Deterrence theory in the cyber-century.

Lessons from a state-of-the-art literature review

Annegret Bendiek[1] und Tobias Metzger[2]

Abstract: Cyberattacks from a variety of perpetrators are constantly rising. To achieve restraint from attacks, deterrence theory has long been considered a valuable concept. How do criticisms of classical deterrence apply in this relatively new domain, and how does cyberdeterrence differ? Can offensive cyber capabilities be effective in deterring adversaries or must kinetic retaliation be "on the table"? Freedman differentiates deterrence-by-retaliation and deterrence-by-denial; immediate versus general; narrow versus broad; and central versus extended deterrence. The authors argue for denial and retaliation as complimentary parts of an overall strategy consisting of resistance, resilience and responses. Unlike nuclear deterrence, cyberdeterrence is not a game of great powers or that of nation-states alone and grey zones serve only rogue actors in the medium-term. Only the establishment of clear rules, similar to the Budapest Cybercrime Convention, enables effective responses to prevailing threats. Agreement on international rules on civilian critical infrastructures would move the superficial discussion on whether or not cyberattacks are legitimate to the more relevant debate on which targets are acceptable, providing clarity for the development of effective military strategy.

Keywords: Cyberdeterrence; Cyberwarfare; Cyberdefence; Deterrence theory; Cybersecurity; Freedman.

1 Introduction

Cyberwar is regularly invoked in journalistic, academic and even political discourse. Yet, apart from the U.S.-Israeli Stuxnet attack on Iran's nuclear facilities in 2010, no cyberattack has ever caused large-scale physical damage. UK Labour last year urged their government to consider the "growing threat of cyberwarfare"[3] and former Defence Secretary Leon Panetta repeatedly warned of the lurking threat of a "cyber Pearl Harbour".[4] While the Stuxnet attack remains the only instance of severe physical damage inflicted via cyberspace, the Internet's increasing pervasiveness and national development of military units bear the risk of militarization. Connecting growing parts of the industry, energy production and society to the Internet fosters economic growth, increases efficiency and, benefitting from the Internet's anonymity, furthers human rights. Nevertheless, U.S. think-tanker Jason Healey cautions that, as the Internet of

[1] German Institute of International and Security Affairs (SWP), Senior Associate in the Research Division EU/Europe, Ludwigkirchplatz 3-4, 10719 Berlin, annegret.bendiek@swp-berlin.org
[2] German Institute of International and Security Affairs (SWP), Research Assistant in the Research Division EU/Europe, Ludwigkirchplatz 3-4, 10719 Berlin, tobias.metzger@swp-berlin.org
[3] [MA14]
[4] [Se12]

Things (IoT) spreads, "a cyberattack will destroy not only ones and zeros, but things made of steel and concrete. And when they break, people will die."[5]

To achieve restraint from attacks, deterrence theory has long been considered a valuable concept. While deterrence will remain an instrument in security policy – any consideration of deterrence theory must acknowledge that its limitation to the military, and more specifically the nuclear, domain is insufficient. In line with this, various authors have discussed the extent to which deterrence theory is applicable to cyberspace. Their findings of appropriateness and limitations are a necessary starting point for policy-making recommendations. How do the criticisms of classical deterrence apply in this relatively new domain, and how does cyberdeterrence differ from its kinetic counterpart? Can offensive cyber capabilities be effective in deterring adversaries? Must kinetic retaliation be "on the table" for deterrence to succeed? Which changes are required for its implementation and where do key challenges lie? Unlike Gaycken and Martinelli, our definition of cyberdeterrence is built on both deterrence of cyberattacks and deterrence by threatening cyberattacks, arguing that rather than being separate, they are different escalatory steps and conceptually cannot be separated.[6] The means of deterrence are part of an overall toolbox and discussing cyber separately would be similar to speaking, for instance, solely about deterrence effects of the navy or air force. The authors build upon Lawrence Freedman's types of deterrence: deterrence-by-retaliation and deterrence–by-denial; "narrow" vs. "broad"; "central" vs. "extended" and "immediate" vs. "general" deterrence. Addressing these questions is essential for determining a deterrence strategy's effectiveness. In its study for the U.S. Air Force, RAND subsumes only questions of punishment as deterrence, while referring to all deterrence-by-denial mechanisms as "defence".[7] It suggests that deterrence-by-denial and deterrence-by-retaliation ought to be considered separately. We argue that they are complimentary. For deterrence to be effective in cyberspace, actors dealing with foreign and security policy can find useful reference points in the state-of-the-art literature on deterrence theory and cyberdeterrence.

2 In theory – Deterrence theory and cyberspace

Throughout the Cold War, deterrence theory was the preferred framework of analysis and of military doctrine to explain the influence of nuclear weapons, and to argue that nuclear powers, fearing the consequences, would not go to war with each other. Some authors have since applied the theoretical framework to cyberspace, as cyberdeterrence[8].

[5] Jason Healey in [SD14a], 31
[6] Cf. [GM13]: Differentiation made between „cybered deterrence", as deterrence by cyber means, and "cyber deterrence", as deterrence of cyberattacks.
[7] Cf. [Li09], 7,8
[8] For in-depth accounts, see [Kn10]; [Ci12]; and [GM13]

While both domains share characteristics, such as the offensive advantage, given the difficulty and costliness of defence[9], significant differences exist.

The emergence of deterrence in military theory dates back to the 1920s/30s when the first flight bombers were considered unstoppable by defensive measures. Then, strategists thought that large-scale attacks on one's cities could only be prevented, if the other side feared counter-attacks of similar or greater magnitude. The first nuclear bombs demonstrated a similar offensive advantage, and Bernard Brodie was among the first to observe that "from now on [the military establishment's] chief purpose must be to avert [wars]".[10] Deterrence theory gained prominence and developed to its present state during the Cold War nuclear stand-off between the USA and the Soviet Union. The term goes back to the Latin "dēterrere", meaning to "frighten from or away", and is defined as "to discourage and turn aside or restrain by fear".[11] "Deterrence is concerned with discouraging others from acting in ways that advantage them but harm you". This definition highlights the two notions of deterrence, firstly the would-be-attacker's turning back due to the other's defences, and secondly restraint for fear of retaliation. The threat of force by *'A'* and the voluntary restraint of *'B'* are central elements. As Freedman puts it: "strategies geared to coercing others to act in ways they might consider harmful but advantage you have been described as compellence or coercive diplomacy".[12] The main difference is that deterrence seeks to preserve the "status quo" and is limited to persuading someone not to do something, while compellence seeks to enforce a change, typically within a frame of urgency.

Deterrence requires two components: the expressed intention of *A* to defend an interest; as well as the ability to achieve the defence or the (perceived) certainty by *B* that interference with *A's* interest will be costly for the attacker, i.e. credibility. However, signalling – used to communicate both the interest to be defended and the threats to be implemented in case of non-compliance – is never straight forward. Freedman stresses the possibility of *A* badly articulating or *B* misunderstanding the threat, thus rendering deterrence ineffective. Signalling or "brandishing"[13] of weapon capability can be very costly, as evident in Israel's military operations aimed at deterring future attacks. The ultimate aim is strategic deterrence, thus creating "internalised deterrence", no longer requiring explicit signalling.[14] Robert Jervis describes an evolution of "three waves of deterrence": Firstly, Brodie's concept of deterrence to avert wars when only the West possessed deployable nuclear weapons; secondly, the rise of the Soviet Union as a nuclear power leading to a bipolar world evoking the question "if nuclear war couldn't be fought how could it be threatened?". At this stage second-strike capabilities deployable from submarines assured Mutually Assured Destruction (MAD), making it

[9] See also [Li14], 5, 96: Calculations put offense at 132 times cheaper than defence, and estimate a $100m "cyber army" as being able to overcome any U.S. cyberdefence, and cf. [Ma12]
[10] [Br46], 31 and cf. [Fr04], 9-11
[11] Search for "to deter", [Ox14]
[12] Both quotes [Fr04], 109
[13] Cf. [Li13]
[14] Cf. [Fr04], 28-32

impossible to inflict sufficient damage to disarm an enemy, which would prevent retaliatory attacks. Game theory was used to evaluate whether cost-benefit evaluations could still favour deterrence postures, issuing a "threat that leaves something to chance"[15] or threatening sanctions using limited (non-nuclear) strikes. U.S. State Secretary Dulles argued for deterrence of would-be-aggressors from a cost-efficiency perspective, calling it "the way to getting maximum protection at bearable cost".[16]

Jeremy Bentham, one of the inventors of deterrence theory in criminology assumes rational individuals capable of performing cost-benefit calculations prior to taking action. The third wave raised considerable doubts about this rational actor model, an important pillar of deterrence theory, arguing that groupthink, misperceptions and bureaucratic politics often overruled mere cost-benefit calculations. Rationality is subjective, and the challenge of deterrence signalling lies in identifying the opposing side's rationale.[17] Furthermore, cost-benefit calculations require clarity and predictability of sentencing and proportionality between punishment and violation. Bentham observed that where arrest is unlikely, severe punishment can help keep up a criminal's expected cost of getting caught, but once punishment is perceived as disproportionate, it loses its effect.[18] In line with the controversy of "rational" actions, signalling can be misunderstood given differences in culture or education. Furthermore, there are claims of fallacies in traditional deterrence theory's sole consideration of state actors. Third-wave theorists Alexander George and Richard Smoke disapproved of the exaggerated role of the military vis-à-vis other foreign policy tools, especially positive inducements.[19] The rise of non-state actors, "rogue states" and "terrorists", brought a possible fourth wave. These actors may be beyond containment by deterrence since they have no nation or citizens to defend. Finally, theorists introduced the differentiation between interest-based and norms-based deterrence in the 1990s to overcome previous challenges. The former aims at deterring challengers of "hard" national interests, while the latter advocates less clear-cut norms. Freedman argues for norms-based deterrence to reinforce "certain values to the point where it is well understood that they must not be violated".[20]

Deterrence-by-retaliation and deterrence-by-denial

To dissuade would-be-offenders from attacking, deterrence theory typically distinguishes two means: denial and retaliation. In the original criminological context, Bentham describes deterrence-by-retaliation, or deterrence-by-punishment, as imposing the "significant likelihood of any culprit being apprehended, brought to trial, found guilty and then receiving a sentence … that will make an impression not only on their future behaviour but on the behaviour of others".[21] This was the core idea of nuclear deterrence until the 1970s when U.S. President Reagan's "Strategic Defense Initiative"

[15] [Sc66] and cf. [Po08]
[16] [Du89] as quoted in [Fr04], 9
[17] Cf. [Mo77]. Bentham's collected works are available online, cf. [Be43]
[18] Cf. [Li09], 29
[19] Cf. [GS74]
[20] [Fr04], 4
[21] [Fr04], 60, 61

(SDI) introduced the notion that is was better "to protect than avenge".[22] The SDI suggested a deterrence-by-denial approach. Deterrence-by-resistance and deterrence-by-resilience are two different approaches within this concept. Resilience is the ability to quickly restore the original shape after an attack, Quick recovery limits potential gains and can convince an opponent not to attack, if the cost of attacking becomes excessive. Both – resistance and resilience – are aimed at denying a would-be-offender's gains, either by building unsurmountable defence structures or by ensuring quick recovery following an attack. Freedman extends the distinction beyond, firstly, the denial versus retaliation debate: He, secondly, introduces "narrow" versus "broad" deterrence, distinguishing whether a particular military operation (e.g. the use of nuclear missiles) or any type of attack is to be deterred; thirdly, "central" versus "extended", thus including the protection of third-party allies in one's deterrence demeanour; fourthly, "immediate" versus "general" deterrence, distinguishing between a crisis situation between known actors and non-state of emergency deterrence against unknown would-be-aggressors.[23]

Many of the same elements apply regarding attacks through cyber means although with some limitations. Signalling, for one, faces comparable challenges, although further complicated by the multitude of actors including non-state groups and individuals. Other challenges are entirely new: In deterrence-by-retaliation, credibility is difficult to establish since demonstrating cyberpower and retaliating immediately and repeatedly is problematic, as outlined below. Although the U.S. and NATO have emphasized their willingness to respond to cyberattacks at a time, place and by means (including kinetic) of their choosing, there can hardly be automaticity in response. The similarity of nation-state and criminal cyberattacks requires time-consuming and costly forensics and close coordination between law enforcement and the military. With their reluctance of admitting to their development of offensive capabilities, and of strategically discussing their legality, Germany and other nations create opaqueness and significant grey areas. Given the offensive advantage, the number of attackers using cheap, readily available tools will continuously rise, empowering non-established powers such as Iran, North Korea or even Daesh/IS. Reversing this trend requires getting serious about agreeing on international norms and improving both defences, especially employees' and citizens' "cyber-hygiene", and about enforcement. Lack of clarity and impunity for attackers is a major roadblock for effective deterrence strategies.

3 In practice – Suitability of cyber: lessons and implications

The U.S. Army analysis that "[f]undamentally, there is no difference between deterrence in the cyber domain than in any other domain"[24] is flawed in at least three regards: First, cross-border differences in law enforcement and legal practice as well as unwillingness to cooperate allow attackers to act with impunity, diminishing the deterrent's credibility.

[22] [Fr04], 19
[23] Cf. [Fr04], 32-42
[24] [Ph13], 16

Second, since "cyberweapons" rely largely on previously unknown, so called zero-day, vulnerabilities and cannot be displayed prior to their use, it is difficult to demonstrate power. Third, deterrence-by-denial differs greatly, since in cyber "you have to work from the assumption that your networks are already compromised"[25], meaning deterrence is constantly failing. In the nuclear context, all intrusions must be deterred, and a single instance of failed deterrence could mean the use of nuclear warheads and large-scale loss of life. Fourth, different from nuclear confrontation, uncertainty arises from the multitude of actors which threaten harm to one's systems and from the difficulty of quickly attributing an attack. Fearing unforeseen escalation, this may hinder immediate retaliation, especially when retaliating with kinetic strikes.

Ambiguously defined interests, misperceived signalling and uncertainty on how to demonstrate force and how to respond, hamper deterrence postures. Melissa Hathaway, former director of a classified high-level effort to establish a U.S. cyberdeterrence strategy, admitted that "we didn't even come close".[26] While being a useful frame of analysis, cyberdeterrence fails to satisfy any of Patrick Morgan's six elements of classical deterrence theory, according to Stevens.[27] Libicki aptly summarizes the core issues to be observed in national cybersecurity efforts: "The ambiguities of cyberdeterrence contrast starkly with the clarities of nuclear deterrence. In the Cold War nuclear realm, attribution of attack was not a problem; the prospect of battle damage was clear; the 1,000th bomb could be as powerful as the first; counterforce was possible; there were no third parties to worry about; private firms were not expected to defend themselves; any hostile nuclear use crossed an acknowledged threshold; no higher levels of war existed; and both sides always had a lot to lose."[28]

3.1 Key challenges: Credibility and capability to display and use force

Beyond technical issues – which are being addressed – deterrence-by-retaliation is a question of credibility and capability. Firstly, could governments effectively deploy cyberforce to respond to attacks, and should they use kinetic force – and if so under which circumstances – to react to cyberattacks, potentially risking escalation? Secondly, should there be stronger declarations of cyberforce to signal one's ability? For deterrence-by-retaliation to work, the capacity to display force is crucial. Only verifiable tests and the demonstration of nuclear weaponry's destructive force convinced nations that future use ought to be feared. Similar demonstration of cyberpower[29] is unlikely because of the "one use only component", meaning that any use reveals significant information necessary to defend against future attacks. Deterrence-by-retaliation requires

[25] Michael Daniel, Cybersecurity Coordinator of U.S. President Obama, at [SD14b]
[26] Hathaway, Melissa quoted in [Ma10]. On cyberdeterrence in U.S. strategy, see [St12], 154
[27] Cf. [Mo03], 8 and cf. [St12], 152: 1. No prevailing military conflict; 2. rational choice models differ for non-state actors; 3. the challenge of attribution complicates immediate retaliation; 4. repeated retaliation and certainty of inflicting severe pain is hampered; 5. the difficulty of demonstrating offensive capabilities lessens credibility; and 6. a multitude of (non-state) actors constantly threatens stability, risking escalation.
[28] [Li09], xvi
[29] For cyberpower's relevance for policy-making, see [Sh11]

a demonstration of force, but most cyberweapons are rendered useless if vulnerabilities are closed.[30] Signalling is complicated by the multitude of actors beyond the great powers, and including the private sector and individuals. Willingness to retaliate requires unmistakable signalling on the interest to be defended. While offensive capabilities have been revealed, Germany has never publicly proclaimed their existence or threatened adversaries with their use in response to attacks. France and the UK publicly announced offensive capabilities, and France stated its intent of becoming a cyberpower in its 2011 cyber doctrine, but both left it unknown when and how cyberweapons could be used.[31] The U.S., on the other hand, emphasized that their response to a cyberattack would not necessarily be via cyberspace, but that all options would be on the table.[32]

Mutually assured destruction, equally, is out of the question, changing the cost-benefit calculation in favour of attack. If the MAD principle does not apply, the consequences of retaliation are considerably less severe, reducing the inherent costs of an attack. This leads to an advantage for the party striking first, according to offence-defence theory, making attacks more likely. While this suggests for other characteristics of offence-defence theory to apply – e.g. that an arms race ensues – the lacking ability to make quick and decisive victory changes the equation. If cyberattacks cannot disarm an opponent, what is the point of rushing into retaliation? An attack may be stopped, but the attacker cannot be disarmed since cyberattacks can be conducted from third-party hardware, including internet cafés, unsecured wireless networks or infected computers, as part of a botnet. Furthermore, the asymmetric nature of cyberspace is a common argument against retaliatory attacks against criminals or "cyber-terrorists". Deterrence fails if there is no valid target to strike back at. The less connected an adversary, the less vulnerable he is to retaliation with cyber means. Retaliation by kinetic means, on the other hand, bears the risk of incurring injuries or deaths where the initial attack did not. Retaliation "requires not only breaking into sufficiently privileged levels, but also figuring out how to induce a system to fail and keep on failing".[33] The ability to induce superficial damage, which is quickly repaired, has no deterrent effect. However, predicting damage ahead of an attack is difficult, as this depends on the other's technical and procedural resilience but also on chance – e.g. whether patches are installed, closing previously existing backdoors. The half-life of exploits creates a "use-it-or-lose-it dilemma".[34] While initial attacks allow for intensive intelligence work, repeated retaliation is costly, if not impossible, since it requires firstly rapid attribution and secondly immediate and continuous knowledge of the target system. Figure two puts this question in context.

[30] Cf. [Li13], xv-xvii and vii-viii
[31] Cf. [Bl13], and [Li13], 25
[32] Cf. [De15]
[33] [Li13], viii
[34] [Li09], 58

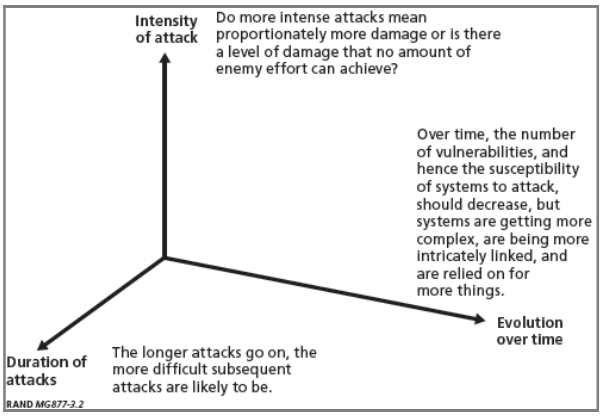

Abb. 1: Limits to retaliation in cyberspace[35]

Additionally, once malware source code such as that of Stuxnet became available online, it revealed flaws in Siemens' operating system, thus enabling their use against other CI. Attackers "may risk handing over ammunition to the enemy as a blueprint for the latter to develop a cyber weapon of its own", as a result of reverse engineering.[36] Instead of taking advantage of such flaws and even buying such information off the black-market, governments must inform the manufacturer to rid systems of backdoors. NSA exploitation of zero-day vulnerabilities and their purchasing with a 2015 budget of $25 million dollars[37] strengthens this business, and may open "Pandora's Box" encouraging others to follow suit.

3.2 How to deter? Deterrence-by-denial and deterrence-by-retaliation

In a cyberspace of easy targets, weak cross-border cooperation and diverging legislation the odds are tilted in favour of attackers. As RAND puts it, "if cyberattacks can be conducted with impunity, the attacker has little reason to stop."[38] Former White House advisor Richard Clarke admits that currently "we cannot deter other nations with our cyber weapons. Nor are we likely to be deterred". According to him, deterrence requires getting "serious about deploying effective cyber defences for some key networks" and "since we have not even started to do that, deterrence theory … plays no significant role in stopping cyber war today". He demands a defensive triad of improving backbone network security, protecting critical infrastructures (CI) and improving the security of military networks and weapons.[39]

Determining the type of defence

[35] [Li09], 60
[36] [Sh14], 78
[37] Cf. [Gr14]
[38] [Li09], xvi
[39] Cf. [CK12]

For deterrence-by-denial to succeed, cost-benefit calculations of attackers must turn negative: "A strong defense deters an attack by convincing an attacker there will be no gains commensurate with the cost of attack".[40] Good defences make an attack less likely to succeed, add credibility to retaliatory measures, reduce the success rate of low-sophistication third-party attacks and make it easier to attribute attacks. As an important differentiation, we note that deterrence-by-denial exists in the pre-event as defence, and in the post-event as resilience. Bologna et al. urge to move from a "fortress" to a "resilience" approach[41], while a RAND study suggests "that, in this medium, the best defense is not necessarily a good offense; it is usually a good defense".[42] The importance of defence is evidenced in NATO's emphasis on cyber-hygiene. If a system is ridded of low-level disturbances, advanced threats are easier to identify and to counter.[43] "Honeypots", entirely sterile systems which capture malware and allow analysts to observe their behaviour in a vacuum, leverage this observation. Another feature of cybersecurity strategies – especially that of the EU – is "strategic dependency management".[44] This includes all measures to secure key components of the supply chain by deciding which levels of strategic independence are required in industry R&D, manufacturing and maintenance of important IT components. It serves to minimize national risks, but often faces accusations of government protectionism. After the Snowden revelations risk management strategies led China to reduce their reliance on U.S. technology, e.g. deciding against Microsoft's Windows 8.[45] It also led to calls within Europe for greater technological sovereignty from U.S. communication infrastructure. Such measures, while sensible in light of deterrence-by-denial, reverse the purpose of the global Internet and defensive measures, e.g. Iran's creation of parallel structures for a national network to prevent high-level cyberattacks, can be bypassed as evidenced in "Stuxnet", which infected the nuclear reactor's software via a USB-stick. While Germany uses offensive means to stop ongoing attacks[46] and works to strengthen cross-border criminal law enforcement, most actions can be seen from a deterrence-by-denial angle. The strong defence network of public and private Computer Emergency Response Teams (CERT) is one component thereof. However, distinguishing between offence and defence is not always easy since evaluating one's own defences requires "penetration-testing" and thus the ability to intrude into systems.

Adding offence to the equation

Other experts strongly disagree with the emphasis on defence, arguing that "in an offense-dominant environment, a fortress mentality will not work".[47] In the nuclear context, deterrence-by-denial approaches were largely inapt given the unbearable costs

[40] [Ph13], 2-4
[41] Cf. [Bo13]
[42] [Li09], 176
[43] Interview [NA13]
[44] Interview [ED14]
[45] Cf. [Re14]
[46] Interview [BM14]
[47] [Ly10]

of defensive failure. In cyberspace, lack of clarity and credibility of punishment encourages cyberattackers to test defences and push their limits, defying deterrence-by-denial. The low cost, as the sum of conducting the attack plus the cost of likely penalties, furthers the offensive advantage, as does the existence of a vast black market, offering anything from zero-day exploits to off-the-shelf services to conduct denial-of-service attacks.[48] The nature of software development renders entirely avoiding bugs, or loopholes impossible, and quickly responding to attacks and closing vulnerabilities entails significant costs for the defender.

3.3 When and whom to deter? Immediate vs. general deterrence and the challenge of attribution

Immediate vs. general deterrence is the question of whether deterrence is aimed at a specific adversary during an ongoing conflict or more generally at any would-be-offender during peacetime. Due to its development time and the need to know a system's vulnerabilities, cyberforce is little appropriate as a concrete deterrence measure once a conflict has erupted. On the other hand, retaliation against cyberattacks cannot be adhoc because of the need for time-consuming forensics. Thus, both general deterrence strategies aiming to deter any attack by cyberforce as immediate deterrence by cyber means alone are improbable. Rather cyber can be an instrument in the broader toolbox. "Active defence" is ill-fitted for various technical, political and legal reasons. Such fully-automated retaliatory attacks may damage computers that are part of botnets without their owner's knowledge, while the botnet as a whole will simply replace the neutralised computer. Furthermore, nation states are not likely to legalise such mechanisms, thereby allowing private sector "corporate vigilantisms" to violate the government's monopoly on the use of force. The multitude of actors entails further difficulties: During conflicts, patriotic and other hackers can engage in disruptive activities to add an unpredictable step of escalation. Governments may not be able to control these groups, weakening the implied promise of deterrence that "if you stop, we stop".[49] Thus, cyber threats are only credible in combination with traditional, threats. Figure 4 introduces one possible model adding levels of retaliation via cyberspace: Firstly, escalation beyond political and economic sanctions using low-level cyberattacks and, secondly, escalating kinetic strikes ahead of the use of tactical nuclear warheads.

[48] Cf. [Sh11], 97 and Interview [ED14]
[49] Cf. [ED14] and [Li], 62

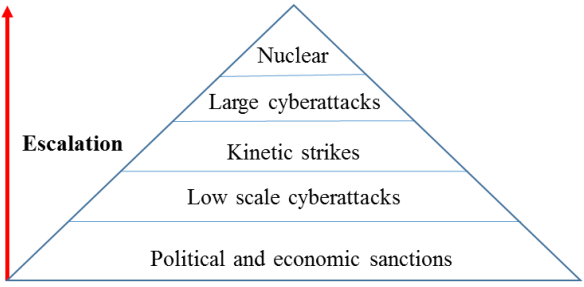

Abb. 2: A possible model of escalation[50]

Since knowing the opponent is crucial in designing defensive or offensive responses, attribution is core. At the moment of an incident, it is difficult to assess whether one is dealing with a technical accident, indiscriminately spreading malware, or a targeted attack. For deterrence-by-retaliation to work, aggressors must be convinced that they will be identified and punished; and attribution must be bulletproof to avoid creating new enemies and to convince third parties that the retaliatory measure is not an attack in itself. A 2013 NATO CCDCOE study lists anonymity as "[denying] identification of malicious actors, thus making deterrence policies futile, the undertaking of diplomatic, political and economic reaction measures difficult, and the application of legal remedies, e.g. countermeasures, impossible".[51] Still, "the attribution problem is partly artificial: As analyses of individual indictments following cyberattacks – both on the U.S. in 2013 and 2014 and Estonia in 2007 – demonstrate, attribution is costly and time-consuming, but possible. It combines technical measures and intelligence operations to pinpoint the culprit. Dealing with state actors, malware source code and the programming style can be mapped against previous incidents. Such forensics revealed similarities between Stuxnet and Flame, intensifying accusations of U.S.-Israeli involvement in both high-complexity tools.[52] Both long-term rivals of Iran were singled out long before Snowden provided proof. The major difficulty lies in prosecution once perpetrators are identified. In the Estonian case, Russia refused to prosecute indicted citizens, saying that their DDoS attacks did not amount to legal violations in Russia.[53]

3.4 What to deter? Narrow vs. broad deterrence

In narrow vs. broad deterrence the core question is "What is to be deterred?". In the current context, and based on initial definitions of operational and strategic cyberwar, a war of cyberattacks is highly unlikely. "Narrow" deterrence-by-retaliation requires unmistakable positioning as to what is acceptable and what is not. The outcomes of U.S. political cyberespionage and simultaneous indictment of Chinese officers for economic

[50] Own figure based on Interview [ED14]
[51] [Zi13], XVI
[52] Interview [NA14] and Interview [EE14]
[53] Interview [NA14]

cyberespionage may set an important precedent. The U.S. draws somewhat artificial lines between political and economic cyberespionage, indicting Chinese officers for industrial espionage while itself facing accusations of spying on OPEC and European arms manufacturers, hacking China's Huawei and supporting, or at least tolerating, British hacking into Belgium telecommunications giant Belgacom.[54] The use of cyberattacks such as Stuxnet risks escalation, given the lack of international agreement on what is acceptable. U.S. President Obama issued an executive order on 1 April 2015 calling for sanctions against individual and organised cyberattackers, especially those conducting attacks on critical infrastructure.[55] This is an important starting point.

NATO CCDCOE's Tallinn Manual[56] insisted on the applicability of Geneva Convention IV on the protection of civilians in cyberspace. This prohibits disproportional attacks against civilian and non-military installations. However, some discard specific "Cyber Geneva Conventions" as unverifiable and unlikely to yield results in the short- to medium-term, insisting instead on internationally agreed principles.[57] The UN's ITU sees hope in "effective norms against cyberaggression […] reining in unacceptable forms of behaviour".[58] International norms on restraint in cyberwarfare appear the only hope to resolving diplomatic upsets about cyberespionage and cyberattacks. The currently prevailing grey zones ultimately only serve non-state and criminal actors. As a first step, nations worldwide must determine what constitutes an illegal attack and warrants responses from the international community. The protection of civilian critical infrastructure should be among the leading goals of international efforts. The UN General Assembly and its working groups on cyber are the most appropriate forum to discuss which targets and which weapons are to be prohibited. Military and political strategists can thereafter develop a narrow deterrence regime against specific actors, to protect certain systems and to single out weapons to be deterred. Cyberpower cannot deter any aggression, but deterrence can be rendered effective by outlining retaliatory measures in case of cyberattacks against a narrow set of targets or using a narrow set of attack mechanisms.

3.5 For whom? Central vs. extended deterrence

The question of central vs. extended deterrence extends beyond whether EU and NATO allies possess collective defence mechanisms, in their respective Article 222 of the Treaty on the Functioning of the European Union and Article 5 of the North Atlantic Treaty. For extended deterrence to be effective, allies would have to signal to would-be-offenders that violations against one ally will result in a joint response. While the articles require considerable thresholds – the loss of human life arguably being one – the increasing connectedness of e.g. energy grids increases the threat for neighbours during

[54] Cf. [Fe14], cf. [Sp13a] and [Sp13b]
[55] [Th15]
[56] See [Sc13]
[57] Interview [EE14]
[58] Cf. [Ha10]

cyberattacks against one country. NATO doctrine recognizes cyber as part of NATO's collective defence and emphasises that retaliation may be "by any means necessary".[59] Similar to any public-private cooperation, however, the contentious item is granting access to one's networks. Public-private and civilian-military cooperation as well as inter-state collective cyberdefence requires a great deal of trust, since technical assistance can only be provided after having been granted wide-ranging access, possibly revealing additional vulnerabilities.[60] Governments may, furthermore, consider it little desirable to perform operational tasks in securing privately-owned critical infrastructure, which may lead to free-riding and insufficient investment in cybersecurity. Given this unwillingness to grant access, NATO-coordinated joint cyberattacks are unlikely. All nations have to, first and foremost, secure themselves. Technologically leading NATO allies should engage in more strategic dialogue with their partners to increase overall preparedness. This can then be the basis for extended deterrence. President Obama strengthened the idea of such concepts, stating that "the United States will respond to hostile acts in cyberspace as we would to any other threat to our country. … We reserve the right to use all necessary means – diplomatic, informational, military, and economic – as appropriate and consistent with applicable international law, in order to defend our Nation, **our allies, our partners**, and our interests."[61]

4 Conclusion and outlook

Previous academic research on deterrence theory and its application to cyberspace, leads to a series of conclusions for policy-making. Theorist and UK foreign policy advisor Lawrence Freedman differentiates 1) deterrence-by-retaliation and deterrence-by-denial; 2) immediate vs. general; 3) narrow vs. broad; and 4) central vs. extended deterrence. Taking into consideration the extensive literature review, cyberdeterrence strategies ought to be inclusive of both retaliatory and denying mechanisms. This chiefly includes four elements: Firstly, resistance by developing strong guidelines against voluntary and accidental disruptions; secondly, resilience to quickly and fully recover from attacks; thirdly, the definition of clear-cut global rules on acceptable practice in peacetime and on the legality of targets during armed conflict; finally, a national strategy of responses within these rules, ranging from criminal prosecution and political condemnation to (economic) sanctions and, finally, measures of active defence and retaliation. Any retaliatory strategy must also address the question of whether kinetic strikes are warranted in the case of sincere damage to the property and lives of citizens.

As "cyberweapons" require target-specific development, quick retaliatory cyberstrikes are impossible. Therefore, and due to the fact that attribution requires time-consuming forensics, retaliation with cyber means cannot be done ad-hoc. Defence and offence must thus be joined into a broader deterrence strategy. Those responsible for national defence

[59] Interview [NA14]
[60] Cf. [Li09], vxiii and Interview [NA14]
[61] [Th11], 14

should focus on protecting their own networks, most importantly communication and weapons control. This will add credibility to any future deterrence strategy. However, deterrence strategies should start not merely from the defence community, but hand-in-hand with activities of the broader government. Currently, rather than strategically thinking about how to tackle threats, authorities are still trying to understand threat agents and vulnerability landscapes. Classical deterrence requires defining interests and drawing redlines. Governments must clarify their interests and prioritize what is to be protected before deterrence postures can be effective. Which services and infrastructures are critical for the country's functioning? Which actors can ensure their security? Purely focusing on the military is inappropriate, and while the military is good at defending its own networks, it has hardly any experience in cooperating with the public and, even less so, the private sector. This, however, is an essential part of cybersecurity. To co-opt the private sector, governments have to create win-win-situations rather than imposing from above, and recognize that, unlike nuclear deterrence, cyberdeterrence is not a game of great powers and not even that of nation-states alone. Neither the European approach of imposed cooperation, as in the draft NIS Directive's mandatory reporting, nor the U.S. approach of largely leaving cybersecurity to market forces appear sufficient. Germany's National Cybersecurity Council (NCSR) and National Cyberdefence Centre (NCAZ) may provide starting points of such integration.

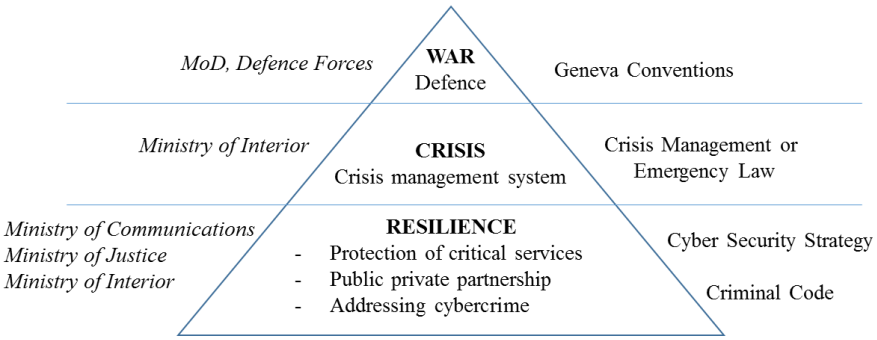

Abb. 3: EEAS figure on a possible inter-ministry division of labour[62]

Only the establishment of clear rules on acceptable practice in cyberspace will enable broad and effective responses to prevailing threats. Nations and individuals offending such treaties would then feel the deterring threat of retaliation with sanctions and criminals could be targeted with legal prosecution across national borders. The current grey zones in cyber conduct serve no one in the medium term. International agreements, similar to the Budapest Convention on Cybercrime[63], are crucial to preserving the internet as an engine of the global economy. The introduction of norms-based deterrence may thus provide a useful framework. Since cross-border cooperation is crucial, nation-states have to agree on acceptable thresholds, e.g. regarding the legality of

[62] Own figure based on [Ti12], 5
[63] Cf. [Co01]

cyberespionage. There is a sense of urgency in statements of high-ranking military and political leaders, acknowledging the risks of making cars, hospitals, energy grids and even prisons controllable via the Internet. Reluctance to tackle these questions politically will complicate matters, as the private sector engages in politics and vigilantism by striking back against botnets or by independently acting against censorship. Governments would most effectively strengthen the private sector by establishing legal clarity.

In order to achieve agreement on norms, it is sensible to build upon existing regimes such as the Geneva Conventions, establishing that offences through cyberspace warrant similar consequences as their traditional counterparts. The NATO CCDCOE's Tallinn Manual and its Peacetime Regime for State Activities in Cyberspace are important and should be seen as the basis for future efforts. In a first step, definitions on civilian critical infrastructures should be agreed upon internationally and undertaking or supporting attacks against such must be prohibited. This would move the discussion on the legitimacy of cyberattacks to the more relevant debate on which targets are acceptable, providing clarity for the development of effective military strategy. To achieve much-needed global progress, international disagreements over the appropriate governance of the internet or the provision of freedoms online should be dealt with separately.

List of references

[Be14] Bendiek, Annegret. “Tests of Partnership: Transatlantic Cooperation in Cyber Security, Internet Governance, and Data Protection.” SWP - German Institute for International and Security Affairs. March 2014.

[Be43] Bentham, Jeremy. The Works of Jeremy Bentham, 11 volumes. Edited by John Bowring. Vols. Edinburgh: William Tait, 1838-1843.

[Bl13] Blitz, James. “UK becomes first state to admit to offensive cyber attack capability.” Financial Times. 29 September 2013.

[BM14] BMVg Cybersecurity Expert, interview by Tobias Metzger. (24 June 2014).

[Bo13] Bologna, Sandro, Alessandro Fasani, and Maurizio Martellini. “From Fortress to Resilience.” In Cyber Security: Deterrence and IT Protection for Critical Infrastructures, edited by Maurizio Martellini, 53-56. Heidelberg: Springer, 2013.

[Br46] Brodie, Bernard, ed. The Absolute Weapon. New York, NY: Harcourt Brace, 1946.

[Ci12] Cilluffo, Frank J., Sharon L. Cardash, and George C. Salmoiraghi. “A Blueprint for Cyber Deterrence: Building Stability through Strength.” Military and Strategic Affairs, Vol. 4, No. 3, 2012: 3-23.

[CK12] Clarke, Richard A., and Robert Knake. Cyber War: The Next Threat to National Security and What to Do About It. New York: Ecco, 2012.

[Co01] Council of Europe. Convention on Cybercrime - Treaty Doc. 108-11, ETS No. 185. 23 November 2001.

[De15] Department of Defense. The DoD Cyber Strategy. Document released on 17 April 2015. www.defense.gov/home/features/2015/0415_cyber-strategy

[Du89] Dulles, John Foster. “The Evolution of Foreign Policy” In US Nuclear Strategy: A Reader, edit. by P. Bobbit, L. Freedman and G. Treverton. London: Macmillan, 1989.

[ED14] EDA Cybersecurity Official, interview by Tobias Metzger. (24 June 2014).

[EE14] EEAS Cybersecurity Official, interview by Tobias Metzger. (27 June 2014).

[Fe14] Ferranti, Marc. “NSA hacked into servers at Huawei headquarters, reports say.” Computerworld. 23 March 2014.

[Fr04] Freedman, Lawrence. Deterrence. Cambridge: Polity Press, 2004.

[GM13] Gaycken, Sandro, and Maurizio Martellini. “Cyber as Deterrent.” In Deterrence and IT Protection for Critical Infrastructures, edited by Maurizio

Martinelli, 1-10. Heidelberg: Springer, 2013.

[Gr14] Grossman, Lev. “Inside the Code War.” TIME, 2014: Vol. 184, No. 3: 20-27.

[GS74] George, Alexander L., and Richard Smoke. Deterrence in American Foreign Policy: Theory and Practice. New York: Columbia University Press, 1974.

[Ha10] Harknett, Richard J., John P. Callaghan, and Rudi Kauffman. “Leaving Deterrence Behind: War-Fighting and National Cybersecurity.” Journal of Homeland Security & Emergency Management, 2010.

[Kn10] Knopf, Jeffrey W. “The Fourth Wave in Deterrence Research.” Contemporary Security Policy, Vol. 31, No. 1 April 2010: 1-33.

[Li09] Libicki, Martin C. “Cyberdeterrence and Cyberwar - Prepared for the United States Air Force.” RAND Corporation. 2009.

[Li13] –. “Brandishing Cyberattack Capabilities.” RAND National Defense Research Institute. 2013.

[Li14] Lieber, Keir. “The Offense-Defense Balance and Cyber Warfare.” In Cyber Analogies, edited by Emily O. Goldman and John Arquilla, 96-107. Monterey: Naval Postgraduate School and U.S. Cyber Command, 2014.

[Ly10] Lynn III, William J. “Defending a New Domain: The Pentagon's Cyberstrategy.” Foreign Affairs Journal. September/October 2010.

[Ma10] Markoff, John, David E. Sanger, and Thom Shanker. “In Digital Combat, U.S. Finds No Easy Deterrent.” The New York Times. 25 January 2010.

[Ma12] Malone, Patrick J. Offense-Defense Balance in Cyberspace: A Proposed Model. Monterey: Naval Postgraduate School, 2012.

[Ma14] Mason, Rowena. “Labour urges strategic defence review to consider cyberwar threat.” The Guardian. 24 March 2014.

[Mo03] Morgan, Patrick. Deterrence Now. Cambridge: Cambridge University Press, 2003.

[Mo77] –. Deterrence: A Conceptual Analysis. Beverly Hills, CA: Sage Publications, 1977.

[NA13] NATO Cybersecurity Official, interview by Tobias Metzger. (30 April 2013).

[NA14] NATO Cybersecurity Official, interview by Tobias Metzger. (17 June 2014).

[Ox14] Oxford English Dictionary. Oxford English Dictionary online - "deter". 2014.

[Ph13] Philbin, Lt. Col. Michael J. “Cyber Deterrence: An Old Concept in a New Domain.” U.S. Army War College: Strategy Research Project. March 2013.

[Po08] Powell, Robert. Nuclear Deterrence Theory: The Search for Credibility. Cambridge: Cambridge University Press, 2008.

[Re14] Reuters. China bans use of Microsoft's Windows 8 on government computers.

20 May 2014.

[Sc13] Schmitt, Michael N., ed. Tallinn Manual on the International Law Applicable to Cyber Warfare. New York: Cambridge University Press, 2013.

[Sc66] Schelling, Thomas. Arms and Influence. New Haven: Yale University Press, 1966.

[SD14a] SDA. “Annual conference.” Conference Report: Overhauling transatlantic security thinking. 15 July 2014.

[SD14b] —. Evening debate: Critical infrastructure protection in the cyber-age. Brussels. 30 June 2014.

[Se12] Secretary of Defence Panetta, Leon E. Remarks by Secretary Panetta on Cybersecurity to the Business Executives for National Security, New York City. 11 October 2012.

[Sh11] Sheldon, John B. “Deciphering Cyberpower: Strategic Purpose in Peace and War.” Strategic Studies Quarterly, Summer 2011: 95-112.

[Sh14] Shaheen, Salma. “Offense-Defense Balance in Cyber Warfare.” In Cyberspace and International Relations: Theory, Prospects and Challenges, edited by Jan-Frederik Kremer and Benedikt Müller, 77-94. Heidelberg: Springer, 2014.

[Sp13a] Spiegel. “Belgacom.” Belgacom Attack: Britain's GCHQ Hacked Belgian Telecoms Firm. 20 September 2013.

[Sp13b] —. “OPEC.” Oil Espionage: How the NSA and GCHQ Spied on OPEC. 11 November 2013.

[St12] Stevens, Tim. “A Cyberwar of Ideas? Deterrence and Norms in Cyberspace.” Contemporary Security Policy, Vol. 33, No. 1 April 2012: 148-170.

[Th11] The White House. International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World. May 2011.

[Th15] The White House. Executive Order. "Blocking the Property of Certain Persons Engaging in Significant Malicious Cyber-Enabled Activities". 01 April 2015.

[Ti12] Tiirma-Klaar, Heli. “Presentation: National Cyber Security Strategies and International Cyber Policy.” European External Action Service. 22 October 2012.

[Zi13] Ziolkowski, Katharina, ed. Peacetime Regime for State Activities in Cyberspace. International Law, International Relations and Diplomacy. Tallinn: NATO CCDCOE, 2013.