The Tallinn Manual: Moving From a Theoretical Position to Universal Adoption in the Governance of Nation-State Cyber Weapons

Katie Passey

NEXTGEN 5.0

ABOUT US

NextGen 5.0 is a pioneering non-profit, independent, and virtual think tank committed to inspiring and empowering the next generation of peace and security leaders in order to build a more secure and prosperous world.

THE AUTHOR

Katie Passey is an aspiring researcher, specialising in online extremism, radicalisation and interventions. She is a postgraduate from the University of Southampton with an MSc in International Security and Risk

COPYRIGHT

CONTENTS

The laws relating to state-offensive weaponry derive from a wider body of law commonly referred to as the law of armed conflict. The purpose of these laws is to alleviate armed conflicts' worst consequences. However, not all weapons law treaties were negotiated on the basis of a commonly felt moral imperative to create the new law in question.[1] It is common for a party to a conflict that was fortunate enough to have an effective weapon, for which the opposition had no counter-measure, to exploit the resulting advantage.[2] Cyberwarfare is a substantial challenge to our conventional thinking about war and armed conflict since the development of nuclear weapons. However, the difference between nuclear weapons and cyber weapons is that where nuclear technology made war unthinkable, cyber weapon technology makes it irresistible.[3] States have been developing their offensive cyber-capabilities and in some cases engaging in military-led cyber-operations against other states. Research has compared this development to a Cool War, where nation-state cyber-operations 'share the Cold War trait of not involving hot conflict on the battlefield, but are warmer than cold because they involve almost constant offensive measures that, while falling short of actual warfare, regularly seek to damage or weaken rivals'.[4]

It is for these reasons that no prior international treaty addressing the dangers cyber weapons pose has been agreed. The NATO Cooperative Cyber Defence Centre for Excellence first created the Tallinn Manual in 2013 which began to decipher the international law applicable to cyberwarfare, but it was designed to produce a non-binding document applying existing law to cyberwarfare.[5] Following a succession of state offensive cyber-operations, from the discovery of Stuxnet on Iran's nuclear facility in 2010, cyberattacks on Ukraine in

---

[1] William H. Boothby, *Weapons and the Law of Armed Conflict* (Oxford: Oxford University Press, 2009) p. 1.

[2] Ibid.

[3]David Rothkopf, 'The Cool War', *Foreign Policy*, 20 February 2013 <http://foreignpolicy.com/2013/02/20/the-cool-war/> [accessed 08 May 2018].

[4] Ibid.

[5] Michael N. Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge: Cambridge University Press, 2013) p. 1.

2015, to the global eruption of WannaCry in 2017, this essay will attempt to adapt the theoretical position assumed by the Tallinn Manual into a binding agreement between the NATO Alliance in relation to the principle of discrimination, booby-traps and starvation, to set the parameters in offensive state cyber-operations.

PRINCIPLE OF DISCRIMINATION

One of the main guiding principles behind the laws of weaponry is the principle of discrimination between combatants and non-combatants, which prohibits weapons that cannot be directed against only military objectives.[6] The Tallinn Manual stated that an indiscriminate cyber weapon qualifies when 'their nature generate effects that are incapable of being controlled and therefore can spread uncontrollably into civilian… computers'.[7]

A prime example of such a cyber weapon is computer worms, which lack the principle of discrimination due to their ability to self-replicate across computer networks. The uncontrollable nature of WannaCry in 2017 illustrated the devastating effect a weaponised computer worm can possess. As can be seen in Fig. 1, WannaCry spread around 150 countries, infecting 400,000 civilian computers in less than a week.[8] An overlooked aspect of this is that WannaCry had the potential to cause harm, even death, to civilians. WannaCry shutdown a large section Britain's National Health Service IT infrastructure; appointments and operations were cancelled and some trusts had to divert patients to other accident and emergency departments.[9] Although it is suspected that WannaCry was unleashed by a North

---

[6] Helen Frowe, *The Ethics of War and Peace: An Introduction* (Abingdon: Routledge, 2011) p. 112.

[7] Schmitt, *Tallinn Manual*, p. 145.

[8] Symantec, 'What you need to know about the WannaCry Ransomware', *Symantec Blogs*, 23 October 2017 <https://www.symantec.com/blogs/threat-intelligence/wannacry-ransomware-attack> [accessed 08 May 2018].

[9] UK Department of Health, 'Investigation: WannaCry cyber attack and the NHS', *National Audit Office*, 25 April 2018 <https://www.nao.org.uk/wp-content/uploads/2017/10/Investigation-WannaCry-cyber-attack-and-the-NHS.pdf> [accessed 08 May 2018] p. 14.

Korean cybergang, the vulnerabilities the computer worm exploited - EternalBlue and DoublePulsar - were stolen from the US National Security Agency.[10] It is difficult to imagine an instance in which a state-actor could have used these exploits and not faced a similar situation of uncontrollable proliferation.

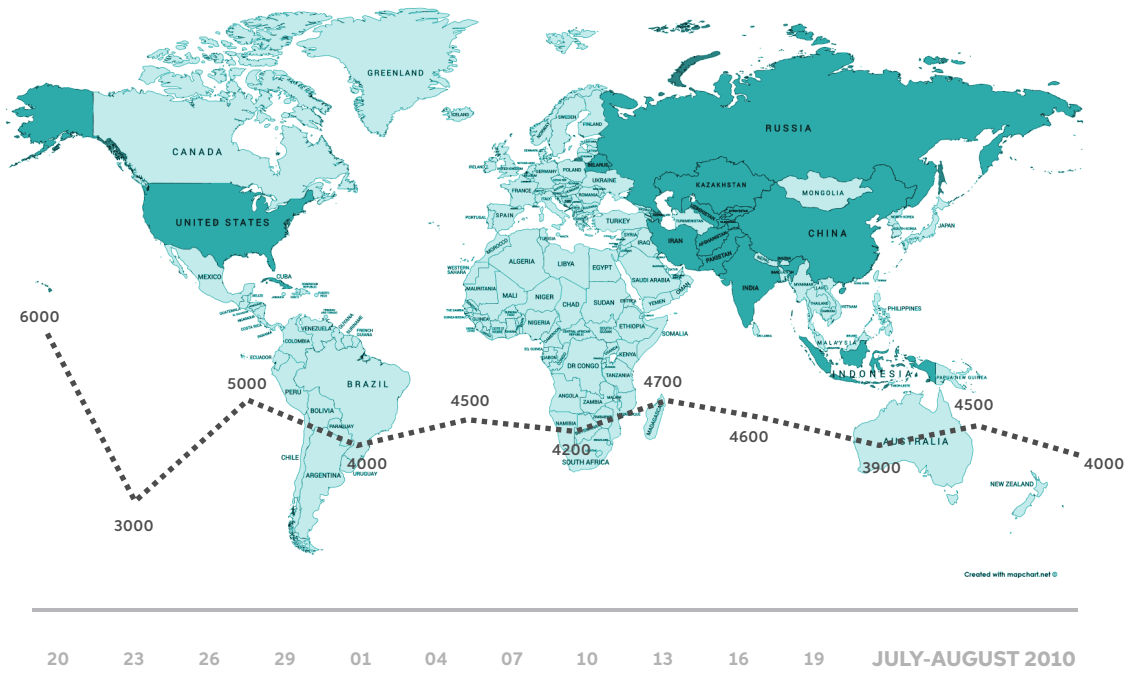
FIGURE 1: Number of exploit attempts blocked by Symantec of Windows vulnerabilities used by WannaCry per day from 11 May 2017 to 22 May 2017, laid over a map showing the affected countries. [Created by Katie Passey].

On the other hand, Stuxnet demonstrates how a computer worm can fulfil the principle of discrimination. Stuxnet is an infamous nation-state cyber-operation that infected the software of fifteen industrial sites in Iran, including a uranium-enrichment plant, discreetly reconfiguring the communication between the computers and

[10] Andy Greenberg, 'Hold North Korea Accountable for WannaCry – and the NSA, too', *Wired*, 19 December 2017 <https://www.wired.com/story/korea-accountable-wannacry-nsa-eternal-blue/> [accessed 08 May 2018].

centrifuges to delay Iran's nuclear armament.[11] Although Stuxnet spread beyond the initial targets in Iran, as shown in Fig. 2, compared to WannaCry, it had an inert infection rate; infecting around 100,000 computers in an estimated twenty countries over a period of two months.[12] Moreover, the computer systems infected outside of the nuclear targets in Iran were not subject to the worm's weapon. Stuxnet was designed to locate a network of 984 converters, running the exact Siemens Step 7 software in Iran's nuclear program before activating; the worm deleted itself if it could not locate the software.

FIGURE 2: Symantec's recorded total rate of infection of new IP addresses by Stuxnet between 20 July 2010 and 19 August 2010, laid over a map showing the affected countries. [Created by Katie Passey].



[11]Paul Mueller and Babak Yadegari, 'The Stuxnet Worm', *University of Arizona*, 2012 <http://www2.cs.arizona.edu/~collberg/Teaching/466-566/2012/Resources/presentations/20 12/topic9-final/report.pdf> [accessed 08 May 2018] p. 1.

[12] Nicolas Falliere, Liam O. Murchu, and Eric Chien, 'W32.Stuxnet Dossier', *Symantec*, February 2011[accessed 08 May 2018] p. 6-7.

Consequently, Stuxnet is considered the most ethical cyber weapon to have delayed Iran's nuclear capacity building. Comparative actions aimed at achieving a similar goal include Operation Opera in 1981, an Israeli airstrike that successfully destroyed Iraq's Tammuz-1 nuclear reactor, but killed eleven people in the process.[13] Another comparative example was an assassination campaign that killed four Iranian nuclear scientists between 2010 and 2012.[14] These assassinations also took place in public, increasing the risk to civilians and causing fear and intimidation akin to terrorism. Given these important factors, the first proposal of this essay towards a cyber weapons treaty is a precision ethos: cyber weapons, in particular computer worms, must be designed conscientiously with regard to the proliferating impact on non-combatants.

CYBER BOOBY-TRAPS

The term booby-trap is defined to mean 'any device or material which is designed, constructed or adapted to kill or injure, and which functions unexpectedly when a person disturbs or approaches an apparently harmless object or performs an apparently safe act'.[15] The purpose of this provision is to protect the civilian population from coming into contact with these devices both during and even after a conflict has expired (for example the ban on landmines and cluster munitions), and to prohibit the treacherous or perfidious use of such devices.

The term 'device or material' is intentionally capable of broad interpretation in order to account for a number of innocuous situations. For example, the legal provision includes booby-traps associated with food or drink, products specifically designed for the use

---

[13] Sasha Polakow-Suransky, *The Unspoken Alliance: Israel's Secret Relationship with Apartheid South Africa* (US: Pantheon Books, 2010) p. 145.

[14] Ronen Bergman, 'When Israel Hatched a Secret Plan to Assassinate Iranian Scientists', *Politico Magazine*, 05 March 2018 <https://www.politico.com/magazine/story/2018/03/05/israel-assassination-iranian-scientists-217223> [accessed 08 May 2018].

[15] Schmitt, *Tallinn Manual*, p. 147.

of children, medical activities and much more. The Tallinn Manual found that there was 'no reason as a matter of law to differentiate between a physical object that serves as a booby-trap and cyber means of achieving an equivalent objective'.[16] However, the Tallinn Manual does dictate that the cyber booby-trap must be deliberately configured to operate unexpectedly, be designed, constructed or adapted to kill or injure, appear innocuous to a reasonable observer, and must not be associated with certain specified objects.[17]

A prime example of a state-sponsored cyber booby-trap is a malware pay-loaded cyber communication. For example, an email containing malicious macros is received by an employee of a water treatment plant purportedly from their physician. The macros suspend the water purification process at the plant, allowing untreated water into the water supply of a military barracks aimed to kill soldiers. This use of malware to deliver a cyber weapon is an unlawful cyber booby-trap because the recipient reasonably believed that the act of opening the email from their 'physician' was safe. Therefore, the second proposal of this essay towards a cyber weapons treaty is to prohibit the use of cyber booby-traps that adhere to the protocols laid out in the Tallinn Manual that derive from Article 6 of Protocol II in International Law.

STARVATION

Starvation of civilians is a prohibited method of warfare that relates to 'attacking, destroying, removing or rendering useless objects indispensable to the survival of the civilian population'.[18] Objects indispensable to the survival of the civilian population currently include items such as foodstuffs, agricultural areas for the production of foodstuffs, crops, livestock, drinking water installations and supplies, and irrigation works.

---

[16] Schmitt, *Tallinn Manual*, p. 147.

[17] Ibid.

[18] Boothby, *Weapons and the Law of Armed Conflict*, p. 46.

The Tallinn Manual acknowledged that these examples are not exhaustive.[19] The term 'starvation' covers not only the meaning of starving, as in killing by hunger or depriving of nourishment, but also the general meaning of deprivation or insufficient supply of some essential commodity, of something necessary to live.[20] As a result, other targets, such as dual-use power supply installations, can constitute an object indispensable to the survival of the civilian population; the elimination of the power supply network will cause considerable disruption to other elements of civilian infrastructure, for example drinking water installations. America's 1991 Operation Desert Storm destroyed thirteen of Iraq's twenty electricity-generating facilities via allied bombs. It highlighted the consequences of disrupting the highly interconnected critical infrastructure, as these attacks on dual-use power facilities caused cascading damage throughout the water purification and sanitation systems.[21] Moreover, there is room to argue that considering the internet itself is increasingly viewed as a general purpose technology - the access of which the United Nations Human Rights Council has considered a human right - denial of access to this commodity also could be considered an act of starvation.[22]

A cyber weapon can have the same impact as the conventional weapons used in Operation Desert Storm, as proven in 2015 following the shutdown of Ukraine's power grid. In December 2015, three energy distribution companies in Ukraine were compromised by a cyberattack, allegedly conducted by Russian Government affiliated hackers Sandworm, which disrupted the electricity supply to three regions in

---

[19] Schmitt, *Tallinn Manual*, p. 227.

[20] Knut Dörmann, 'Preparatory Commission for the International Criminal Court: The Elements of War Crimes', *IRRC*, 83 (2001), 461-487 (p. 475).

[21] Michael Knights, 'Infrastructure Targeting and Postwar Iraq', *The Washington Institute*, 14 March 2003 <http://www.washingtoninstitute.org/policy-analysis/view/infrastructure-targeting-and-postwar-iraq> [accessed 08 May 2018].

[22] Emma Boyle, 'A New Resolution from the UN Condemns countries that deliberately disrupt the internet access of their citizens', *The Independent*, 5 July 2016 <https://www.independent.co.uk/life-style/gadgets-and-tech/un-declares-online-freedom-to-be-a-human-right-that-must-be-protected-a7120186.html> [accessed 07 July 2018].

Ukraine.[23] The energy companies were able to restore service relatively quickly. However, the power-cut had the potential to disrupt other critical infrastructure, such as Internet and drinking water installations. Therefore, the final proposal of this essay towards a cyber weapons treaty is to prohibit the intentional deprivation of civilians through a cyber means of objects indispensable to their survival, including Internet and dual-use power facilities.

CONCLUSION

One challenge to adapting the theoretical position assumed by the Tallinn Manual into a binding agreement between the NATO Alliance is the reality that not all weapons law treaties are negotiated on the basis of a commonly felt moral imperative to create the new law in question. Consequently, this essay anticipates difficulties with the NATO Alliance to agree with the cyber weapons treaty rules suggested by this essay. Differences of opinion in relation to nuclear weapons similarly divide NATO members, mainly as a result of whether nuclear weapons represent an essential hedge against the unknown.[24] Nevertheless, NATO has continually maintained a commitment to arms control as an integral part of its security policy to seek stability and security for all. NATO, despite challenges, has been able to establish an effective policy towards nuclear warfare; the same could be achieved with cyber weapons in spite of hesitations.

A further challenge to adapting the theoretical position assumed by the Tallinn Manual into a binding agreement between the NATO Alliance relates to regulation and response. The lack of definitive attribution behind attacks makes it very easy for States to maintain a façade of adhering to the formal obligations under this cyber weapons treaty, while nevertheless exercising and developing the capabilities to

---

[23] Andy Greenberg, "Crash Override': The Malware that Took Down a Power Grid', *Wired*, 06 June 2017 <https://www.wired.com/story/crash-override-malware/> [accessed 08 May 2018].

[24] Malcolm Chalmers and Simon Lunn, 'NATO's Tactical Nuclear Dilemma', *RUSI*, March 2010 <https://rusi.org/sites/default/files/201003_op_natos_tactical_nuclear_dilemma.pdf> [accessed 08 May 2018] p.11.

conduct a cyber-operation in defiance with the treaty. The lack of attribution also makes it difficult for NATO to measure an appropriate response when an Alliance member has been subjected to/or used a weapon prohibited by this treaty. For example, in 2007, Estonia experienced extensive cyberattacks, allegedly conducted by Russia. Estonia requested that the cyberattack trigger NATO's Article 5, which commits NATO to respond to attacks on any member of the Alliance as permitted under the UN Charter provision in Article 51 for collective self-defence.[25] However, NATO did not respond with a counterattack. Taking into consideration the Cool War trait of offensive cyber operations involving almost constant offensive measures and the inability to provide clear evidence to attribute another state, Article 5 should only be triggered in the most extreme circumstances; otherwise it may be subject to overuse and risk escalating tensions to all-out war when resolution may have been achieved through diplomatic discussion. These matters are beyond the scope of this essay, but it is crucial that follow up research is conducted on how to determine compliance with such a cyber weapons treaty, how to establish responsibility in the event of a cyberattack on a NATO nation, or by a NATO nation, with a weapon prohibited by this treaty, and how to measure an appropriate response.

[25] Ian Traynor, 'Russia accused of unleashing cyberwar to disable Estonia', *The Guardian*, 17 May 2007 <https://www.theguardian.com/world/2007/may/17/topstories3.russia> [accessed 08 May 2018].

ABOUT US

We are a non-profit virtual think tank committed to inspiring and empowering the next generation of global leaders in order to build a more secure and prosperous world.

FOLLOW US

Facebook: **@NextGen50ThinkTank**
Twitter: **@NextGen_50**
Linkedin: **@next-gen50**
Instagram: **@nextgen_50**