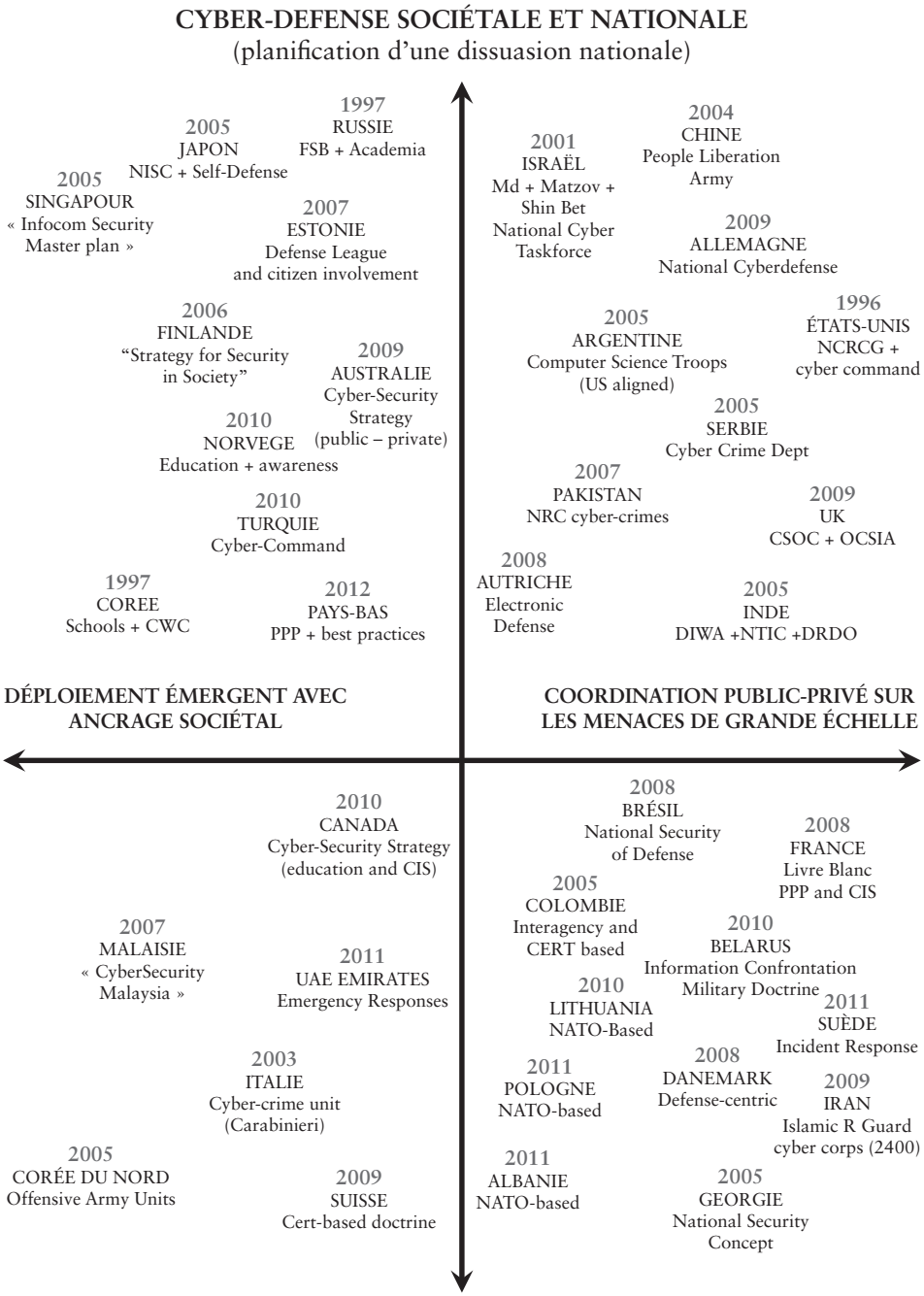
Différentes classes d'attaques demandent des réponses différentes. Les attaques spontanées et à effets immédiats (classe I) peuvent être contenues avec une sécurité de l'information robuste, qui peut inclure des systèmes d'apprentissage pour contingenter des attaques encapsulant de l'intelligence artificielle. La plupart des doctrines nationales ont une compréhension mature et des réponses appropriées pour ce type d'attaques. Les attaques « sur commande » à effets courts (crime informatique organisé, vol de capacités, *phishing* et *cracking* – classe II) demande la coordination de réponses techniques inter-juridictionnelles. Des systèmes de détection à base de signatures, ou à bases de données comportementales, sont généralement suffisants pour contrecarrer ce type d'attaques, si la législation peut être appliquée. Les attaques « sociétales » (hacktivistes, groupes temporaires motivés par des enjeux politiques, sociétaux ou économiques – classe III) demande un engagement dans des guerres de perception, des guerres de l'information et des capacités d'attribution de sens pour répondre à leur déploiement émergent et distribué. Finalement, les campagnes offensives coordonnées à intelligence comportementale (classe IV) requièrent des réponses transversales, qui incluent une dissuasion proactive « au-delà du simple enjeu technique » et « au-delà des revendications de surface ». Les menaces de classes III et IV réclament des capacités d'interprétation à très large échelle, impliquant soit un apprentissage cognitif capable d'assimiler de larges quantités de renseignement humain (III), soit de l'apprentissage machine capable de surveiller le comportement de très larges réseaux (IV).

Notre analyse de l'évolution des doctrines de lutte contre la cybercriminalité sur la période 1994-2013 donne des résultats mitigés. Les doctrines de « puissance – souveraineté » (classe IV), mettant généralement l'emphase sur le développement de larges unités spécialisées, sont obnubilées par la protection des infrastructures critiques, et développent, plus ou moins officiellement, des capacités offensives. Bien qu'elles puissent délivrer des politiques de dissuasion efficaces envers les attaques lancées par des États, elles entraînent aussi une rigidité face aux menaces émergentes, qui poussent les gouvernements qui en sont porteurs à délaisser leur implication dans le changement sociétal. Le risque de ces doctrines est d'être déconnecté des mouvements de *hacking* émergents, et de manquer de réactivité face à des guerres cognitives distribuées. Les doctrines à dominante « résilience sociétale » (Classe III), de leur côté, sont plus sensibles aux mouvements d'opinion, essaye de tirer parti de l'espace public et se focalise sur l'appréhension de possibles guerres de l'information. La motivation pour de telles doctrines n'est pas, pour autant, toujours ancrée dans des aspirations de défense des libertés individuelles. La numérisation de la société y est clairement identifiée comme à la fois une menace et une opportunité pour la cyberdéfense et le cyberdéveloppement. Finalement, les doctrines de classes I et II (« ordre social » et « technocratique ») ne diffèrent entre elles que sur la perception

qu'elles ont de la notion de contrôle. Les premières favorisent un contrôle à la source, tandis que les secondes favorisent un contrôle à l'aval (normes de comportement, normes techniques). Les doctrines technocratiques souffrent d'une perception différée des changements technologiques, sont principalement inspirées par une philosophie de réponse aux incidents, et sont plutôt le fait d'entrants tardifs dans la discipline. Les doctrines favorisant l'ordre social souffrent généralement d'un manque de vision nationale, ou ont bâti leurs politiques par emprunt (ou par alignement) sur des visions nationales importées.



Le graphique suivant illustre le positionnement des 38 stratégies nationales de lutte contre la cybercriminalité et de conduite de la cyber-défense (l'année indique la date du premier document étudié). Les résultats illustrent l'arbitrage entre des politiques nationales qui se concentrent sur une vision régalienne de la cyber-défense de celles qui sont guidées par une volonté d'implication, de surveillance (ou de soutien) des racines sociétales de la cybercriminalité. Il est intéressant de noter que les doctrines russes étudiées se situent plus sur le versant sociétal que leurs contreparties nord-américaines ou chinoises. On retrouve cette différence dans les tensions récurrentes qui opposent les régulateurs russes et leurs homologues européens ou américains. D'une part, les représentants russes refusent de dissocier la sécurité de l'information, prise dans son sens le plus large, et incluant donc l'information civile au même titre que l'information d'État ; d'autre part, la notion de « cybersécurité » est rejetée comme une construction politique visant à imposer une neutralité de l'information, conçue comme un « stock », indépendante de la sécurité des « systèmes », créés comme des moyens de transport et de traitement de l'information. La négociation d'un cadre de régulation international et partagé devient dès lors une confrontation doctrinale entre une perception « pan-technologique » de la régulation (anglo-saxonne) et une perception sociétale, ne dissociant pas la sûreté de l'État de celle de l'information prise dans sa totalité. On retrouve cette distinction sur le graphique suivant, où la Russie, tout en poursuivant une stratégie de cyber-défense nationale, se situe sur le versant gauche du graphique (déploiement émergent avec un ancrage sociétal).


CYBER-DEFENSE SOCIÉTALE ET NATIONALE
(planification d'une dissuasion nationale)
1997 RUSSIE FSB + Academia
2005 JAPON NISC + Self-Defense
2005 SINGAPOUR « Infocom Security Master plan »
2007 ESTONIE Defense League and citizen involvement
2004 CHINE People Liberation Army
2001 ISRAËL Md + Matzov + Shin Bet National Cyber Taskforce
2009 ALLEMAGNE National Cyberdefense
2006 FINLANDE "Strategy for Security in Society"
2009 AUSTRALIE Cyber-Security Strategy (public – private)
2010 NORVEGE Education + awareness
2005 ARGENTINE Computer Science Troops (US aligned)
1996 ÉTATS-UNIS NCRCG + cyber command
2005 SERBIE Cyber Crime Dept
2007 PAKISTAN NRC cyber-crimes
2009 UK CSOC + OCSIA
2010 TURQUIE Cyber-Command
2008 AUTRICHE Electronic Defense
1997 COREE Schools + CWC
2012 PAYS-BAS PPP + best practices
2005 INDE DIWA +NTIC +DRDO
DÉPLOIEMENT ÉMERGENT AVEC ANCRAGE SOCIÉTAL
COORDINATION PUBLIC-PRIVÉ SUR LES MENACES DE GRANDE ÉCHELLE
2010 CANADA Cyber-Security Strategy (education and CIS)
2008 BRÉSIL National Security of Defense
2008 FRANCE Livre Blanc PPP and CIS
2005 COLOMBIE Interagency and CERT based
2007 MALAISIE « CyberSecurity Malaysia »
2011 UAE EMIRATES Emergency Responses
2010 BELARUS Information Confrontation Military Doctrine
2010 LITHUANIA NATO-Based
2011 SUÈDE Incident Response
2003 ITALIE Cyber-crime unit (Carabinieri)
2008 DANEMARK Defense-centric
2011 POLOGNE NATO-based
2009 IRAN Islamic R Guard cyber corps (2400)
2005 CORÉE DU NORD Offensive Army Units
2009 SUISSE Cert-based doctrine
2011 ALBANIE NATO-based
2005 GEORGIE National Security Concept
CYBER-DEFENSE TECHNIQUE ET JURIDICTIONELLE
(défense des infrastructures critques - sans doctrine nationale)

La plupart des stratégies nationales étudiées qui appuient leurs politiques nationales sur des fondements techniques et juridictionnels répondent à un « retard ressenti » du point de vue technique. Elles sont généralement ancrées dans une vision de contingentement technique, et favorisent la coordination technique et juridictionnelle. La plupart des difficultés sociétales liées au numérique sont « rattrapées » par des mises à jour doctrinaires bien postérieures à leur émergence. En somme, les doctrines continuent à s'ancrer dans le quart supérieur droit (classe IV), quand les innovations de rupture émergent dans le quart supérieur gauche (classe III).

Les techniques utilisées par la cybercriminalité sont stables sur la période allant de 1990 à 2012. Les menaces persistantes avancées (APT) ne sont pas le résultat d'une disruption dans l'exploitation de nouvelles vulnérabilités, mais plutôt le produit d'un changement paradigmatique dans les technologies périphériques (principalement : l'apprentissage machine, l'automation, la reconfiguration combinatoire). De tels changements paradigmatiques prolifèrent grâce à l'obsolescence des infrastructures si vieillissantes. L'exploitation de vulnérabilités génériques permet la construction de combinaisons avancées. L'interopérabilité, en surcouche de systèmes vieillissants, accroît la probabilité d'une exploitation à la volée de telles vulnérabilités. Dans un tel contexte, les fabricants de logiciels, en poussant des solutions focalisant sur la vulnérabilité des points d'accès (IPS, IDS), ralentissent l'investissement dans les technologies d'apprentissage comportemental avancé (en maintenant des systèmes à performance réduite, mais forte profitabilité, de détection basé sur les signatures).

III. LA RÉPONSE DES ÉTATS ET LES ÉVOLUTIONS DE LA RÉGULATION

Très peu des doctrines étudiées identifient les retards technologiques comme des vulnérabilités systémiques. Les mesures de confiance et de sécurité (CSBMs) sont ainsi adossées à une perception technique qui augmentent les vulnérabilités plutôt que les réduire, et souffrent généralement d'un angle mort, aussi bien en matière d'attribution que de détection, sur les attaques à intelligence comportementale avancée (AI, *morphing*, etc.). Les doctrines de classe II (technocratique) et I (ordre social) produisent une connaissance verticale et juridictionnelle (silos) tandis que l'évolution des menaces est horizontale (AI) et transversale. Les campagnes avancées récentes (APT1, Blaster-worm, etc.) ont montré les limites de la coopération juridictionnelle dans la réponse à des attaques aux attributions le plus souvent impossibles, porteuses d'« exploits » inconnus ou non découverts (« *zero day*s ») et utilisant de l'apprentissage causatif pour s'adapter aux défenses techniques communes.

La plupart des doctrines analysées présentent ainsi une perception datée de l'identification et de l'attribution. L'attribution est assimilée dans la plupart des doctrines avec un point géographique (ou plusieurs), une intention centrale, et une perspective légaliste de la traçabilité. L'effacement des traces d'intrusion est maîtrisé depuis longtemps par les attaquants, ne pouvant mener qu'à la conclusion que les efforts diplomatiques sont tournés vers la résolution d'un problème qui a perdu sa pertinence technique bien avant 2007.

L'objet d'un cadre de régulation partagé entre États est de faciliter une lutte coordonnée contre la cybercriminalité par le partage d'information, l'accord mutuel d'accès dans le cadre d'enquêtes internationales, mais également d'établir une norme comportementale partagée afin de prévenir l'escalade conflictuelle lors de confrontations pouvant engager la souveraineté des cosignataires. Ces cadres de régulation concernent aussi bien la société civile, les entreprises que les États. La négociation d'un cadre régulé de réponses aux menaces cyber-criminelles (les « contre-mesures ») répercutent les différends engendrés par les désaccords doctrinaires que nous avons évoqués plus haut, et reflètent les retards techniques respectifs des différentes nations participant à ces négociations.

IV. L'INCAPACITÉ D'UN ACCORD DE RÉGULATION SUR DES FONDEMENTS TECHNIQUES

Il n'existe pas de régulation internationale globalement ratifiée, en 2014, permettant d'établir quelles sont les règles d'engagement, les fondements de la légitimité de telles contre-mesures, et ne serait-ce même qu'un accord de principe entre les nations sur l'idée que la proportionnalité et la « contre-offensive » puissent être légalement justifiées[15]. Les contre-mesures peuvent être soit de nature logicielle, soit de nature humaine, et sont le plus souvent composées de ces deux types d'intervention. Une contre-mesure logicielle est un ensemble de lignes de commandes déclenché à partir d'un système de détection reposant sur un raisonnement analytique programmé (intelligence artificielle, corrélation statistique, bayésiens, etc.). La contre-mesure « humaine » consiste à mener le travail d'investigations et d'enquêtes (audit, forensics) par des moyens qui peuvent être humains (*social engineering*, entretiens, enquêtes sous couvert) ou informatiques (traçage inverse du chemin d'attaque, tests de pénétration, etc.). Cette seconde partie peut impliquer de retracer le chemin d'une attaque qui est passée par des serveurs physiquement localisés dans un ou plusieurs pays étrangers.

15 V. D. Fleck, « Searching for International Rules Applicable to Cyber Warfare - A Critical First Assessment of the New Tallinn Manual », *Journal of Conflict & Security Law*, vol.18(2), 2013, pp.331-351.

Le problème pourrait être anodin ou devenir la simple prolongation des traités internationaux encadrant la criminalité, la guerre, les trafics humains et le grand banditisme si la nature même de l'information et des systèmes d'information ne rendait pas impossible cette simple évolution.

D'une part, il est difficile d'établir une destination criminelle *a priori* d'un système d'information ou d'une information elle-même. Par exemple, le recours au secret et au chiffrement peut être une garantie du respect des libertés individuelles et de l'anonymat tout autant que le signe d'une activité criminelle. De fait, les lanceurs d'alerte utilisent les mêmes technologies que les organisations criminelles pour rendre anonymes leurs communications et échanges (*Tor network*, *darkpools*). Le « cyberlibertaire », le *whistleblower*, le « résistant », le « rebelle » ou le « cyberterroriste » sont souvent une même personne selon qu'on la perçoit d'un côté différent de la barrière. On ne peut pas identifier l'intention de malveillance à partir de « l'arme » en possession du « cyberattaquant » pour la bonne et simple raison qu'une commande utilisée sur un réseau n'est pas une arme. Leur combinatoire complexe peut en devenir une, mais les techniques d'offuscation de code (qui consiste à masquer un code dans une forêt de codes placebo mais non pour autant dénués de sens) et de chiffrement sont si librement accessibles, que le travail d'audit et d'investigation scientifique d'un code complexe peut prendre plusieurs semaines. Une approche fonctionnelle et typologique du caractère offensif ou criminel d'une information ou d'un système d'information est donc vouée à l'échec[16].

D'autre part, l'architecture même des systèmes d'information, pour des raisons d'efficience, n'est pas délimitée par ses finalités mais par la performance d'opérations qui peuvent être distribuées, parcellaires, partagées, voire « fragmentées » de façon aléatoire des systèmes *peer-to-peer*. Dès lors, la géographie de l'information, de sa destination comme de sa création, devient le résultat d'un processus technique qui va permettre de dissimuler, que cela soit intentionnel ou pas, son origine, son but, sa chronologie et son propriétaire[17].

V. ANONYMAT, ATTRIBUTION ET PREUVE

16 V. E. Talbot Jensen, « Cyber Warfare and Precautions Against the Effects of Attacks », *Texas Law Review*, vol.88, 2010, pp.1533-1570.

17 M. Van Eeten et al., « The State and the Threat of Cascading Failure across Critical Infrastructures: The Implications of Empirical Evidence from Media Incident Reports », *Public Administration*, vol.89(2), 2011, pp.381-400.

D'INTENTIONNALITÉ

La difficulté réside dans l'interprétation de la responsabilité des différentes parties prenantes (opérateurs de télécommunications, États, fabricants de logiciels, etc.), dans l'expression finale qui peut résulter de ces cheminements complexes d'information. Dans l'article 32 de la convention de Budapest de 2001, l'obligation de communication d'information revient à la personne qui en a l'autorité légale. Mais est-ce la personne qui a accédé à cette information ? Celle qui a fourni le moyen de transport ? Celle dont le pays est à l'origine du robot logiciel qui a assuré l'échange anonyme de pair-à-pair (P2P) de cette information ? Et dans le cas d'une attaque robotique, doit-on punir le fabricant du robot, le ou les différents pays ayant laissé ce robot opérer, ou simplement le dernier utilisateur « conscient » de ce robot... ou bien le robot lui-même ? (Interdiction par la catégorisation du logiciel robotisé en « arme de guerre »).

Ce travail de réflexion est mené par le Conseil de l'Europe, qui coordonne l'évolution de la Convention de Budapest sur le cybercrime de 2001, signée par 60 pays ; mais avec, selon l'Office des Nations Unies contre la drogue et le crime (UNODC), plus d'un tiers de la population mondiale en situation de connectivité, et une prévision de connectivité à 70 % de la population planétaire en 2017, la question des « contre-mesures » et de l'autorité à les conduire devient un enjeu qui dépasse de très loin la problématique de la coopération transfrontalière. La duplication ou la transposition par allégorie (maritime, territoriale, spatiale) d'un droit international quelconque à un droit de la cyberdéfense est illusoire, sinon utopique.

Avec plus de 70 % des communications mondiales en 2013 étant de machine-à-machine sans intermédiation humaine, il peut être dangereux, du point de vue de la simple efficacité du droit, d'attribuer à ce monde de machines les règles d'encadrement et de contrôle du crime humain. D'autre part, la vitesse des micro transactions et la célérité avec laquelle une attaque peut être métamorphosée en une soixantaine de ses variantes, ne laissent aucune avance au législateur, à l'exécutif et aux forces de l'ordre pour résoudre des enquêtes judiciaires si les règles du jeu sont celles de la criminalité courante[18]. Le problème est alors double : on ne peut exercer la loi car elle n'aura aucun terrain viable d'exercice ; et lorsque ce terrain est rendu possible, on ne peut financer son déploiement ou le simple devoir d'enquête. Que ce soit en Europe, aux États-Unis, en Russie, en Chine, en Afrique ou au Brésil, très peu de cas de « cyber-crimes » ont été effectivement traités par le système judiciaire,

18 K. Geers, « The challenge of cyber attack deterrence », *Computer Law & Security Review*, vol.26(3), 2010, pp.298-303.

et beaucoup moins sont sanctionnés[19]. Les mécanismes de coopération judiciaire, d'enquête et de police sont très souvent inopérants en matière de cybercriminalité, non pas par déficit de talents, car contre une idée reçue les forces de l'ordre disposent de talents décisifs dans ce domaine, mais par le caractère impraticable de la coopération technique sur la cybercriminalité[20].

Partager un « cas » de cybercriminalité revient à partager une vulnérabilité de son système de défense car la frontière de l'innovation bouge à chaque cas. Dès lors, si un État coopère avec un autre, il refusera de le faire sur les « zéro jours » (attaques dont la faille n'a pas été identifiée), de peur de révéler à d'autres États son état de l'art défensif, son réel degré de maîtrise de la cyberdéfense, voire une faille de cybercriminalité qui pourrait être exploitée à d'autres fins (espionnage d'État, cyberguerre). Le « zéro jours » est tout autant une vulnérabilité avouée dans ses propres défenses que la concrétisation d'une opportunité offensive, si l'on possède la solution pour résoudre, ou si l'on peut exploiter cette vulnérabilité chez un adversaire[21]. Comme chaque acteur cherche à créer des asymétries décisives face à des puissances étrangères qui ne sont pas toujours bienveillantes, il n'existe aucune incitation à divulguer une information sur ce type de vulnérabilités et de possibilités d'exploitation de failles. Dès lors, il est impossible aujourd'hui de résoudre ou répondre à une attaque d'origine transfrontalière (dont le cheminement implique plusieurs pays), tant l'accès transfrontalier à la donnée *réelle* est difficile[22]. Il ne s'agit pas ici des données elles-mêmes de l'attaque (le code de celle-ci), mais aussi des données concernant ses moyens de transport, des serveurs sur lesquels elle a pu laissé des traces exploitables et des machines qui ont, même involontairement, servi de vecteurs de diffusion et d'attaque[23].

Contrairement aux armements traditionnels, les « armes » cybernétiques (mis entre guillemets tant cette expression est abusive) ne font pas l'objet d'une manufacture spécifique, d'un numéro de série ou de l'utilisation d'outils de production spécialisés. L'énergie motrice d'une attaque, son « fuel » pour prendre une analogie

19 J. Goldsmith, « How Cyber Changes the Laws of War », *European Journal of International Law*, vol.24(1), 2013, pp.129-138.

20 V. C.C. Demchak and P. Dombrowski, « Rise of a Cybered Westphalian Age », *Strategic Studies Quarterly*, vol.5(1), 2011, pp.31-62.

21 F. Li, A. Lai and D Ddl, « Evidence of Advanced Persistent Threat: A Case Study of Malware for Political Espionage », *6th International Conference on Malicious and Unwanted Software* (Malware 11), IEEE, 2011, pp.102-109.

22 D. Denning, « Reflections on Cyberweapons Controls », *Computer Security Journal*, vol.16(4), 2000, pp.43-53.

23 P. Brunst, « Legal Aspects of Cyber Terrorism », in Centre of Excellence – Defence Against Terrorism (ed.), *Legal Aspects of Combating Terrorism*, Amsterdam, IOS Press, 2008, pp.63-76.

anglo-saxonne, peut par exemple résider dans la bande passante que le système d'attaque a pu capturer. Il s'agit des attaques en DDoS[24] dont l'objectif est de générer un emballement de requêtes sur les machines visées afin de provoquer leur panne par saturation. Ce type d'attaque correspond le plus souvent au troisième des *Trente-Six Stratagèmes*[25], c'est-à-dire à emprunter l'épée pour défaire l'adversaire, et il n'y a pas meilleure « épée » numérique que la bande passante anonyme, sous-utilisée et souvent mal protégée que l'on trouve dans le foyer de chacun, dans des établissements publics mal sécurisés (universités), avant même d'avoir recours à un service criminel spécialisé de location de telles capacités. Ces attaques se multiplient dès la fin des années 2000 visant les serveurs racines du Département de Défense (DoD) américain, les serveurs de l'Icann (février 2007), les systèmes de paiement, les systèmes de sécurité ou, comme ce fut le cas en Estonie en avril 2007, les services d'urgence. Les régulateurs sont prompts à attribuer ces attaques à des groupes organisés, voire à la criminalité organisée ou à des États adverses, mais la réalité technique est très différente de ses projections étatiques. Il est donc possible qu'une cyberattaque, même menée par un groupuscule par ailleurs financé par un gouvernement, repose intégralement sur des capacités (serveurs, bande passante) « empruntées » à des clients tout à fait innocents, et complètement ignorants d'avoir contribué au façonnage global d'une arme de guerre cybernétique.

Les cyberattaques résistent ainsi aux typologies et aux « classes » qu'on aimerait leur attribuer. Une attaque en DDoS peut avoir pour finalité sa fonction première : noyer des serveurs de requête afin de les mettre en berne, pour le *fun*, pour la revendication politique, pour une bravade entre *hackers*… ou pour déstabiliser un État. Il n'y a pas de réelle corrélation entre l'ampleur des moyens mobilisés, le « faciès numérique » du *cracker*/*phracker*/*hacker*/cyberterroriste, qui sont, là encore, des dénominations qui n'engagent que l'éthique personnelle de ceux qui veulent bien les porter. La différence entre le *hacker* et le *cracker*, termes hérités des années pionnières du *hacking*, ne tient qu'à une volonté individuelle d'utiliser, ou pas, des moyens techniques hors d'un cadre et d'un propos éthiques. On n'est pas *hacker* par le type de technique, ou de signatures de code, que l'on laisse derrière soi, mais par la destination que l'on fait de son code, et la vision du monde qu'on lui attache.

24 *Distributed Denial of Service : Déni de service distribué*.

25 F. Kircher (trad. par F. Kircher, calligraphies de A. Huchant) (éd.), *Les trente-six Stratagèmes : traité secret de stratégie chinoise*, Paris, Payot & Rivages, 1995, pp.420-479.

VI. UN DÉSACCORD SUR LA NOTION DE "FORCE"

Au-delà de la capacité d'anonymat de l'attaquant, le caractère indifférencié du corps d'expertise technique, des moyens technologiques, et même du support culturel (ou culture alternative) de ce qu'on appelle des attaques cybernétiques interdit donc tout catégorisation *ex ante*. Or, la régulation contemporaine de la guerre repose, avec l'article 2(4) de la Charte des Nations Unies, sur la catégorisation *ex ante* de l'usage de la force :

> « Les membres de l'Organisation s'abstiennent, dans leurs relations internationales, de recourir à la menace ou à l'emploi de la force, soit contre l'intégrité territoriale ou l'indépendance politique de tout autre État, soit de toute autre manière incompatible avec les buts des Nations Unies »[26].

L'article 2(4) est complété par l'article 51 de la Charte qui stipule que :

> « [a]ucune disposition de la présente Charte ne porte atteinte au droit naturel de légitime défense, individuelle ou collective, dans le cas où un Membre des Nations Unies est l'objet d'une agression armée »[27].

Qu'est-ce que la « force » en matière cybernétique ? Et à partir de quel moment, de quelle quantité de bande passante, de quelle caractérisation de l'attaque, peut-on parler de « force armée » cybernétique ?

Le préfixe « cyber » est lui-même utilisé avec tant de déclinaisons, que les traités internationaux, s'intéressant à la régulation des espaces numériques, ont échoué à se mettre d'accord sur sa signification ou son étymologie. Le terme cybernétique est bien entendu un héritage des travaux des pionniers de cette discipline, comme Miller, Gallanter et Pribram[28], et dont l'origine, dans les années 1940, est la volonté

26 Charte des Nations Unies, art.2-§-4.

27 *Ibid*, art.51. V. O. Schachter, « The Right of States To Use Armed Force », *Michigan Law Review*, vol.82(5/6), 1984, 1620, p.1624.

28 G.A. Miller, E. Galanter and K.H. Pribram, *Plans and the Structure of Behavior*, New York, Rinehart and Winston, 1960.

d'une étude systématique et scientifique de l'interaction du vivant et des machines[29]. Son projet initial, dans les pas d'Alan Turing, est de montrer qu'une manipulation symbolique autonome peut être conduite par une machine. Déjà, Polanyi posait dans son article de 1952 la question de l'expérience de la conscience et du jugement responsable qu'il considère comme un critère beaucoup plus pertinent que l'opposition pensée – calcul pour opposer hommes et machines[30]. Il rejoignait en cela les réflexions de McKay, dans le même volume qui affirmait que la libre volonté des hommes ne pouvait être assimilée à l'indétermination des machines. Un système cybernétique est donc, dans cette perspective, un système capable d'un comportement délibéré poursuivant un propos conscient, que ce dernier soit le résultat de sa propre évolution ou qu'il lui ait été appris par l'homme. La cybernétique de Wiener, dont il invente le terme en ignorant qu'Ampère l'a déjà utilisé en 1873 pour désigner l'art de gouverner, n'était pas une science du contrôle et du signal, mais plutôt une science du gouvernement de la vie par la maîtrise de ses communications, humaines et/ou machines[31].

Une « arme cybernétique », si on s'en réfère à l'étymologie, est donc un dispositif visant à pervertir ou détruire les systèmes de commandement des systèmes de support du vivant et/ou les machines les composant. On peut d'ores et déjà noter que cette définition ne présuppose pas l'implication d'un pilotage ou d'une commande directe de nature humaine. De fait, les avancées dans le comportement artificiel et le raisonnement autonome des machines permettent aujourd'hui d'envisager l'existence d'un système cybernétique totalement autonome, conduisant aussi bien des opérations offensives que défensives[32]. En ce sens, les attaques DDoS sont bien des agressions cybernétiques, mais toute agression utilisant une technologie de l'information ne peut pas être définie comme une « cyberattaque ». Un acte d'espionnage, qu'il soit mené par des moyens humains (*humint*) ou électroniques (*sigint*) reste un vol ou un détournement d'information. Ainsi, la définition de la Maison Blanche, qui englobe toute organisation poursuivant des buts criminels

29 Le terme « cybernétique » a pour origine les articles suivants : A. Rosenblueth, N. Wiener and J. Bigelow, « Behaviour, Purpose and Teleology », *Philosophy of Science*, vol.10, 1943, S. 18-24 ; W.S. McCulloch and W. Pitts, « A Logical Calculus of the Ideas Immanent in Nervous Activity », *Bulletin of Mathematical Biophysics*, vol.5, 1943, pp.115-133 ; Pour une note de la naissance de la cybernétique : M. Polanyi, « The Hypothesis of Cybernetics », *British Journal for the Philosophy of Science*, Vol.2(8), 1952, pp.312-315 ; L'article offre une synthèse d'un débat sur la cybernétique menée dans le même journal par K.R. Popper, vol.1, pp.194-195 ; J.O. Wisdom, vol.2, p. 1 ; D.M. MacKay, vol 2, p.120 ; F.M. Walshe, vol 2, pp.161-163 ; W. Mays, vol 2, pp. 249-250.

30 Polanyi, *supra* note 29, p.315.

31 Lire à ce propos : D.J. Clark, « Enclosing the Field : from "Mechanization of Thought Processes" to "Autonomics" », thèse de doctorat, Université de Warwick, sept. 2002, pp.88-91.

32 P. Baumard, « Using Machine Foreknowledge to Enhance Human Cognition », in O. Pourret, P. Naïm and B. Marcot (eds.), *Bayesian Belief Networks: A Practical Guide to Applications*, New York, Wiley, 2008, pp.365-375.

par le bais de technologies de l'information ou en s'y connectant, tend à assimiler toute nouvelle forme de criminalité à de la « cybercriminalité », y compris les manipulations d'opinion, les déstabilisations par des campagnes de guerres de l'information, ou le vol de propriété intellectuelle, qui sont très éloignés d'une activité « cybernétique » *stricto sensu*. Ainsi, dès sa création en novembre 2001, la convention de Budapest sur la cybercriminalité intègre les actions de vol d'identité, les attaques contre l'intégrité d'un système, le vol de données, la fraude, les contenus offensifs, la violation de *copyright*, les crimes « informationnels » de natures xénophobes et racistes, aussi bien que le « cyber-terrorisme », défini par la convention de Budapest comme l'usage de technologies de l'information pour propager la terreur[33].

On retrouve dans la convention de Budapest cette velléité de faire de la guerre de l'information une « guerre de la connaissance »[34] qui est déjà dans l'arrière-plan des doctrines de cyber-défense américaines à la fin des années 1990. Cette volonté d'imposer une *Pax Cybernetica* américaine éloigna un peu plus la convention de Budapest de la lisibilité nécessaire à établir un usage de la force selon l'article 2(4) de la Charte des Nations Unies. À vouloir créer un accord cadre qui couvre tous les aspects du cyberespace, de l'expression de la violence symbolique, jusqu'à l'acte de sabotage des infrastructures vitales d'un pays, la convention de Budapest a échoué à échoué à atteindre un but qui était inconciliable : celui de mettre sur une même échelle de force le voyeurisme numérique, la pornographie, la liberté d'expression et l'attaque en sabotage d'une centrale nucléaire[35]. Là encore, la régulation bute non pas sur la question de l'attribution, mais sur celle de l'établissement de preuve d'une intentionnalité. Le résultat est un cadre de régulation qui substitue la coercition à l'observation réelle et factuelle de la force comme légitimation de l'intervention. Dès lors, un tel cadre de régulation est prompt à instaurer un climat de Guerre froide[36]. Dans le cas de Stuxnet ou de Flame, deux campagnes d'attaques avancées avec des objectifs industriels précis, on peut effectivement parler d'un

33 T. Remus, « Cyber Attacks and International Law of Armed Conflicts: A "Jus Ad Bellum" Perspective », *Journal of International Commercial Law and Technology*, vol.8(3), 2013, pp.179-189.

34 P. Baumard, « From Information Warfare to Knowledge Warfare: Preparing for the Paradigm Shift », in A.D. Campen, D.H. Dearth and R.T. Goodden (eds.), *Cyberwar: Security, Strategy, and Conflict in the Information Age*, Fairfax, AFCEA International Press, 1996, pp.147-160. Lire également dans la meme collection : J.P. MacIntosh, « Connectivity: The Space, Tempo and Exploitation of Risk in the Information Age », in A.D. Campen and D.H. Dearth (eds.), *Cyberwar 2.0: Myths, Mysteries and Reality*, Fairfax, AFCEA International Press, 1998.

35 Lire à ce propos la synthèse de M.C. Waxman, « Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4) », *Yale Journal of International Law*, vol.36, 2011, pp.421-459.

36 M.N. Schmitt, « Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework », 37 *Columbia Journal of Transnational Law*, vol.37, 1998-1999, 885, p.905.

usage d'une « force de coercition » ou d'une « force d'interférence », car même si l'attribution n'a pas été concluante, le caractère ciblé de l'attaque, jusque dans la programmation et les codes de Stuxnet et Flame, montrent la spécificité de la construction de l'attaque pour obtenir un effet de force. Mais l'interférence obtenue par un agent humain infiltré qui révèle un dispositif d'écoute électronique, ou par les fuites systématiques de WikiLeaks, sont des « forces » bien plus déstabilisatrices à l'échelle d'une nation qu'un sabotage avorté sur une centrale nucléaire. C'est là qu'entre en jeu l'article 51 de la Charte des Nations Unies, et la « permission » de répondre soi-même, au titre de l'auto-défense à une « attaque armée », remettant en cause « l'intégrité territoriale ». Le problème est, bien sûr, qu'il est difficile de se décider à répondre lorsqu'on ne possède pas d'échelle d'évaluation des délits et des peines, et lorsque l'on ne sait pas si « l'État d'en face » ne possède pas lui-même une échelle de réponses très différente[37] ; et comme le souligne Gaycken, cela ouvre la voix à un nombre incalculable de guerres déclenchées par des parties tierces, sans même que les belligérants n'en est réellement conscience[38].

VII. AFFAIRE D'ETAT OU MATIÈRE PRIVÉE ?

Dès lors, en matière de cybernétique, la guerre fera-t-elle loi ? Les tensions que la négociation internationale pour une régulation de la cyber-défense pourraient laisser présager que l'on cherche à préparer une guerre et à lui trouver ses supports juridiques. La convention de Budapest de novembre 2001, qui effectivement devait réguler la cybercriminalité n'a toujours pas été ratifiée par la Chine et la Russie en septembre 2014. S'il n'y a pas de traité signé par les Russes et les Chinois, c'est à cause du problème, assez essentiel, de « définition » de la cyber-sécurité et de la cyber-défense. D'un côté, les Occidentaux séparent la cybercriminalité (lutte contre la criminalité dans le cyberespace), la cyber-défense (actions stratégiques des États dans ce milieu) et la cyber-sécurité (encouragé en cela par les industriels, qui espèrent un cadre de régulation aménagé)[39]. D'un autre, les Russes considèrent de longue date que ces notions n'existent pas individuellement, mais qu'il y a une seule et entière « sécurité de l'information », incluant autant le contenant que le contenu. Toutes les initiatives de régulation ont feint d'ignorer l'obstacle, jusqu'à la dernière

37 J.A. Lewis, « Multilateral Agreements to Constrain Cyberconflict », *Arms Control Today*, June 2010.

38 S. Gaycken, « The Necessity of (Some) Certainty - A Critical Remark Concerning Matthew Sklerov's Concept of 'Active Defense' », *Journal of Military and Strategic Studies*, vol.12(2), 2010, pp.4-6.

39 V. à ce propos l'étude de l'Office des Nations Unies contre la drogue et le crime (ONUDC) : <http://www.unodc.org/documents/organized-crime/UNODC_CCPCJ_EG.4_2013/CYBERCRIME_STUDY_210213.pdf> [Dernière visite, le 01/10/2014].

initiative d'établir un glossaire commun pour tous les membres de ce tour de table mondial[40]. Mais les Russes ou les Chinois ne peuvent renier leurs engagements constitutionnels respectifs, qui sanctuarisent cette conception d'une information, contenant et contenu, perçue comme une entité inséparable.

Rien, pourtant, ne laissait présager que les règles de l'engagement et de la légitimité de la guerre soient durablement transformées par l'avènement d'une société numérique. Comme le souligne Goldsmith, personne dans la première décennie d'existence du « réseau des réseaux » ne s'est jamais soucié de questions de sécurité. Jusqu'en 1988, lorsque Robert Tappan Morris, un étudiant de Cornell, introduisit un ver dans l'Internet naissant dont le but expérimental était de mesurer la taille du réseau, et qui finit par mettre hors d'usage 10 % des 60 000 machines qui composaient alors le net[41]. C'est en 1988 que naît le premier programme officiel de la DARPA sur la sécurité des réseaux, avec comme perspective celle de la résilience à des scénarios d'attaques sur les infrastructures… Mais c'est à Berlin, déjà, que se sont cristallisés les vrais enjeux et questions futures de la cyber-sécurité autour de l'affaire « Hagbard », pseudonyme de Kar Werner Koch, et de la naissance du Chaos Computer Club en 1981 autour de Markus Hess, Hans Heinrich Hübner ou Otto Brezinski[42]. Le détournement de la technologie n'est pas dissociable de la poursuite d'un but politique et, dès lors, pour les Russes autant que pour les Américains qui s'affrontent dans cette première cyberguerre autour de la chute du Mur de Berlin, l'espace numérique naissant est déjà un terrain de conflit géopolitique.

Si la finalité politique d'une campagne de déstabilisation ou de sabotage n'est pas difficile à établir, au moins *a posteriori*, il est très difficile de définir ce qui constitue un « acte de guerre » du point de vue cybernétique. Le *Chaos Computer Club*, dès sa fondation, poursuit un but de démonstration des dangers potentiels d'un mauvais usage des technologies de l'information pour la liberté d'expression, les libertés individuelles et la démocratie, s'invitant au débat politique par la prouesse technique. Ces démonstrations sont nombreuses entre 1981 et 2014, et la plus symbolique est sans doute l'opération visant à démontrer la facilité avec laquelle les machines à voter électroniques néerlandaises pouvaient être interceptées en moins de trois minutes par un petit groupe organisé pour en changer les résultats. Réalisé en 2006 en coopération avec une fondation néerlandaise pour les libertés individuelles, cet exploit du CCC a abouti à une révision de la cour constitutionnelle

40 V. <http://www.unodc.org/unodc/en/commissions/CCPCJ/session/22.html> [Dernière visite, le 01/10/2014].

41 Goldsmith, *supra* note 19, p.129.

42 K. Hafner and J. Markoff, *Cyberpunk: Outlaws and Hackers on the Computer Frontier*, New York, Touchstone, 1995.

néerlandaise de l'encadrement du vote électronique[43], puis à l'abandon de ce dernier[44]. On comprend dès lors qu'on ne puisse aveuglément appliquer une image du droit public, ou du droit privé, dans le domaine « informationnel » sans prendre auparavant quelques précautions. Doit-on juger l'action du CCC sous l'angle du droit commun et, dès lors, l'interdire et la sanctionner dès ses premiers pas ? Ou doit-on la juger par sa finalité de « lanceur d'alerte » et, en ce cas, en saluer la prouesse et la noblesse sur sa finalité ?

VIII. UN CADRE DE RÉGULATION AMBIGU

La première ambiguïté de l'exercice d'un jugement sur un crime informatique réside donc dans sa « non déterminabilité » *ex ante*. L'opération menée par le CCC pour démontrer le caractère manipulable des machines de vote est exactement la même que celle qu'aurait mené un gang criminel qui poursuivrait le but de manipuler les résultats d'un vote national[45]. Le vecteur lui-même n'est pas caractérisable *a priori* comme porteur d'une intention malveillante. Ce qui serait vrai pour un virus, un code malicieux visant à paralyser ou détruire un système logique, ne l'est jamais pour un détournement de fonctionnalité ou de finalité d'usage[46].

Avant même de pouvoir établir l'auteur d'un crime informatique (« attribution »), il est difficile d'établir l'intention réelle d'un test, d'une découverte et d'une exploitation de faille, d'une ingénierie inversée, ou encore d'un détournement de fonctionnalité mal protégé, -ce que l'on peut regrouper sous le vocable *hacks*. Ainsi, les approches qui consistent à vouloir caractériser *a priori* les acteurs du cyberespace entre les *wicked* et les « bien intentionnés »[47]dénotent soit une grande naïveté et méconnaissance du sujet ou, tout au contraire, une position politique affirmée cherchant à promouvoir une forme de « délit de faciès numérique » dans le cadre de régulation.

---

43 Pour une documentation plus détaillée, lire *Die Datenschleuder*, publication du CCC, disponible en ligne : <http://ds.ccc.de/> [Dernière visite, le 01/10/2014].

44 B. Jacobs and W. Pieters, « Electronic Voting in the Netherlands: from early Adoption to early Abolishment », *Foundations of Security Analysis and Design V, Lecture Notes in Computer Science*, vol.5705, 2009, pp.121-144.

45 *Ibid* ; M.C. Libicki, *Cyberdeterrence and Cyberwar*, Santa Monica, RAND, 2009.

46 P.C. Reich et al., « Cyber Warfare: A Review of Theories, Law, Policies, Actual Incidents - and the Dilemma of Anonymity », *European Journal of Law and Technology*, vol.1(2), 2010, pp.1-58.

47 V. notamment : C.C. Demchak, « Complexity, Rogue Outcomes and Weapon Systems », *Public Administration Review*, vol.52(4), 1992, pp.347-355 ; C.C. Demchak, *Wars of Disruption and Resilience: Cybered Conflict, Power, and National Security*, Athens, University of Georgia Press, 2011.

La seconde ambiguïté réside dans la téléologie présumée de l'agression informatique : si son but est le système informatique de l'adversaire, ce système peut-il être assimilé à l'adversaire lui-même ? Faut-il dès lors établir un encadrement légal entre les machines et leurs propriétaires, pour identifier ces « actes de guerre » ou « actes criminels », ou faut-il laisser ce délicat sujet dans le « brouillard de guerre » ? Certains auteurs assimilent l'homme, la machine et l'ennemi sans aucune inhibition : l'un vaut indistinctement l'autre[48]. D'autres préfèrent y voir un distinguo et encourager l'émergence d'un cadre normatif spécifique à l'usage de la force sur les réseaux informatiques[49]. Mais l'agression par des moyens informatiques est-elle vraiment similaire à un acte de guerre ? Peut-on comparer une « cyberattaque » à un usage de la force armée selon l'article 2(4) ou une « attaque armée », selon l'article 51, de la Charte des Nations Unies ? Qu'en est-il des frontières – juridique, économique, sociale – entre la cyberguerre et le crime informatique ?

La question est moins triviale que le laisse entendre sa formulation. Elle pose non seulement la question de la neutralité des technologies de l'information, mais surtout celle de la frontière entre droit privé, droit commercial et droit public dans le cyberespace. Si l'on considère la technologie informatique comme un élément neutre, on doit dès lors accepter l'idée que l'établissement de la preuve, de l'attribution de l'agression, ne peuvent plus être résolus par un simple audit informatique[50].

IX. UN « BROUILLARD DE GUERRE » OPPORTUN

Réguler le cyberespace, c'est s'acheter un passeport d'ingérence dans les affaires internes de ses voisins. C'est du moins comme cela que sont perçus de nombreux articles de ce traité, notamment l'article 32 qui concerne l'accès transfrontalier aux « données » lors de crises, d'incidents, ou d'investigations. Les Russes et les Chinois y sont particulièrement opposés. Mais il n'y a pas que cela. La régulation de la contre-mesure, c'est-à-dire de la légitimité à engager des représailles et à estimer sa

48 V. A.W. Ezekiel, « Hackers, spies, and stolen secrets: protecting law firms from data theft », *Harvard Journal of Law & Technology*, vol.26(2), 2013, pp.649-668.

49 C'est notamment l'axe adopté par la proposition de cadre de régulation du centre d'excellence de l'OTAN de Tallin : M.N. Schmitt (ed.), *The Tallinn Manual on the International Law Applicable to Cyber Warfare: Prepared by the International Group of Experts at the Invitation of the NATO Cooperative Cyber Defence Centre of Excellence*, Cambridge, Cambridge University Press, 2013.

50 S.J. Shackelford, « From Nuclear War to Net War: Analogizing Cyber Attacks in International Law », *Berkley Journal of International Law*, vol.25(3), 2009, pp.192-250.

proportionnalité, est évidemment au cœur des différends[51]. L'Ouest pousse une vision de l'attribution des attaques proche de celle du droit commun, qui associerait donc la responsabilité des crimes à leur origine géographique, et le « droit de réponse » à une forme d'extra-territorialité accordée par ces traités. Mais, depuis leur rédaction, les technologies ont changé et l'attribution géographique des campagnes d'attaques avancées modernes est quasiment impossible[52]. On finirait par avoir un cadre de régulation efficace contre le petit crime commun, suffisamment naïf ou mal équipé pour ne pas masquer son origine d'attaque, et incapable d'encadrer les attaques d'États, ou celles du crime organisé, capables de soustraire technologiquement à la régulation.

Certains partenaires de ces négociations y voient une « carte blanche » donnée aux nations qui ont un avantage technologique pour mener une « guerre limitée » ou une « guerre sale » électronique permanente[53]. Le problème est qu'en matière de crime numérique, détenir l'arme, le mobile et le lieu du crime est loin d'être suffisant pour établir une intentionnalité et l'identité du coupable ! Les cyberguerres répondent à un paradigme dominant qui est celui de la guerre limitée ou du conflit contre-insurrectionnel : les moyens y sont empruntés, détournés, déplacés et subversifs par nature[54]. L'usage intensif de capacités robotiques (*botnets*, etc.) rend complexe la tâche d'établir une intentionnalité ou une source. Le modèle proposé par le cadre de régulation actuel est celui des conflits conventionnels, incluant, par exemple, des notions comme l'escalade du conflit. Mais qui aujourd'hui seraient les premières victimes d'une « escalade », par exemple sur les infrastructures vitales, d'un cyber-conflit ? Probablement ceux qui dépendent, d'un point de vue économique, des infrastructures numériques. Les Chinois ou les Russes ont fait des choix stratégiques différents en matière de développement d'infrastructure nationale d'information. Ils ont des réseaux suffisamment compartimentés, en partie avec des technologies propriétaires, développées depuis le milieu des années 1980 ; certes, sans doute plus lentes, mais beaucoup moins sensibles à une escalade de dommages en cas de conflit.

Le problème de fond de la régulation reste la machine, c'est-à-dire le point d'ancrage de l'économie numérique. Pourquoi les Américains sont-ils si pressés sur le sujet ?

---

51 V. J.A. Lewis, « Aux armes, citoyens: Cyber security and regulation in the United States », *Telecommunications Policy*, vol.29(11), 2005, pp.821-830.

52 M. Van Eeten and J.M. Bauer, « Emerging Threats to Internet Security: Incentives, Externalities and Policy Implications », *Journal of Contingencies and Crisis Management*, vol.17(4), 2009, pp.221–232.

53 S. Li, « When Does Internet Denial Trigger the Right of Armed Self-Defense? », *Yale Journal of International Law*, vol.38(1), 2013, pp.179-216.

54 P.D. Allen and C.C. Demchak, « The Palestinian-Israeli Cyberwar », *Military Review*, vol.83(2), 2003, pp.52-59.

Ils mettent en avant l'enjeu du vol de propriété intellectuelle avec une perte de 4,8/5 milliards de dollars selon le rapport du Congrès[55], ce qui est une estimation faible, avec des rapports alarmistes sur les infrastructures critiques par ailleurs. Cet empressement cache une vulnérabilité majeure, provenant de l'industrie elle-même. Aujourd'hui, le point de création de valeur, le point d'ancrage des modèles économiques, et le point de confrontation des cyberconflits est le même : votre machine, ma machine. Avec un commerce électronique bientôt dominant les transactions commerciales, une formation des opinions directement dépendante d'interfaces hommes-machines (Google, Facebook ou MSFT parmi d'autres), le levier de création de richesse nationale se trouve être, aussi, le potentiel champ de bataille numérique[56]. C'est ce qui explique sans doute la divergence des agendas entre les majors qui veulent un modèle où chaque citoyen est le simple titulaire d'une licence d'exploitation de ses propres données personnelles et le modèle de souveraineté sur la propriété intellectuelle, qui suit une logique nationale. La transparence et l'accès que réclament les États ne sont pas forcément dans l'intérêt des industriels du numérique, et *vice-versa*. Cela a conduit certains exploitants à refuser de continuer ce genre d'échanges, non par préoccupation pour la défense des libertés personnelles, qu'ils continuent de bafouer allégrement, mais pour éviter que la régulation internationale de la cyberdéfense vienne perturber leur modèle économique.

X. ADAPTER LE DROIT À L'AUTOMATION COMPORTEMENTALE

La question de la contre-mesure est avant tout la question de la part que l'on souhaite laisser à l'automate et aux algorithmes de raisonnement autonomes dans la société de demain, ainsi que de la part que l'on souhaite réserver au privilège d'être humain. L'état de l'art technique est incapable aujourd'hui de rattraper la courbe de production et d'inventivité des technologies capables de perturber, déstabiliser, espionner, détourner ou détruire des infrastructures d'information globales, et historiquement non conçues pour gérer de tels problèmes de sécurité. Il semble que les régulateurs courent après des ennemis imaginaires et qu'ils aient eux-mêmes créé l'aporie dans laquelle se situe, depuis 2001, la négociation pour établir un cadre réglementaire et légal au cyberespace.

55 Executive Office of the President of the United States, Administration Strategy on Mitigating the Theft of U.S. Trade Secrets, February 2013, disponible en ligne : <http://www.whitehouse.gov//sites/default/files/omb/IPEC/admin_strategy_on_mitigating_the_theft_of_u.s._trade_secrets.pdf> [Dernière visite, le 01/10/2014].

56 P. Baumard, « Numérisation, liberté et sécurité », *Politique Internationale*, n°135, 2012, pp.371-377.

La production des contre-mesures, que sont aujourd'hui les très inefficaces « découvertes de failles », passe par une organisation de marché déficiente : celui des « exploits » (espaces de marché numérique fermé où sont achetées les découvertes de vulnérabilités « zéro jours ») ; mais c'est, trop souvent, un marché de dupes qui a connu une inflation sans précédent dans les cinq dernières années, sans améliorer l'efficacité de la défense[57]. Que ce soit du point de vue de son équation économique, ou de celui des progrès technologiques de la ligne de défense, la stratégie qui consiste à vouloir appliquer les articles 2(4) et 51 de la *Charte des Nations Unis* à la gestion du cyberespace risque bien de provoquer l'inverse du résultat escompté. L'enjeu des « contre-mesures » pose ainsi la question de savoir si c'est le plus compétent, le souverain ou le gestionnaire des infrastructures, qui doit prendre les armes et répondre à l'attaquant. La doctrine de la « compétence » est celle des technologistes. Elle pousse bien sûr à l'extraterritorialité des fournisseurs, et derrière eux, du ou des pays qui en sont les hébergeurs ou actionnaires. Ce scénario est souvent identifié comme celui des États-Unis, où la suprématie du secteur américain de la cybersécurité et de l'appareil d'État pour la cyber-défense permettrait d'établir un modèle qui consiste à confier au fournisseur de la technologie de cybersécurité la responsabilité des contre-mesures. Cette doctrine s'oppose au modèle « souverainiste » vertical des Russes qui ne sépare pas l'information en contenants et en contenus. Les Chinois, sur ce point, suivent les Russes, et même les dépassent en en faisant un objectif affiché de sécurité économique nationale. Les Européens tentent de faire front commun sur les interceptions, sur la protection des données personnelles avec le projet étudié actuellement par la Commission européenne.

Les positions à l'intérieur de l'Europe sont bien sûr très variées. Les Allemands ont une perception différente de la nôtre de la protection des données personnelles, dont les souvenirs douloureux sont encore récents. Cela se traduit effectivement dans leurs événements domestiques (flou des images sur Google View, refus du recueil de données personnelles, émergence du Parti pirate promoteur de ces thèmes). Les Français, on le sait, ont une relation différente à la donnée privée, souvent perçue comme paradoxale au pays des droits de l'homme. Il y a sans doute là une carte diplomatique que nous n'avons pas su jouer, tant nous sommes absents, du point de vue stratégique, des corps doctrinaires dominants. La régulation de la cyberdéfense et de la cybersécurité est un enjeu qui ne mobilise pas l'opinion et,

57 S. Ransbotham, S. Mitra and J. Ramsey, « Are Markets for Vulnerabilities Effective? », *MIS Quarterly*, vol.36(1), 2012, pp.43-64 ; R. Anderson et al., « Measuring the cost of cybercrime », in *Proc. of WEIS'12*, 2012. Les découvertes de vulnérabilités exploitables (« exploits ») peuvent s'échanger sur des places de marché anonymes entre 20 et 300 K$ (marchés criminalisés), tandis que la déclaration légitime de découvertes de failles envers les fournisseurs est généralement récompensée entre 2 et 5 K$ en 2013.

pourtant, il s'agit de la régulation la plus transformatrice des cinquante prochaines années. Réguler le recours à la sanction, à la protection et à l'interdiction dans le monde numérique, c'est définir les frontières des sociétés civiles demain et, très certainement, les positions de puissances économique et politique de ces sociétés.

CONCLUSION

Comprendre la psychologie sociale du développement des menaces devient critique alors que nous entrons dans une période pionnière étrangement similaire aux années de *phreaking* (1972-1987). L'amélioration de la portabilité de l'apprentissage machine (embarquée, distribuée et autonome) est curieusement absente de la majorité des hypothèses de travail des doctrines étudiées. Cet état des lieux est peut-être dû à la transposition de modèles d'escalade (course à l'armement, concentration, capacités décisives) emprunté aux doctrines militaires pour aborder les enjeux de cybercriminalité. Les capacités offensives cybernétiques ne répondent pas aux modèles traditionnels d'escalade ou de renforcement. Elles fondent la malveillance de leur capacité offensive sur leur nature transformationnelle, leur déploiement distribué et la supériorité de leur apprentissage autonome.

L'autonomie d'apprentissage des sytèmes d'attaques soulève aussi bien la question de l'attribution judiciaire que celle de l'organisation du progrès technique de la filière de cyber-sécurité. D'une part, les vecteurs d'attaque, c'est-à-dire les programmes ou codes malveillants, ne portent pas toujours une signature ou une identification pouvant être rapprochées d'un auteur spécifique. De nombreux composants d'attaques avancées sont assemblés à partir de librairies d'application en *open source,* appartenant au domaine libre sur lequel repose très majoritairement le progrès technique, aussi bien pour les défenseurs que les attaquants. Quand bien même, une partie du code utilisé pour une attaque serait "signé" ou imputable à un auteur identifié, cela ne veut pas dire qu'il ait été créé dans un but malveillant. La recherche en cybersécurité repose sur la possibilité de mener des études scientifiques sur les attaques et les contre-mesures, et par nécessité, d'en explorer le caractère offensif. D'autre part, la prédiction comportementale et l'apprentissage non-supervisé ont connu depuis dix ans des avancées spectaculaires qui laissent présager une rupture radicale dans ce que l'on pourrait appeler des attaques "métamorphiques", c'est-à-dire capables de se transformer de façon autonome, sans supervision ou sans chemin préalablement programmé. Les systèmes de défense seront de plus en plus dépendant d'apprentissages de causalité complexe, également autonomes et non-supervisés, qui de leur côté aussi, adopteront des conceptions métamorphiques.

Dès lors, les débats virulents qui ont opposé Russes, Chinois, Américains et Européens depuis 2001 paraîtront certainement triviaux. Doit-on attribuer la responsabilité d'une escalade entre deux automates d'attaque et de défense à leurs concepteurs respectifs ? Qu'advient-il lorsque l'intégralité des codes et des librairies logicielles utilisées par les deux côtés proviennent du domaine libre et n'autorisent aucune identification ? Doit-on en sanctionner et en réglementer *l'usage* au risque de pénaliser une recherche fondamentale essentiel au progrès technique de la cyberdéfense ? Le probème est d'autant plus complexe qu'une technologie de cybercriminalité comportementale, reposant sur l'intelligence artificielle, peut *ne pas présenter* de composants pouvant être qualifés, pris un à un, comme "sensibles" pour la sécurité. C'est bien là le changement futur de paradigme qui pourra rendre inopérant du point de vue technique des méthodes fondées sur l'identification de codes malicieux, et impraticable du point du vue juridique, l'établissement d'une préméditation fondée sur la recherche d'un "code malveillant" dans le cadre d'une enquête judiciaire.

Les algorithmes de prédiction comportementale font déjà partie de nos sociétés. Ils préviennent les risques d'emballement sur les marchés de change ou de matière première. Ils évitent les collisions d'avions de ligne. Ils contribuent à la maintenance des infrastructures à risques critiques. Leur extension à la croisée des sciences du vivant, de la robotique et de l'intelligence artificielle est inéluctable. Tout autant que le sera leur détournement et leur usage malveillant. L'histoire du "hacking" et de son frère ennemi, la "cybercriminalité" nous montre que l'exploration libre, la créativité et la défense des libertés fondamentales d'exploration et d'expression ont été essentielles à créer une société qui peut globalement résister à des menaces sur sa résilience. La courte histoire de la "cyber-régulation" nous a montré un tout autre visage: celui d'un "brouillard de guerre" prémédité et tactique, maintenant la réflexion technique sur un corpus technologique inadéquat, et substituant le conflit idéologique à la préparation raisonnée des enjeux futurs de la société.