5

WHO DID IT? ATTRIBUTION OF CYBER INTRUSIONS AND THE *JUS IN BELLO*

William Banks[1]

I. INTRODUCTION

The central concepts that make up the law of armed conflict (LOAC) have not been easy to adapt to cyber operations.[2] In addition to their kinetic history and orientation, the core LOAC principles do not in most instances anticipate the kind of cyber-specific analysis that should accompany the use of increasingly advanced cyber systems and tools in conflict. Cyber operations

---

[1] Director, Institute for National Security and Counterterrorism, Board of Advisers Distinguished Professor, Syracuse University College of Law and Maxwell School of Citizenship & Public Affairs, Syracuse University. The author appreciates the helpful feedback from the participants in the Lieber Institute for Law and Land Warfare workshop on the Impact of Emerging Technology on the Law of Armed Conflict at the USMA, West Point, October 2017, and thanks Taylor Henry, Syracuse University College of Law, J.D. 2018, for excellent research assistance.

[2] *See* OFFICE OF GEN. COUNSEL, U.S. DEP'T OF DEFENSE, DEPARTMENT OF DEFENSE LAW OF WAR MANUAL § 16.2.1, at 988 (2015), https://www.defense.gov/Portals/1/Documents/DoD_Law_of_War_Manual-June_2015_Updated_May_2016.pdf [hereinafter DOD LAW OF WAR MANUAL]; Michael Gervais, *Cyber Attacks and the Laws of War*, 30 BERKELEY J. INT'L L. 525, 579 (2012).

rarely cause physical damage, much less injury or death.[3] More often they cause cyber harm—by corrupting, manipulating or stealing data, denying access to a website, or interfering temporarily with the functionality of information systems. Or they indirectly disrupt or damage objects that are not part of the cyber domain. Measuring the harm from a cyber incident and calculating that harm in ways that the LOAC credits remains challenging, as does defining and distinguishing civilian and military objects, and accounting for the indirect effects of cyber operations.[4] Nor has the LOAC settled on a legal status for critical national security-related components of the cyber domain, including data and dual-use infrastructure.

---

[3] *See* Sue Halpern, *US Cyber Weapons: Our 'Demon Pinball,'* N.Y. REV. BOOKS (Sept. 29, 2016), http://www.nybooks.com/articles/2016/09/29/us-cyber-weapons-our-demon-pinball/ (describing the software worm Stuxnet that destroyed thousands of centrifuges at the Natanz nuclear enrichment facility between 2008 and 2010); DAVID E. SANGER, CONFRONT AND CONCEAL 188–225 (2012) (same, including the Olympic Games mission of the Obama administration); KIM ZETTER, COUNTDOWN TO ZERO DAY: STUXNET AND THE LAUNCH OF THE WORLD'S FIRST DIGITAL WEAPON (2014) (same).

[4] *See* Michael N. Schmitt & Eric W. Widmar, *"On Target": Precision and Balance in the Contemporary Law of Targeting*, 7 J. NAT'L SEC. L. & POL'Y 379, 395 (2014); Noam Lubell, *Lawful Targets in Cyber Operations: Does the Principle of Distinction Apply?*, 89 INT'L L. STUD. 252, 268−69 (2013); Peter P. Pascucci, *Distinction and Proportionality in Cyberwar: Virtual Problems with a Real Solution*, 26 MINNESOTA J. INT'L L. 419, 431, 448−49 (2017).

Intuitively, we do not think of cyber weapons in the same way we do kinetic arms. Cyber seldom involves the use of force, is not thought of as constituting an armed attack, and is not by itself likely to trigger or become an armed conflict. Yet the LOAC commentary has gone to great lengths in recent years to show how these terms and principles derived from kinetic warfare can be applied to the cyber domain.[5]

Meanwhile, little attention has been paid to what might be thought of as a threshold question that could be usefully posed in applying traditional LOAC principles to cyber incidents in armed conflict: Who is responsible for the cyber operations absorbed by the States in an armed conflict? Is the enemy in the kinetic conflict responsible for the incoming cyber intrusions? Tracing the source and responsibility for a cyber operation can be challenging, and the possibilities for proxies, anonymity, and spoofing add uncertainties and complexity to an already daunting task.[6]

---

[5] *See, e.g.*, Harold Hongju Koh, *International Law in Cyberspace*, 54 HARV. INT'L L. J. ONLINE 1, 4–5 (2012), http://www.harvardilj.org/wp-content/uploads/2012/12/Koh-Speech-to-Publish1.pdf; TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS 3 (Michael N. Schmitt & Liis Vihul eds., 2017) [hereinafter TALLINN MANUAL 2.0].

[6] *See* John P. Carlin, *Detect, Disrupt, Deter: A Whole-of-Government Approach to National Security Cyber Threats*, 7 HARV. NAT'L SEC. J. 391, 396–97, 409 (2016); David D. Clark & Susan Landau, *Untangling Attribution*, 2 HARV. NAT'L SEC. J. 323, 327, 329 (2011), http://harvardnsj.org/wp-content/uploads/2011/02/Vol-2-Clark-Landau.pdf; CLEMENT

Why bother? There are at least two good reasons to determine responsibility for cyber operations during armed conflict. First, in today's climate of increasing cyber conflicts between states and between states and non-state criminals, hackers, and terrorists, adversaries expect that they may use cyber operations to attack anonymously with impunity. In short, attribution may easily be assumed but mistaken. Sometimes mistaken attribution does not make a cyber operation unlawful, but at other times it does. Second, international law, including the LOAC, applies to cyber activity in armed conflict only when there is "a nexus between the cyber activity in question and the conflict."[7] In other words, the cyber operations have to be connected in some way to the armed conflict for the LOAC to apply. That determination cannot be made confidently in the cyber domain without an attribution process that looks beyond the machines involved to the persons or entities responsible for what those computers or systems do. Although LOAC targeting and precautions analysis takes into account some of the same intelligence that would be part of an attribution process, establishing state responsibility and attribution *before* responsive targeting could strengthen the lawful application of the LOAC in armed conflict.

This chapter concludes that even a rudimentary process designed to attribute cyber intrusions may accomplish important objectives in armed conflict. First, States responsible for

---

GUITTON, INSIDE THE ENEMY'S COMPUTER: IDENTIFYING CYBER ATTACKERS 5, 11 (2017); Koh, *supra* note 8, at 6, 8.

7 TALLINN MANUAL 2.0, *supra* note 5, R. 80(5), at 376.

harmful cyber operations would be on notice that they may be held accountable for their cyber activity, including unlawful acts. Second, military commanders would have more reliable guidance in response targeting during the armed conflict, whether through cyber or non-cyber means. Better targeting guidance could, in turn, enhance compliance with LOAC. Relatedly and third, States may avoid making unlawful mistakes in the armed conflict—targeting civilian but arguably dual-use cyber infrastructure or failing to take available precautions knowing more about potential targets—because of weak or non-existent efforts to attribute incoming attacks. In the aggregate, attribution of cyber attacks in an armed conflict may act as a deterrent to unlawful uses of cyber tools and serve to better protect civilians, particularly if the attributed attacks expose an enemy State's cyber attacks against civilians or civilian cyber infrastructure.[8]

---

[8] Clark & Landau, *supra* note 6, at 352. To date, there are no clear examples of a civilian population being severely affected by cyber operations during armed conflict. *See* Cordula Droege, *Get Off My Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians*, 94 INT'L REV. RED CROSS 533, 539 (2012); *see, e.g.*, MICHAEL CONNELL & SARAH VOGLER, CTR. FOR STRATEGIC STUDIES, RUSSIA'S APPROACH TO CYBER WARFARE 18 (2016), https://www.cna.org/CNA_files/PDF/DOP-2016-U-014231-1Rev.pdf ("[T]he overall impact of the [Russian] cyberattacks [in Georgia] was minimal – Georgia's IT infrastructure was limited in 2008, and the Georgian government was eventually able to reroute most of its through servers in other countries."). *But see* David Hollis, *Cyber War Case Study: Georgia 2008*, SMALL WARS FOUNDATION (Jan. 6, 2011), http://smallwarsjournal.com/jrnl/art/cyberwar-case-study-georgia-2008 (describing Georgian

citizens being unable to access government web sites for information and instructions during armed conflict with Russia); CONNELL & VOGLER, *supra*, at 19 ("Russia has been able to compromise the Ukrainian government and military's ability to communicate and operate, thereby undermining the legitimacy and authority of Ukrainian political and military institutions."). In December 2015, Ukraine was subjected to what is believed to be the first cyberattack on another country's electric power grid. *Id*. at 20. Cyber attacks took three Ukrainian power distribution centers offline, causing outages that affected more than 220,000 citizens for periods spanning from one to six hours. *Id*. The overall effect of the attack has been described as limited, although the power company's distribution centers were not fully operational for several months. *Id*. The attackers also executed a telephone denial of service attack on the power company's call center, preventing customers from being able to call customer support during the outages. ROBERT M. LEE ET AL., ELEC. INFO. & SHARING CTR., ANALYSIS OF THE CYBER ATTACK ON THE UKRAINIAN POWER GRID 12 (2016), http://www.nerc.com/pa/CI/ESISAC/Documents/E-ISAC_SANS_Ukraine_DUC_18Mar2016.pdf. *See also* Eric Talbot Jensen, *Cyber Warfare and Precautions Against the Effects of Attacks*, 88 TEX. L. REV. 1533, 1540 (2010) (noting the "natural integration" of cyber attacks with future kinetic attacks, a trend that will "almost certainly" continue).

International law has relatively little to say about the obligations of States to identify the perpetrator of cyber intrusions,[9] and such law as exists resides in the *jus ad bellum*. There the law of attribution aims to identify and place responsibility for internationally wrongful acts.[10] The analysis is conducted after the fact regarding incidents below the armed conflict or attack threshold. The essence of the legal rule is simple: There can be no state responsibility for internationally wrongful acts until those acts have been attributed to a state.[11]

The *jus ad bellum* law of attribution has no bearing on legal obligations during an armed conflict. Yet the modest *ad bellum* attribution requirements, further addressed in Part I below, may provide guidance in evaluating whether and to what extent some legal principles for attribution of cyber intrusions could be usefully extended to the *jus in bello*. The LOAC does not address cyber attribution, apparently presuming that cyber intrusions that occur during an armed conflict are simply a part of the conflict, subject to LOAC principles. However, it is not necessarily the case that cyber intrusions suffered by a State are attributable to the other State engaged in the armed conflict. Nor is all cyber activity during an armed conflict necessarily connected to the hostilities between the states in conflict.[12] For the same reason, nor is all cyber activity during an armed

---

[9] TALLINN MANUAL 2.0, *supra* note 5, at 79–110.

[10] *See id.* R. 14, at 84–87.

[11] *Id.* R. 14(1), at 84.

[12] TALLINN MANUAL 2.0, *supra* note 5, R. 80(6), at 376 (citing the example of a cyber operation in pursuit of commercial secrets undertaken by State A while in armed conflict with State B).

conflict necessarily subject to the LOAC. Assuming a sort of corporate enemy posture for hostile acts during an armed conflict may facilitate operational decision making. Yet errors in cyber attribution could lead to unlawful responses directed at the enemy State and failures to identify and respond in lawful ways to other cyber intruders.

This chapter explores in a preliminary way the potential benefits of adding an attribution component to the LOAC. The chapter asks whether the principles of the LOAC are well served by treating the enemy State as the functional equivalent of a corporate enemy in an armed conflict, including a presumption that any cyber attacks suffered are its responsibility. The difficulty of accommodating dual-use cyber infrastructure and the data resident on many cyber systems within traditional LOAC doctrine underscores the shortcomings in protecting civilians and civilian objects during armed conflict. The article also preliminarily considers elements of an attribution process that could be grafted onto the LOAC for the cyber components of armed conflict.

It is important to clarify that the category of cyber operations considered in this chapter include those that do not rise to the level of attack under relevant law, defined in Additional Protocol I (API) as "acts of violence against the adversary, whether in offence or defence."[13] This

---

The International Group of Experts (IGE) convened to develop *Tallinn 2.0*, were split on whether the commercial secrets operation would be subject to LOAC. *Id*.

[13] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts art. 19, June 8, 1977, 1125 U.N.T.S. 3

chapter assesses cyber operations that produce a cyber effect during an armed conflict. A cyber effect may consist of adverse effects on an information system or access to a public facing website and corruption, manipulation, or loss of data even where there is no corresponding impact on the functionality of a cyber system.[14] The widespread colloquial use of the term "attack" or "cyber attack" to refer to various types of malicious cyber activities are not necessarily "attacks" under the LOAC.[15] The 2015 Department of Defense (DoD) Law of War Manual lists characteristics that render a cyber operation not an "attack" under the LOAC, including defacing a government webpage; a minor, brief disruption of internet services; briefly disrupting, disabling, or interfering with communications; and disseminating propaganda.[16] Although the DoD Law of War Manual concludes that cyber operations that are not attacks are not restricted by the rules that apply to

---

[hereinafter Additional Protocol I]. *Tallinn 2.0* defines cyber attack as "a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects." TALLINN MANUAL 2.0, *supra* note 5, R. 92, at 415.

[14] Pascucci, *supra* note 4, at 453.

[15] DOD LAW OF WAR MANUAL, *supra* note 2, §16.5.2, at 996.

[16] *Id.*

attacks, including targeting restrictions,[17] the law remains unsettled. [18] Thus, it remains unclear whether all cyber activity that occurs during armed conflict is subject to the LOAC, or whether instead the cyber operations must be connected in some way to the conflict or cross some threshold of harm to civilians before the LOAC applies.[19] In any event, some cyber operations that are not "attacks" according to the LOAC nonetheless produce a cyber effect and will be considered here on the assumption that they take place in connection to an armed conflict and that their attribution may serve the humanitarian objectives of the LOAC.

Arguably the corporate enemy presumption should not extend to cyber operations conducted during armed conflict. A growing array of cyber operations may occur during armed conflict. Some are reasonably presumed to be a means for conducting a military campaign, but

---

[17] ABA STANDING COMM. ON LAW AND NAT'L SEC., DOD LAW OF WAR MANUAL REVIEW WORKSHOP REPORT 65–66 (2016) [hereinafter SCOLANS REPORT], https://www.americanbar.org/content/dam/aba/administrative/law_national_security/DOD%20REVIEW%20OF%20THE%202015%20LAW%20OF%20WAR%20MANUAL%20-%202016%20Workshop.authcheckdam.pdf.

[18] One subset of the IGE believed that the LOAC applies to any cyber activity conducted by a party to an armed conflict against its opponent, while another group indicated that the cyber activity must have been undertaken in furtherance of the hostilities for LOAC to apply. TALLINN MANUAL 2.0, *supra* note 5, R. 80(6), at 376.

[19] *Id.*

others are less clearly connected to the conflict or are not related to the conflict at all. These cyber operations may originate from a third state or multiple states and could be the responsibility of these other states or a non-state actor.

Consider three examples:

- During the 1999 NATO bombing campaign to force Serbian military units out of Kosovo, widespread cyber attacks through direct denial of service (DDoS) and virus-infected emails occurred against NATO and member State governments and militaries. At the time, some media reports pointed to Serbian military responsibility for the attacks.[20] Other reports pointed to the Serbian Black Hand and Russian Hacker Brigade as responsible for some of the attacks. Another report claimed that another set of attacks was "clearly tied" to

[20] Jason Healey, *Cyber Attacks Against NATO, Then and Now,* ATLANTIC COUNCIL (Sept. 6, 2011), http://www.atlanticcouncil.org/blogs/new-atlanticist/cyber-attacks-against-nato-then-and-now. These cyber incidents included an upsurge of defacements of DOD websites. *Id*. *See also* Ellen Messmer, *Serb supporters sock it to NATO, U.S. Web sites*, CNN (Apr. 6, 1999), http://edition.cnn.com/TECH/computing/9904/06/serbnato.idg/index.html ("The same week a U.S. F-117A stealth fighter was lost over Yugoslavia, a NATO Web server here was shot down by denial-of-service attacks, which NATO sources strongly suspect came from the Serbian military, not independent hackers.").

nationalist Chinese hackers.[21] It remained unclear whether any of the hackers worked directly for the Yugoslav/Serbian military.[22]

- In the summer of 2008, tensions between Georgia and its South Ossetia region prompted Georgia to send troops to South Ossetia. Russia responded by launching air attacks throughout Georgia and invading South Ossetia.[23] In response to Georgian attacks in South Ossetia, Russia conducted airstrikes on Georgian targets and sent military units into South Ossetia to support the separatist region in its conflict with Georgia.[24] As the kinetic conflict began, Georgian government web sites began to crash.[25] Some experts opined that the

---

[21] Healey, *supra* note 20.

[22] Kenneth Greers, *Cyberspace and the changing nature of warfare*, SC MEDIA (Aug. 27, 2008), https://www.scmagazine.com/cyberspace-and-the-changing-nature-of-warfare/printarticle/554872/. *See also* MYRIAM DUNN CAVELTY, CYBER-SECURITY AND THREAT POLITICS: U.S. EFFORTS TO SECURE THE GOLDEN AGE 77 (2007) ("The question remains whether any of these attacks were state-sponsored…").

[23] JIM NICHOL, CONG. RESEARCH SERV., RUSSIA-GEORGIA CONFLICT IN AUGUST 2008: CONTEXT AND IMPLICATIONS FOR U.S. INTERESTS 2 (2009), https://fas.org/sgp/crs/row/RL34618.pdf.

[24] Michael Schwirtz, Anne Barnard & C.J. Chivers, *Russia and Georgia Clash Over Separatist Region*, N.Y. TIMES (Aug. 8, 2008), http://www.nytimes.com/2008/08/09/world/europe/09georgia.html?mcubz=0.

[25] *Id.*

Georgia case was the first time a known cyber attack had coincided with a shooting war.[26] Cyber attacks in July of 2008 against Georgian Internet infrastructure appear to have been a dress rehearsal for the kinetic conflict with Russia. The attack spread to computers throughout the Georgian government after Russian troops entered South Ossetia.[27] Some sources maintained that the cyber operations could not be attributed to Russia because there was no evidence that the Russian government conducted or facilitated the attack, or that

[26] John Markoff, *Before the Gunfire, Cyberattacks,* N.Y. TIMES (Aug. 12, 2008), http://www.nytimes.com/2008/08/13/technology/13cyber.html?mcubz=0.

[27] *Id. See also* ENEKEN TIKK ET AL., COOP. CYBER DEF. CTR. OF EXCELLENCE, CYBER ATTACKS AGAINST GEORGIA: LEGAL LESSONS IDENTIFIED 12 (2008), http://www.ismlab.usf.edu/isec/files/Georgia-Cyber-Attack-NATO-Aug-2008.pdf.

the evidence was unclear.[28] Several other sources attributed the attack to the Russian government.[29]

---

[28] Markoff, *supra* note 26. The independent, non-profit research institute, U.S. Cyber Consequences Unit (US-CCU), determined that the cyber attacks were carried out by "civilians with little to no direct involvement on the part of the Russian government or military." U.S. CYBER CONSEQUENCES UNIT, OVERVIEW BY THE US-CCU OF THE CYBER CAMPAIGN AGAINST GEORGIA IN AUGUST OF 2008 2 (2009), http://registan.net/wp-content/uploads/2009/08/US-CCU-Georgia-Cyber-Campaign-Overview.pdf. *See also* TIKK ET AL., *supra* note 27, at 12 ("[T]here is no conclusive proof of who is behind the DDoS attacks, even though finger pointing at Russia is prevalent by the media. There seems to be a widespread consensus that the attacks appeared coordinated and instructed."); and *see* Oona Hathaway et al., *The Law of Cyber-Attack*, 100 CAL. L. REV. 817, 837–38 (2012) (noting that as Russian forces invaded South Ossetia, private hackers – not the Russian government – orchestrated a cyber attack, and that although the Russian government "stood by" while the attack was "openly" committed, it was not the party that planned and executed the attack).

[29] *See, e.g.*, MINISTRY OF JUSTICE OF GEORGIA, CYBER ATTACKS AGAINST GEORGIA 6 (2011), http://www.dea.gov.ge/uploads/GITI%202011/GITI2011_3.pdf; *see also* David J. Smith, *Russian Cyber Strategy and the War Against Georgia*, NATOSOURCE (Jan. 17, 2014), http://www.atlanticcouncil.org/blogs/natosource/russian-cyber-policy-and-the-war-against-georgia.

- In 2011, the Norwegian military suffered a cyber attack by malicious software one day after joining NATO operations in Libya. No groups took credit for the attack. [30]

In light of the distinctive challenges presented by cyber operations in the battle space—their non-kinetic but potentially serious harms to civilians, along with the potential anonymity of attackers and their use of deception techniques—it may be prudent and perhaps legally advisable for States to develop and agree upon principles for attribution of cyber operations during an armed conflict. It is prudent because misdirected cyber or kinetic responses can cause harmful effects on innocent parties or States and because errors could unnecessarily escalate existing conflicts. It is legally important if mistaken assumptions of state responsibility lead to LOAC violations.

In Part II, this chapter will summarize why attribution of cyber intrusions remains challenging. Part III will review briefly a few aspects of the LOAC that are hardest to apply to cyber operations. I will argue that paying attention to attribution of cyber incidents in armed conflict could lessen some of these doctrinal challenges in applying the LOAC to cyber. In Part IV I will suggest modest enhancements to the LOAC analysis that would incorporate attribution of cyber intrusions for at least some categories of cyber operations in an armed conflict.

---

[30] Healey, *supra* note 20; *Norway army says faced cyber attack after Libya bombing*, ABS–CBS NEWS (May 19, 2011, 11:06 PM), http://news.abs-cbn.com/global-filipino/world/05/19/11/norway-army-says-faced-cyber-attack-after-libya-bombing.

II. ATTRIBUTION AND INTERNATIONAL LAW

Because the internet facilitates anonymous communications and "was not designed with the goal of deterrence in mind,"[31] attribution of cyber intrusions can be challenging, all the more so when the intruders purposefully hide their tracks. The practice of attributing cyber attacks is a relatively recent phenomenon. As cyber intrusions have proliferated in recent years, States have invested in doing attribution well and, as a result, deterring and coercing States and other cyber intruders into complying with societal norms.[32] When attribution is done badly or not at all, States lose credibility and likely effectiveness in dealing with those who would harm the State and its citizens.[33] These risks hold for state-on-state interactions across the spectrum of cyber operations—from espionage to destructive attacks on infrastructure.

The United States takes seriously meeting the challenges of cyber attribution. Former Director of National Intelligence (DNI) Director James Clapper opined a few years ago that "definitive, real-time attribution of cyber attacks—that is, knowing who carried out such attacks and where these perpetrators are located" is the most important challenge faced by the United

---

[31] Clark & Landau, *supra* note 6, at 323.

[32] Thomas Rid & Ben Buchanan, *Attributing Cyber Attacks*, 38 J. STRATEGIC STUD. 4, 4 (2015).

[33] *Id.*

States.[34] In its 2015 *Cyber Strategy*, the U.S. Department of Defense emphasized the importance of attribution:

> Attribution is a fundamental part of an effective cyber deterrence strategy as anonymity enables malicious cyber activity by state and non-state groups. On matters of intelligence, attribution, and warning, DoD and the intelligence community have invested significantly in all source collection, analysis, and dissemination capabilities, all of which reduce the anonymity of state and non-state actor activity in cyberspace. Intelligence and attribution capabilities help to unmask an actor's cyber persona, identify the attack's point of origin, and determine tactics, techniques, and procedures. Attribution enables the Defense Department or other agencies to conduct response and denial operations against an incoming cyberattack.[35]

---

[34] James R. Clapper, Dir. of Nat'l Intelligence, Worldwide Threat Assessment to the Senate Select Committee on Intelligence (Jan. 21, 2012), https://www.dni.gov/files/documents/Newsroom/Testimonies/20120131_wwta_as_delivered_remarks.pdf.

[35] DEP'T OF DEF., THE DEPARTMENT OF DEFENSE CYBER STRATEGY 11–12 (2015).

Despite the U.S. rhetoric, there is little law to guide attribution. Attribution is hard because States do not usually carry out cyber attacks transparently.[36] Instead, they use technical tools to hide their responsibility and rely on non-state proxies to carry out cyber activities for them.[37] Indeed, the United States has only rarely officially attributed a malicious cyber operation to another State—China following widespread corporate espionage in 2014,[38] North Korea following the

[36] *See* Matthew J. Sklerov, *Solving the Dilemma of State Responses to Cyberattacks: A Justification for the Use of Active Defenses Against States Who Neglect Their Duty to Prevent*, 201 MIL. L. REV. 1, 8 (2009).

[37] *See* JEFFREY CARR, INSIDE CYBER WARFARE 29, 139–40 (O'Reilly 2010).

[38] *See* Robert Chesney, *DOJ's Summary of the Charges in the Chinese Economic Cyberespionage Case*, LAWFARE (May 19, 2014, 10:55 PM), https://www.lawfareblog.com/dojs-summary-charges-chinese-economic-cyberespionage-case; *See also* Susan E. Rice, Nat'l Sec. Advisor, Remarks on the U.S.-China Relationship at George Washington University, (Sept. 21, 2015), https://obamawhitehouse.archives.gov/the-press-office/2015/09/21/national-security-advisor-susan-e-rices-prepared-remarks-us-china (describing a U.S.–China agreement prohibiting cyber commercial espionage for commercial gain a little more than a year after the indictments).

Sony hack in 2014,[39] and Russia following the DNC hack in 2016.[40] Notably, none of these incidents and attributions occurred during armed conflict.

When cyber operations are launched alongside or to facilitate kinetic strikes in an armed conflict, attribution will in all likelihood be assumed. Indeed, Jody Prescott, a senior fellow at the United States Military Academy (USMA) noted that "[w]ith cyber operations conceivably moving at near light speed, commanders in cyber warfare will likely need to rely extensively upon autonomous decision-making processes (ADPs) to be effective."[41] For example, during a hypothetical armed conflict between States A and B, it may be reasonable to assume that attacks

---

[39] *See* Herb Lin, *Learning from the Attack Against Sony*, LAWFARE (Jan. 23, 2015, 10:38 PM), https://www.lawfareblog.com/learning-attack-against-sony#.

[40] *See* Press Release, Dep't of Homeland Sec., Joint Statement from the Department of Homeland Security and Office of the Director of National Intelligence on Election Security (Oct. 7, 2016), https://www.dhs.gov/news/2016/10/07/joint-statement-department-homeland-security-and-office-director-national; Ellen Nakashima, *U.S. government officially accuses Russia of hacking campaign to interfere with elections*, WASH. POST (Oct. 7, 2016), https://www.washingtonpost.com/world/national-security/us-government-officially-accuses-russia-of-hacking-campaign-to-influence-elections/2016/10/07/4e0b9654-8cbf-11e6-875e-2c1bfe943b66_story.html?utm_term=.655d3d88e3a6.

[41] JODY M. PRESCOTT, AUTONOMOUS DECISION-MAKING PROCESSES AND THE RESPONSIBLE CYBER COMMANDER (2013), https://ccdcoe.org/cycon/2013/proceedings/d2r1s6_prescott.pdf.

on the command and control systems, classified communications networks, or weapons guidance systems are the result of actions taken by the enemy State.[42] However, even during a conventional state-on-state armed conflict it is not necessarily the case that all cyber intrusions suffered by State A were caused by State B, even those apparently originating in State B. Nor will the origins of all cyber activity be known, certainly not in the real time dynamics of an armed conflict. Private actors could be responsible for any of the cyber operations, as could another State or proxies of another State or a terrorist organization. Machine attribution could trace malware to computers or systems in State C, which may or may not be controlled by neutral State C. Or malware could be coming from sources in several States, and State responsibility is not immediately clear.

Attribution has been characterized as more art than science.[43] In fact, significant strides have been made in attribution of cyber events in the last decade, making the task "more nuanced, more common, and more political" than has typically been acknowledged.[44] Attribution is measured in degrees of certainty, and requires input from a range of actors. In the United States,

---

[42] *See* Hathaway et al., *supra* note 32, at 838.

[43] Rid & Buchanan, *supra* note 36, at 4; Clark & Landau, *supra* note 9, at 350 ("[Attribution] is not actually a technical issue at all, but a policy concern with multiple solutions depending on the type of technical issue . . . . to be solved . . . . [S]olutions . . . . lie outside the technical realm, and are instead in the space of law, regulation, multi-national negotiation, and economics.")

[44] Rid & Buchanan, *supra* note 32, at 6.

much of the evidence to support attribution is off-line and involves traditional interviews and examination of equipment.[45] The attribution efforts may themselves be thwarted or slowed down by adversaries, often using cyber tools to spoof their location or identity.[46]

Although considerable advances in detection technology enable States to more reliably identify the machines that have disseminated cyber attacks than in the past,[47] identifying the persons, organizations, or States that are responsible for the cyber attack remains challenging.[48] Even finding and seizing the offending computer is unlikely to reveal the sponsors of an attack.[49] The problems are in part due to technical means of deception and anonymity, but are also due to the vagaries of the process of fixing responsibility for cyber attacks and the malleability and open-endedness of the little attribution law that currently exists in the *jus ad bellum*.[50]

---

[45] *See* Herbert Lin, *Attribution of Malicious Cyber Incidents: From Soup to Nuts*, 70 J. INT'L AFFAIRS 75, 92 (2017); Carlin, *supra* note 6, at 414.

[46] *See* Carlin *supra* note 6, at 409; *see also* Lin, *supra* note 49, at 82.

[47] *See, e.g.*, Carlin, *supra* note 6, at 416; Lin, *supra* note 49, at 82–83.

[48] Carlin, *supra* note 9, at 409; Lin, *supra* note 49, at 84.

[49] *See* GUITTON, *supra* note 6, at 47.

[50] *See* William Banks, *State Responsibility and Attribution of Cyber Intrusions After Tallinn 2.0*, 95 TEX. L. REV. 1487, 1494–97 (2017).

The customary law of state responsibility and attribution is largely drawn from the long-term work of the International Law Commission (ILC) and its Rules on State Responsibility. The ILC rules were commended to the member states by the UN General Assembly in 2012 and have become the authoritative guidepost for public international cyber law.[51] The starting point is that "a State bears international responsibility for a cyber-related act that is attributable to the State and that constitutes a breach of an international legal obligation."[52] Thus, attribution is *required* before a State may be found legally responsible for a cyber intrusion. Once attributed, States are legally responsible for an internationally wrongful act. Establishing factual attribution remains challenging in many instances, as does setting legal requirements for arriving at attribution.

The 2017 *Tallinn Manual 2.0 on the International Law of Cyber Operations*[53] summarizes the extant customary international law on state responsibility and attribution. In essence, states are responsible for cyber-related acts of their own officials, agents, contractors, non-state actors, and other states to the extent they actually control the operations.[54] States do not escape legal responsibility for internationally wrongful acts by perpetrating them through proxies.[55] Below the use of force threshold, States are responsible for a "cyber-related act . . . that constitutes a breach

---

[51] TALLINN MANUAL 2.0, *supra* note 5, at 79 n.112.

[52] *Id*. R. 14, at 84.

[53] TALLINN MANUAL 2.0, *supra* note 5.

[54] *See generally*, *id*. at R. 15, at 87–92.

[55] *Id*. R. 17, at 94–95.

of an international legal obligation."[56] The act may violate a treaty, customary international law, or other "general principles of law."[57]

Outside an armed conflict, international law forbids cyber intrusions that violate the prohibition on intervention.[58] Based on the international law principle of sovereignty, the principle forbids coercive intervention by cyber means.[59] *Tallinn 2.0* reports that state-on-state cyber intrusions that are not coercive but are "detrimental, objectionable, or otherwise unfriendly"[60] are not international law violations. As confirmed by the International Court of Justice (ICJ) in the *Nicaragua* judgment, "the element of coercion . . . forms the very essence of [] prohibited intervention."[61] What constitutes coercion? According to *Tallinn 2.0*, "coercion is not limited to

---

[56] *Id*. R. 14, at 84; *see, e.g.*, Phosphates in Morocco (It. v. Fr.), Preliminary Objections, 1938 P.C.I.J. (ser. A/B), No. 74, at 28 (June 14) ("This act being attributable to the State and described as contrary to the treaty right of another State, international responsibility would be established immediately as between the two States."); United States Diplomatic and Consular Staff in Tehran (U.S. v. Iran), Judgment, 1980 I.C.J. 73, ¶¶ 29–30 (May 24).

[57] TALLINN MANUAL 2.0, *supra* note 5, R. 14(2), at 84.

[58] *Id*. R. 66(1), at 312.

[59] *Id*. at 312–13.

[60] *Id*. R. 15(7), at 85.

[61] Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. 14, ¶ 205 (June 27).

physical force, but rather refers to an affirmative act designed to deprive another State of its freedom of choice . . . to force that State to act in an involuntary manner or involuntarily refrain from acting in a particular way."[62] The General Counsel of the DoD indicated in a January 2017 memorandum to the Combatant Commands and other senior military and civilian lawyers in the Pentagon that coercion is a prerequisite for unlawful intervention and that even attributed non-coercive cyber intrusions do not violate the non-intervention principle and are "largely unregulated by international law at this time."[63]

To date, state practice on intervention is based on kinetic examples; the analogy to cyber may not be persuasive. The leading case is *Nicaragua*, where the ICJ found that United States support of the Nicaraguan Contras in 1983 and 1984 through financial support, training, supply of weapons, intelligence, and logistical support breached the principle of non-intervention and constituted a threat to use force, thus coercing the government of Nicaragua.[64] In any case, physical damage or injury is not necessary for a cyber intrusion to be an internationally wrongful

---

[62] TALLINN MANUAL 2.0, *supra* note 5, R. 66(18), at 317.

[63] Dep't of Def., Memorandum for Commanders in the Combatant Commands, International Law Framework for Employing Cyber Capabilities in Military Operations (2017) (on file with author). The Memorandum acknowledges that the "exact contours that might violate the principle of non-intervention are not clear, and will continue to develop with state practice over time." *Id.*

[64] *See* Nicar. v. U.S., 1986 I.C.J. 14, ¶ 202, 205, 251.

act.[65] For example, a State that launches a targeted and highly disruptive distributed denial of service (DDoS) operation against another State may have acted coercively and engaged in a prohibited intervention if the operation is intended to cause the victim State to change its conduct, such as in relation to a third State.[66]

The International Group of Experts (IGE) that provided the analysis in *Tallinn 2.0* acknowledged the "uncertainty as to the attribution of cyber operations" and agreed "that as a general matter, States must act as reasonable States would in the same or similar circumstances when considering responses to them."[67] The IGE elaborated:

> Reasonableness is always context dependent. It depends on such factors as, *inter alia*, the reliability, quantum, directness, nature (e.g., technical data, human intelligence), and specificity of the relevant available information when considered in light of the attendant circumstances and the importance of the right involved. These factors must be considered together. Importantly in the cyber context, deficiencies in technical intelligence may be compensated by, for example, the existence of highly reliable human intelligence.[68]

---

[65] *See* TALLINN MANUAL 2.0, *supra* note 5, at R. 14(8), at 86, R. 66(16)–(17), at 317.

[66] *Id*. R. 66(19), at 318.

[67] *Id.* at 81.

[68] *Id*. at 81–82.

The IGE opined that "as a general matter the graver the underlying breach . . . , the greater the confidence ought to be in the evidence relied upon by a State considering a response[69] . . . because the robustness of permissible self-help responses . . . grows commensurately with the seriousness of the breach."[70] Notwithstanding the best work of the IGE, because attribution judgments that determine state responsibility remain to some extent uncertain, and because there is no robust international or domestic law understanding on how much evidence suffices for attribution of state responsibility, the attribution bar is set very low by international law.

In addition, the legal standards for attribution are malleable to the extent that the evidence of attribution is not required to be shared publicly[71] and normally is not. In addition, the evidence

---

[69] In support of its position, the IGE cited: Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. 161, ¶ 33 (Nov. 6) (separate opinion of Judge Higgins); Corfu Channel (U.K. v. Alb.), Judgment, 1949 I.C.J. 4, ¶ 17 (Apr. 9); Application of Convention on Prevention and Punishment of Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 108, ¶¶ 209–10 (Feb. 26); and Application of Convention on Prevention and Punishment of Crime of Genocide (Croat. v. Serb.), 2015 I.C.J. General List No. 118, ¶ 178 (Feb. 3).

[70] TALLINN MANUAL 2.0, *supra* note 5, at 82.

[71] *See* Brian J. Egan, Remarks on International Law and Stability in Cyberspace at Berkeley Law School (Nov. 10, 2016), https://www.law.berkeley.edu/wp-

leading to attribution is often based on intelligence collection rather than testable machine-derived data. As a result, the legal criteria for attribution decisions from the *jus ad bellum* distill to a subjective reasonableness.[72] For example, it may be difficult to tell whether cyber intrusions were ordered by a State, tolerated by a State that knew about them, or carried out by proxies for the State that followed their own loosely governed agenda.

Ultimately, the decision to assign responsibility for a cyber attack to a State is a political decision, based on a combination of digital forensics and intelligence intercepts rather than a set of established legal criteria.[73] Actual (beyond technical, machine) attribution in a State sponsored attack rarely takes place quickly, except when strategic or political considerations incentivize rapid attribution.[74] Indeed, the more time investigators have to collect evidence for attribution, the more reliable the attribution judgment is likely to be. Strategic reasons may also give States cause for delaying attribution or never making it public.[75] States can normally make an initial guess about the perpetrators of a cyber intrusion in the national security realm, but obtaining conclusive evidence of sponsorship is difficult.

---

content/uploads/2016/12/egan-talk-transcript-111016.pdf; TALLINN MANUAL 2.0, *supra* note 5, at 83.

[72] *See* Banks, *supra* note 50, at 1505–06.

[73] *Id*. at 1510–11.

[74] GUITTON, *supra* note 6, at 138.

[75] *Id*. at 154, 160, 185.

In addition, the time it takes to produce a high confidence attribution judgment can impact the lawful responses to cyber operations below the armed conflict threshold. For example, mistaken attribution can lead to an unlawful response even if the State made a reasonable attribution judgment and implemented countermeasures.[76] The IGE concluded that "as a general matter the graver the underlying breach . . . the greater the confidence ought to be in the evidence relied upon by a State considering a response."[77] The more severe the injury, the less certain attribution needs to be, and the stronger the planned response, the greater the confidence in attribution. When intrusions are not severe, the State can accumulate more data for attribution.[78] Judgments are heavily influenced by what is at stake politically. Although attribution is necessarily probabilistic, the process serves its purpose if it convinces the responsible State (and victim State

[76] TALLINN MANUAL 2.0, *supra* note 5, at 82–83.

[77] *Id*. at 82; *See also* Application of Convention on Prevention and Punishment of Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), 2007 I.C.J. 108, ¶¶ 209–10 (Feb. 26) (discussing the implicitly proportionate connection between the degree of one country's offense and another country's response); Corfu Channel (U.K. v. Alb.), 1949 I.C.J. 4, 17 ¶ 39 (Apr. 9) ("A charge of such exceptional gravity against a State would require a degree of certainty").

[78] *See* TALLINN MANUAL 2.0, *supra* note 5, at 82.

citizens) that a response to the cyber intrusion was called for.[79] As suggested in Part IV, such a sliding scale approach to attribution may be portable to the *in bello* world of armed conflicts.

The architecture of the Internet has changed little over the last two decades. Burdened by a largely insecure structure, the art and science of attribution are evolving, but only gradually. The good news is that better intrusion detection systems now flag breaches in real or nearly real-time.[80] At the same time, improvements in adaptive, resilient networks help deter offensive intrusions.[81] The bad news is that the intruders are learning, too, and encryption and other deception advances greatly complicate forensic identification.[82] Meanwhile, States and non-state actors often act in the cyber realm with relative impunity when no or only negligible sanctions follow from being outed.[83] Indeed, a 2017 Council on Foreign Relations Memorandum opined that even a major cyber attack on the U.S. electric power grid could be carried out on the likely mistaken assumption that the attack could not be attributed. Even an unfounded expectation that another State could attack the

[79] *See* GUITTON, *supra* note 6, at 66.

[80] *E.g.*, Lin, *supra* note 45, at 108; GUITTON, *supra* note 6, at 137–46.

[81] *E.g.*, Lin, *supra* note 45, at 106–07

[82] *See* Rid & Buchanan, *supra* note 32, at 31–32.

[83] *See* Banks, *supra* note 50, at 1511–12.

United States anonymously and with impunity could lead to devastating consequences.[84] Under such circumstances, a "lax de-facto norm of negligible consequences"[85] may emerge, even during an armed conflict. The dangers of complacency—increasing harms from cyber intrusions following lax attribution and only modest enforcement of norms—enhance the value of undergirding the LOAC with *in bello* attribution components. Only if States invest in accountable attribution mechanisms will any new international law on attribution have practical value.[86] A dilemma for the United States is that we benefit from the absence of express norms because we have the most offensive tools. But our society is also the most vulnerable to cyber intrusions.[87]

---

[84] *See* Robert K. Knake, *A Cyberattack on the U.S. Power Grid: Contingency Planning Memorandum No. 31*, COUNCIL ON FOREIGN RELATIONS 3, 4 (April 3, 2017), https://www.cfr.org/sites/default/files/pdf/2017/03/ContingencyPlanningMemo31_Knake.pdf.

[85] Rid & Buchanan, *supra* note 32, at 33.

[86] *E.g.*, Egan, *supra* note 71, at 11–12; Rid & Buchanan, *supra* note 32, at 31–33.

[87] *See* Banks, *supra* note 50, at 1511–12. The 2015 DoD Law of War Manual claims that "[a]s a matter of U.S. policy, the United States has sought to work internationally to clarify how existing international law and norms, including law of war principles, apply to cyber operations." DOD LAW OF WAR MANUAL, *supra* note 2, § 16.1, at 985. Others have suggested that "lingering ambiguity with respect to what the U.S. regards as lawful and unlawful actions in the cyber domain [may] actually serve U.S. interests." SCOLANS REPORT, *supra* note 17, at 61.

Ironically, we also have the best attribution capabilities and can therefore sleuth out and identify the states and non-state actors engaged in unlawful cyber operations, even during armed conflict.

It remains to be seen how well the *jus ad bellum* law on state responsibility and attribution, limited as it is, may be applied in armed conflict. It is certainly true that the LOAC has an interconnected patchwork of principles and doctrinal rules that serve to protect civilians in armed conflict from the impacts of cyber intrusions. In theory, because attribution occurs *before* response targeting analysis and its application of the LOAC principles of distinction, proportionality, and precautions, something like the *ad bellum* law on state responsibility and attribution may improve the application of LOAC when cyber is part of armed conflict.

III. Challenges in Applying the LOAC to Cyber

Consider a hypothetical illustration:

> State A is engaged in an armed conflict with State B. The conflict is primarily kinetic, although both States have utilized cyber means in attempts to degrade the command and control of each other's military. Meanwhile, various additional cyber operations have impacted civilian infrastructure in State A, including civilian networks and operations that directly or indirectly support the ongoing military campaign. The victimized networks include civilian contractors that supply and provide logistical support to the State A military, and the infrastructure the military relies on for its operations, including ports, railroads, and electricity. No State or

> anyone else has claimed responsibility for the cyber intrusions in State A. Preliminary machine attribution indicates that the attacks have originated primarily inside State B, although the dissemination of malware has exploited computers at various locations around the world.

Are the cyber intrusions in State A "attacks"? If so, although the offending machines may be targeted by State A as military objectives, may or must State A attribute state responsibility for the attacks to State B before targeting the facilities or entities responsible for the incoming cyber attacks? If the intrusions do not qualify as attacks, does their occurrence during an armed conflict permit a cyber or kinetic military response? Is attribution required for those operations?

A. Which cyber intrusions are subject to the LOAC? Must they be attributed?

Most cyber intrusions that cross sovereign boundaries do not violate international law. Outside an armed conflict, for those operations where the impact does not constitute an "attack" according to the LOAC, only cyber intrusions that constitute an internationally wrongful act—coercive cyber intervention—are clearly proscribed by international law. Thus, an isolated cyber intrusion that is neither an attack nor an internationally wrongful act may have no international legal consequences.

According to *Tallinn 2.0*, "a situation involving hostilities, including those conducted using cyber means" is an armed conflict.[88] In addition, intrusions that cause cyber harm but not physical damage during an armed conflict are subject to the LOAC. As *Tallinn 2.0* explains, the 2007 cyber operations targeting Estonia did not trigger the LOAC because the "situation did not rise to the level of an armed conflict,"[89] while the cyber operations that occurred between Georgia and Russia in 2008 and now in the ongoing conflict between Ukraine and Russia are subject to the LOAC because those conflicts involved hostilities rising to the level of armed conflict.[90]

Are all cyber intrusions during armed conflict subject to the LOAC? According to one view held among the IGE, the LOAC "governs any cyber activity conducted by a party to the armed conflict against its opponent."[91] (This view presumes attribution or dismisses its importance.) Another group maintained that the LOAC applies only when "the cyber activity [is] undertaken in furtherance of the hostilities."[92] (Still no mention of attribution.) All members of the IGE agreed "that there must be a nexus between the cyber activity in question and the conflict for the law of armed conflict to apply to that activity."[93]

---

[88] TALLINN MANUAL 2.0, *supra* note 5, R. 80(2), at 375.

[89] *Id*. R. 80(3), at 376.

[90] *See id*.

[91] *Id*. R. 80(6).

[92] *Id*.

[93] *Id*. R. 80(5).

The IGE acknowledged "that the application of the law of armed conflict to cyber operations can prove problematic [because] it is often difficult to identify the existence of a cyber operation, its originator, its intended object of attack, or its precise effects."[94] However, the experts agreed that questions of fact regarding the existence, purpose, or origins of a cyber operation "do not prejudice the application of the law of armed conflict."[95] Because of the vagaries of applying the LOAC to cyber activities, the IGE agreed that the Martens Clause[96] would provide general law of nations protections for cyber activities conducted in the course of an armed conflict.[97]

---

[94] *Id*. R. 80(10), at 377.

[95] *Id*.

[96] Convention No. IV Respecting the Laws and Customs of War on Land, preamble, Oct. 18, 1907, 36 Stat. 2227; Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, art. 63, Aug. 12 1949, 6 U.S.T. 3114, 75 U.N.T.S. 31; Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea, art. 62, Aug. 12, 1949, 6 U.S.T. 3217, 75 U.N.T.S. 85; Convention Relative to the Treatment of Prisoners of War, art. 142, Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135: Convention Relative to the Protection of Civilian Persons in Time of War, art. 158, Aug. 12, 1949, 6 U.S.T. 2516, 75 U.N.T.S. 287; Additional Protocol I, *supra* note 17, at art. 1(2).

[97] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 80(12), at 378.

Although the definition of "attack" in the LOAC is clearly focused on kinetics, the colloquial understanding of what constitutes a "cyber attack" has a broad, almost all-encompassing meaning, ranging from destructive attacks to exfiltration to denial of service.[98] Additional Protocol I, the *Tallinn Manual*, and the *DoD Law of War Manual* agree that "attack" is the pertinent triggering concept for invoking LOAC principles.[99] The *Tallinn Manual* defines a cyber attack as "a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects."[100] The focus of analysis is on the effects or consequences of a cyber operation, and the harm or damage to objects must be more than *de minimus*.[101] The *Tallinn Manual* definition clearly is not limited to kinetic force. For example, interference with the functionality of a computer or system may qualify as an attack.[102]

---

[98] *See, e.g.*, Lubell, *supra* note 4, at 255–56.

[99] Additional Protocol I, *supra* note 13, at art. 1; TALLINN MANUAL 2.0, *supra* note 5, R. 92(2), at 415; DOD LAW OF WAR MANUAL, *supra* note 2, § 16.5.1, at 994.

[100] TALLINN MANUAL 2.0, *supra* note 5, R. 92, at 415.

[101] *See* Michael N. Schmitt, *Cyber Operations and the Jus in Bello: Key Issues*, 87 INT'L L. STUD. 89, 94 (2011).

[102] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 92(6), at 417. A majority of the IGE concluded that interference with functionality amounts to damage if restoring the system requires replacing components. *Id*. R. 92(11), at 417.

At the same time, some categories of cyber harm are not cyber attacks and may not trigger LOAC principles standing alone. As Michael Schmitt has recognized, the effects of cyber operations that cause "inconvenience, disruption, disorder or irritation. . . might . . . be severe, as in significant interference with the economy, transportation system or other critical infrastructure."[103] Yet such cyber operations do not by themselves initiate an armed conflict even if the effects on civilians are significant.[104] To the extent the LOAC is not in force, the important principles such as distinction and proportionality do not apply to protect civilians.

Tallinn Manual 2.0 confirms that once an armed conflict exists, cyber operations that cause cyber harm are subject to the LOAC.[105] Cyber harm might become the benchmark for invoking the rules designed to shield civilian populations from harm. Part III will consider whether international law could insist that state responsibility for cyber operations that cause cyber harm during an armed conflict be attributed in order to add accountability for the harms to civilians during conflict and to improve responsive targeting by cyber means.

B. Cyber and the LOAC Principle of Distinction

[103] Schmitt, *supra* note 101, at 103.

[104] *See Id*. at 104.

[105] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 80, at 375.

Codified in API, the bedrock principle of distinction requires that "Parties to the conflict . . . at all times distinguish between the civilian population and combatants and between civilian objects and military objectives."[106] Accordingly, the LOAC prohibits cyber attacks in an armed conflict that are uncontrollable, unpredictable, or that otherwise do not discriminate between civilian and military objectives.[107] Article 52(2) of API enforces the principle of distinction by stating that

> Attacks shall be limited strictly to military objectives . . . [which] are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage.[108]

What objects in the cyber domain are obviously lawful targets under the Protocol? Weapon guidance systems, classified military networks, the factory that makes the software for the network

---

[106] Additional Protocol I, *supra* note 13, at art. 48.

[107] *See id*. at art. 52(2) (requiring that targets serve a military purpose and their attainment produces a definite military advantage); *id*. at art. 51(4) (forbidding weapons that cannot be limited to a military objective).

[108] *See id*. at art. 52(2).

or guidance system for the weapon, for openers. Likewise, some cyber attacks would be clearly unlawful, including a cyber attack on a hospital, museum, or place of worship.[109]

Beyond the easy cases, in the cyber domain the principle of distinction may be seriously compromised. Machine attribution is often straightforward, so that a computer may be targeted, but establishing State responsibility is often challenging. An easy illustration involves spoofing. The computer that "shot" at the victim State was taken over, so the computer is a military objective, but its owner is not. Attribution would attempt to determine who is responsible for the spoofed cyber activity before a targeting analysis is undertaken.

The distinction principle is enforced by the military objective definition quoted previously. In many instances, the "nature" of the object cannot be determined without knowing who owns or controls it. Civilian telecommunications infrastructure is not a lawful target, while military communications infrastructure that relies on the same internet backbone may be targeted in an armed conflict. Following the same military objective criteria, "purpose" and "use" determinations in the cyber domain require knowing about ownership, or at least control, and, thus, attribution. The latter two components of the military object definition are confounded in the cyber realm by the fact that just about every cyber installation could be considered a dual-use object[110] and thus a

[109] *Id*. at 85(4)(d).

[110] Dual-use objects can have military and civilian purposes. Jensen *supra,* note 8, at 1535, 1544 n.76; Hathaway et al., *supra* note 28, at 852–53; Droege, *supra* note 8, at 562–63.

military objective.[111] While objects in the physical world are theoretically capable of being dual-use, most are not in fact.[112] Because the military uses the same cyber infrastructure that civilians use for their purposes, that infrastructure may in general be lawfully attacked during an armed conflict.[113]

As technologically advanced States attain greater sophistication in the use of the cyber domain for strategic purposes, the cyber infrastructure will present increasingly significant targets in future armed conflicts. Based on the conventional LOAC conception of what counts as a military object, all civilian cyber infrastructure that transmits military communications and data are dual-use and could be seen as lawful military objectives.[114] As Robin Geiβ and Henning Lahmann argued in 2013, "there simply is no difference between a military and a civilian computer; any computer and basically any part of the larger cyber infrastructure can be used to serve the military and the civilian constituency either interchangeably or simultaneously."[115] According to Article 52(2) of API, a wide range of civilian cyber assets would qualify as legitimate military objectives because their neutralization or destruction would offer a definite military advantage. Of course,

---

[111] *See* Robin Geiβ & Henning Lahmann, *Cyber Warfare: Applying the Principle of Distinction in an Interconnected Space*, 45 ISR. L. REV. 381, 383 (2012).

[112] *Id*. at 389.

[113] *Id*. at 383.

[114] *See* Jensen, *supra* note 8, at 1542.

[115] Geiβ & Lahmann, *supra* note 111, at 389.

these permissive targeting principles do not apply unless the civilian infrastructure is actually used by the military.

Despite widespread criticism of the sweeping potential for the military objective definition to reach virtually the entire civilian cyber infrastructure,[116] the recent trends, particularly in the United States, are to expand the definition to include war-sustaining objects.[117] The United States includes war-sustaining objects as lawful military objectives, defined in the Handbook on the Law of Naval Operations as "[e]conomic objects of the enemy that indirectly but effectively support and sustain the enemy's war-fighting capability. . . ."[118] War-supporting or war-sustaining objects would include a factory that makes a computer guidance system for a weapon or the software that

---

[116] *See* Geiβ & Lahmann, *supra* note 111, at 390; Lubell, *supra* note 4, at 272; *see e.g.* TALLINN MANUAL 2.0, *supra* note 5, R. 101(6), at 446 ("In theory, strict application of the definition of military objective could lead to the conclusion that the entire Internet can become a military objective if used for military purposes. . . . [i]n this regard, particular attention must be paid to the requirement to conduct operations in a manner designed to minimise harm to the civilian population.").

[117] *E.g.*, Geiβ & Lahmann, *supra* note 111, at 390; DOD LAW OF WAR MANUAL, *supra* note 2, § 5.7.8, at 213.

[118] DEP'T OF THE NAVY & DEP'T OF HOMELAND SEC., THE COMMANDER'S HANDBOOK ON THE LAW OF NAVAL OPERATIONS § 8.2.5 (2007), http://www.jag.navy.mil/documents/NWP_1-14M_Commanders_Handbook.pdf (hereinafter "Commander's Handbook").

runs on a classified network. Under the U.S. approach, it would also be lawful to "launch cyber attacks against the enemy State's oil export industry if the war effort depends on revenue from oil sales."[119]

The U.S. interpretation is contrary to the views of the IGE in *Tallinn Manual 2.0*. The *Tallinn 2.0* Rule 100 concedes that "[c]yber infrastructure may qualify as a military objective."[120] However, on the issue of war-sustaining objects, a majority of the IGE rejected the extension of the "military objective" treaty language "on the ground that the connection between war-sustaining activities and military action is too remote."[121] The majority would limit permissible targets to those objects that are war-fighting or war-supporting. An example of the latter is a factory producing hardware or software for use by the military.[122]

Following the U.S. approach in the DoD *Law of War Manual* and API, the *Tallinn 2.0* IGE confirmed that dual-use objects and facilities are military objectives "without qualification."[123] However, the IGE also carefully parsed several examples in the commentary on the dual-use and related Rules, acknowledging several categories of hard cases and affirming the duty to

---

[119] TALLINN MANUAL 2.0, *supra* note 5, R. 101(1), at 441.

[120] *Id.* at R. 100, at 436.

[121] *Id.* R. 100(19), at 441.

[122] *Id*.

[123] *Id*. R. 101(1), at 445.

expeditiously resolve any doubts about the legal status of cyber infrastructure as a military or civilian object.[124] In any event, the API standard and the U.S. stance generally call into question LOAC protections for civilian cyber infrastructure. Although cogent proposals have been made to limit cyber operations against dual-use infrastructure to the least disruptive action and to preclude war-sustaining objects from being subject to cyber attack,[125] there is no indication that the LOAC or state practice has adopted these reforms.

Arguably, Article 56(1) of API and its exemption from attack due to severe humanitarian consequences where objects otherwise qualify as military objectives may protect civilian cyber infrastructure in some settings.[126] If, for example, components of the dual-use cyber infrastructure are impacted by military cyber operations (as they almost surely would be), the Protocol could be read to limit such actions where the consequences for the functionality of civilian cyber traffic are significant.[127] Unfortunately, Article 56(1) justifies an exemption only when there may be "severe losses among the civilian population"[128] and is unlikely to protect against the loss of functionality

---

[124] *See id*. R. 101–102, at 445–51.

[125] *See* Pascucci, *supra* note 4, at 456; Int'l Law Ass'n Study Grp. on the Conduct of Hostilities in the 21st Century, *The Conduct of Hostilities and International Humanitarian Law: Challenges of 21st Century Warfare*, 93 INT'L L. STUD. 322, 335–40 (2017).

[126] *See* Geiβ & Lahmann, *supra* note 111, at 391.

[127] *Id*.

[128] *Id*. at 392.

of the internet short of the destruction of important infrastructure components.[129] Alternatively, steps could be taken to segregate military and civilian networks, following the Article 58(a) obligation to take passive precautions in an armed conflict.[130] However, the obligation applies only "to the maximum extent feasible," and in any case, the passive precautions duty does not override lawful dual-use targeting.[131] In addition, the precautions requirement in Article 58(c), requiring States to take other necessary precautions to protect the civilian population and civilian objects from the dangers resulting from military operations, has not but could be construed to require ensuring continuing cyber functionality of, for example, the electric grid during an armed conflict.[132] Consistent with the Protocol, the DoD *Law of War Manual* notes that the "obligation to take feasible precautions may be of greater relevance in cyber operations . . . because this obligation applies to a broader set of activities than those to which other law of war rules apply."[133]

Overall, the dual-use phenomenon makes it unlikely that States will take steps based on the principle of distinction to limit military cyber operations that impact civilian systems. Even in the face of the U.S. position on war-sustaining objects, however, an attribution process embedded in analyzing the incoming cyber operations, if implemented before the response targeting process,

---

[129] *Id*.

[130] *See* Additional Protocol I, *supra* note 13, at art. 58(c).

[131] Geiβ & Lahmann, *supra* note 111, at 393.

[132] *Id*. at 395.

[133] DOD LAW OF WAR MANUAL, *supra* note 2, §16.5.3, at 997.

could ameliorate the risk of targeting mistakes, highlight responsibility for the provoking attack, and possibly provide alternative targeting options.

In theory, the principle of proportionality affords greater flexibility than the principle of distinction toward the same objective. Article 51(5)(b) of API sets up a balancing, where incidental loss of civilian life, injury to civilians, damage to civilian objects or a combination thereof is prohibited if it is excessive in relation to the concrete and direct military advantage anticipated.[134] The principle clearly applies in the cyber domain.[135] Yet the relevant criteria for consideration in a proportionality assessment are "loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof."[136] A loss of functionality is apparently not part of a proportionality calculus.[137] Thus, the principle would be an important factor for destructive cyber operations, but not for those that cause cyber harm but do not destroy any objects. Although creative arguments have been made to extend proportionality analysis to incorporate cyber harms,[138] the law has not embraced such a change so far.

---

[134] Additional Protocol I, *supra* note 13, at art. 51(5)(b).

[135] Eric Talbot Jensen, *Unexpected Consequences From Knock-on Effects: A Different Standard for Computer Network Operations?*, 18 AM. U. INT'L L. REV. 1145, 1158–61 (2003).

[136] Additional Protocol I, *supra* note 13, at art. 51(5)(b).

[137] *See* Geiβ & Lahmann, *supra* note 111, at 397.

[138] *Id*. at 396–98.

C. What to do about Data

May data be lawfully treated as a civilian "object" protected from attack for LOAC purposes? The answer remains unclear. The *Tallinn Manual* 2.0 IGE agreed by a majority that data should not be considered an "object . . . at least in the current state of the law" because "data is intangible and therefore neither falls within the 'ordinary meaning' of the term object, nor comports with the explanation of it offered in the ICRC Additional Protocols 1987 Commentary."[139] The ICRC Commentary indicates that the term "object" refers to something "visible and tangible."[140] Of course, the commentators' view was only an understanding, not part of the language, and it was observed in 1987 before the growing significance of the cyber domain.

The implication of the IGE understanding is that a cyber operation aimed at corrupting, manipulating, or destroying data resident on a computer or cyber system does not constitute an attack and is not subject to the distinction principle so long as the operation does not affect the functionality of the computer or system. An operation that does affect functionality of computers or cyber systems was thought by the IGE as "sometimes" qualifying as an attack.[141] Exceptions

[139] TALLINN MANUAL 2.0, *supra* note 5, R. 100(6), at 437.

[140] Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 §2008 (Yves Sandoz et al. eds., 1987).

[141] TALLINN MANUAL 2.0, *supra* note 5, R. 100(6), at 437.

are recognized where the attack on data leads to injuries or physical damage.[142] A minority of the IGE argued that at least certain data (such as social security data, tax records, and bank accounts) should be within the scope of the targeting rule and protected by distinction so that such critical data are not lost and the civilian population thereby victimized. For the minority, the severity of the consequences of a cyber operation matter more than the nature of the harm.[143] None of the analyses of how to treat data in the LOAC have to date considered attributing state responsibility for incoming cyber activity that harms data on cyber systems.

Certainly, the overall approach to the LOAC taken by the United States is to focus on practical impacts of military operations when striving to protect civilians. In the cyber world, the focus should properly be on harm to the cyber system, including resident data.[144] At the same time, it would be similarly helpful to reframe the criteria for military objective in a cyber setting to focus on whether the data offers a definite military advantage or demonstrable military purpose.[145] Data not meeting the test for military objective would be civilian objects and thus protected in applying LOAC principles.

D. Neutrality

---

[142] *See id*.

[143] *See id*. R. 100(7).

[144] *See* Lubell, *supra* note 4, at 268.

[145] *See* Pascucci, *supra* note 4, at 455.

During international armed conflicts, the law of neutrality applies to cyber operations and to cyber infrastructure located within or owned by a neutral State.[146] The law of neutrality protects neutral States and their citizens from the armed conflict while it protects the States in conflict against actions taken by the neutral State for the benefit of one of the States in conflict.

In the cyber domain, the territorial boundaries that can signify neutrality are not easily applied in part because internet pathways host traffic that may be routed through neutral States' cyber infrastructure regardless of its origins or destinations. As such the core principle of the law of neutrality—that States in conflict are prohibited from conducting hostilities within neutral territory—is not easily applied.[147] Attacks on neutral cyber infrastructure are, of course, forbidden, but parsing when an attack on a belligerent State that impacts infrastructure in a neutral State is unlawful is difficult, and the law remains unsettled.[148] Analysis is complicated because computers in the neutral State may be exploited by another State for its armed conflict ends without the knowledge of the neutral State. The main objective of neutrality analysis can be spoofed.[149]

---

[146] *See* DOD LAW OF WAR MANUAL, *supra* note 2, § 16.4, at 993; TALLINN MANUAL 2.0, *supra* note 5, chap. 20(1), at 553.

[147] TALLINN MANUAL 2.0, *supra* note 5, R. 150(4), at 555.

[148] *Id*.

[149] Hathaway et al., *supra* note 28, at 859.

Likewise, using cyber means to conduct armed conflict in neutral territory is unlawful.[150] The same principle applies to remotely conducting cyber operations in neutral territory.[151] According to the *Tallinn 2.0* IGE, the principle applies to private individuals or groups only if their conduct is attributable to a State in an international armed conflict.[152] Thus, an element of attribution is already part of the law of neutrality. Extending the analysis from private participants to States could become a straightforward part of neutrality law in the cyber domain. In the same way it is essential to know whether apparently private actions can be attributed to a State, it is important to know that an incoming cyber attack is attributable to a neutral State or to a third State that has exported malware through a neutral State's cyber infrastructure.

Although the *Tallinn 2.0* IGE agreed with the prevailing customary international law that State parties in conflict do not violate the law of neutrality by using the Internet to the extent components of it are located in neutral territory, a majority of the experts concluded that transmitting cyber weapons across a neutral State's cyber infrastructure violates international law. This conclusion was based on a provision of Hague Convention V that prohibits movement of munitions or supplies of war across the territory of a neutral State.[153] Illustrating the unsettled

---

[150] TALLINN MANUAL 2.0, *supra* note 5, R. 151, at 556.

[151] *Id*. R. 151(1).

[152] *Id.* R. 151(2).

[153] *Id.* R. 151(5–6), at 557; Convention No. V Respecting the Rights and Duties of Neutral Powers and Persons in Case of War on Land, art. 2, Oct, 19, 1907, 36 Stat. 2310.

nature of neutrality law in cyber, the United States DoD *Law of War Manual* interpreted the Hague Convention not to prohibit routing even destructive cyber weapons through a neutral State.[154]

The basic and fundamental attribution question reappears: In light of spoofing capabilities, and the dynamic features of malware, the legal questions about neutrality cannot be answered reliably without a process that fixes State responsibility. The same questions should be asked and answered before deciding whether a neutral State has knowingly allowed a State in conflict to use its cyber infrastructure for military purposes.[155] "Knowingly" presumes a duty on the part of neutral States.

IV. COULD ATTRIBUTION REVIEW IMPROVE THE ADAPTATIONS OF LOAC TO CYBER?

The development of cyber weaponry may in some ways make armed conflict less violent and thus less costly in human suffering. At the same time, the cyber domain also expands greatly the available targets in armed conflict. Yet, once an armed conflict has begun there is no legal requirement that incoming cyber activity is attributed to the enemy or some other state or non-state entity. The apparent mainstream view in the LOAC is that once an armed conflict has begun it is lawful for each side to presume that incoming cyber operations are the responsibility of the enemy

[154] DOD LAW OF WAR MANUAL, *supra* note 2, § 16.4.1, at 993–94; SCOLANS REPORT, *supra* note 17, at 64–65.

[155] TALLINN MANUAL 2.0, *supra* note 5, R. 152(5), at 599.

in the armed conflict.[156] In other words, the enemy is corporate, and the hostile acts that surround the core conflict are part and parcel of that conflict. While the corporate enemy concept is appealing from an operational perspective and may be a practical imperative in some conflict situations, a more nuanced approach that pays careful attention to the nature and source of ongoing cyber operations may enhance enforcement of LOAC requirements.

This part of the chapter will argue that an attribution process for cyber intrusions in an armed conflict would augment existing LOAC protections, particularly those that drive response targeting. In one important sense, LOAC compliance is about timing. Targeting analysis mirrors the kind of analysis that determines attribution, or at least similar questions are asked and answered. But attribution should occur *before* response targeting and is a separate inquiry. In colloquial terms, you have to know who is shooting at you in an armed conflict before you can lawfully shoot back. Target identification becomes an adjunct to attribution, where much of what is learned through an attribution process serves the targeting analysis.

---

[156] The Tallinn 2.0 IGE opined that the LOAC "does not embrace activities of private individuals or entities that are unrelated to the armed conflict," (*Id*. R. 80(8), at 377), and that the "applicability of [LOAC] does not depend upon the qualification of the situation under the jus ad bellum." *Id*. R. 80(9). Otherwise, the significant ongoing debate concerning application of LOAC to cyber is what it means for cyber activity to be "in the context of an armed conflict." *Id*. R. 80(5), at 376.

Incorporating a cyber attribution process during armed conflict could more precisely identify state (or some other) responsibility for cyber harms. As a consequence, States would be better able to isolate the nature and degree of response targeting called for by the cyber intrusion, taking care to meet traditional LOAC principles of distinction, proportionality, and precaution. Attribution could also assist in setting the metrics in some particular cyber response settings, such as determining the circumstances and scope of lawfully targeting critical infrastructure in the responsible State. For example, to the extent that attribution of enemy state responsibility is established with high confidence, greater discretion to target dual-use critical infrastructure could lawfully follow. Lesser confidence in attribution could demand more discrimination analysis in response targeting. International law could incorporate a sort of sliding scale regarding attribution—the greater the confidence in state responsibility, the more discretion should be used in targeting dual-use infrastructure; with less confidence, more attention should be paid to carefully parsing civilian impacts.

A. Calibrating Cyber Intrusions

Of course, cyber operations are not monolithic. Distinguishing among the types of cyber intrusions may help to calibrate application of the key LOAC principles during an armed conflict. Where the type of cyber operation is likely to cause significant cyber harm to civilians, attributing the source even some time after the fact may better protect civilians by deterring aggressive cyber operations in the future. To be sure, some cyber intrusions may be more amenable to attribution

than others, and the risk of harm to civilians may be greater in some kinds of operations, thus meriting enhanced attention toward attribution and state responsibility.

For example, the LOAC could distinguish operations that have as their objective shutting down or otherwise interfering with the functionality of a computer network—command and control systems or communications networks, for example, such as through a DDoS. A second category of cyber operations seeks to corrupt or destroy data on a computer system, not the system itself, while a third type attempts to take control of a system for the purpose of manipulating some physical object, such as a missile system, a dam, or an electric grid. In the latter case, the target is the physical thing, and the cyber operation is part of the means and methods of attack.[157] For the category of cyber operations that attempts to take control of a physical object, such as the enemy's missile defense system or the controls on its water supply, the LOAC analysis is more or less unaffected by the cyber means of impacting the object.[158]

Regarding cyber operations that target networks and data resident on them, the LOAC analysis is more complicated. To a degree, an attribution process is mirrored by traditional LOAC targeting analysis. For a target to constitute a military objective, the State is required to have

---

157 Lubell, *supra* note 4, at 255.

158 *See id*. at 256.

knowledge of the target's nature, including State ownership or responsibility.[159] Although the prospective targeting judgments are not now based on attribution of a cyber operation that has already occurred, the judgments about what is a military or civilian object and their derivative inquiries approximate the analysis that the responding State should undertake in assessing responsibility for incoming cyber activity.

Cyber operations during armed conflict that attempt to impact the functionality of computer systems or impact the data resident on them may cause significant harm to civilians. An attribution process for those operations could serve to more clearly identify the state responsible for the intrusions and, once held publicly accountable, deter excessive or especially harmful cyber operations. Depending on how the LOAC principles of distinction, proportionality, and precaution are applied to cyber operations, an attribution process could be attempted for cyber operations significantly impacting civilians during armed conflict or only those constituting "attacks" as understood in the LOAC.[160] If the practical assessment of the level of harm caused by a cyber

[159] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 100(8), at 438. Objects may qualify as military objectives due to their nature, location, purpose, or use. *Id*. The object must also make "an effective contribution to military action." *Id*. R. 100(15), at 440.

[160] Lubell, *supra* note 4, at 260–61; TALLINN MANUAL 2.0, *supra* note 5, R. 92, at 415 (defining a "cyber attack" for the purposes of LOAC as a "cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction").

operation is used as the measure of whether there has been an attack,[161] operations that impact the functionality of the targeted system to the extent that components must be replaced are attacks[162] and could be made subject to attribution.

A cyber operation that manipulates, destroys, or corrupts data on a computer or server in a way that does not affect or destroy the functionality of the computer or system is not an attack on an object and apparently is not regulated by the LOAC or its distinction and proportionality principles. Although a minority of the *Tallinn 2.0* IGE pointed out that the traditional LOAC principle would not protect "essential civilian datasets such as social security data, tax records, and bank accounts,"[163] contrary to the overarching goal of protecting civilians during armed conflict, the IGE confirmed that data is not an object for LOAC purposes.[164] An attribution process for incoming cyber activity that targets essential civilian data resident on computer systems could expose enemy overreach in armed conflict. Establishing state responsibility for these attacks on data could also lead to a more nuanced LOAC approach to treating data as an object in targeting.

---

[161] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 92(3), at 415 ("The crux of ['acts of violence'] lies in the effects that are caused.").

[162] Lubell, *supra* note 4, at 265–66; *see* TALLINN MANUAL 2.0, *supra* note 5, R. 92, at 415 (including operations that are reasonably expected to cause damage or destruction to objects in the definition of cyber attack).

[163] TALLINN MANUAL 2.0, *supra* note 5, R. 100(7), at 437.

[164] *Id*. R. 100(6).

Similarly, the approach taken by the United States that treats war-supporting or war-sustaining objects as valid military targets broadens the scope of dual-use and civilian components of critical infrastructure that are vulnerable to attack consistent with the LOAC.[165] Although a majority of the *Tallinn 2.0* IGE rejected the United States view "on the ground that the connection between war-sustaining activities and military action is too remote,"[166] theirs is expert opinion, not law. In any event, the debates and fine contextual lines between war-fighting, war-supporting, and war-sustaining activities illustrate that a one-size-fits-all rule for dual-use targets does not serve well the overarching LOAC objective of protecting civilians in armed conflict. As with attacks on data, an attribution process for incoming cyber activity that targets various categories of dual-use but war-related cyber infrastructure could expose enemy excesses, serve to model response attacks, and possibly deter future such attacks.[167]

One widespread type of cyber intrusion is the DDoS attack. These operations involve coordinated botnets where virus-infected hijacked "zombie" computers overwhelm servers by systematically and continuously visiting designated websites.[168] DDoS attacks are typically carried out by networks of hackers, but State involvement is often suspected, as was the case with

---

[165] *See* DOD LAW OF WAR MANUAL, *supra* note 2, § 5.7.6.2, at 210.

[166] TALLINN MANUAL 2.0, *supra* note 5, R. 100(19), at 441.

[167] *See id*. R. 100(20)–(22), at 442.

[168] *See* Hathaway et al., *supra* note 28, at 837–38.

suspected Russian State involvement in the 2007 Estonia and 2008 Georgia DDoS attacks.[169] Conclusive attribution was not established for these and other DDoS operations because of the

[169] *Id*. at 838; *See* CONNELL & VOGLER, *supra,* note 8, at 13 (describing how the attack on Estonia lasted for about a month, "forcing most sites to either shut down or sever their international connections" and preventing the country from communicating with the outside world). *See also* Ian Traynor, *Russia accused of unleashing cyberwar to disable Estonia*, THE GUARDIAN (May 16, 2007), https://www.theguardian.com/world/2007/may/17/topstories3.russia (describing one Estonian citizen as stating, in the immediate aftermath of the flurry of DDoS attacks, "[t]he cyber-attacks are from Russia. There is no question. It's political."); TIKK ET AL., *supra* note 27, at 12 (2008) (describing the "wide public understanding that the attacks were at least tolerated by the Russian authorities, if not coordinated or supported by them," based Russia's large-scale collusion of interest between South Ossetia and the Russian government and because "the coordination of and support to attacks took place mainly in the Russian language and was conducted on Russian or Russia-friendly forums"); *see* Noah Shachtman, *Top Georgian Official: Moscow Cyber Attacked Us – We Just Can't Prove It*, WIRED (Mar. 11, 2009), https://www.wired.com/2009/03/georgia-blames/ (describing a Georgian National Security Council official as stating that there is "plenty of evidence" that the attacks were "directly *organized*" by the Russian government without providing any evidence to conclusively link Moscow to the attacks); *see* Brian Krebs, *Report: Russian Hacker Forums Fueled Georgia Cyber Attacks*, WASH. POST (Oct. 16, 2008, 3:15 PM),

anonymity created by the botnets utilizing unsuspecting computers from around the world.[170] Although temporarily shutting down websites causes inconvenience and delay in transacting

---

http://voices.washingtonpost.com/securityfix/2008/10/report_russian_hacker_forums_f.html (noting that after an "exhaustive inquiry," there is "no smoking gun in the hands of the Russian government," although the attack was coordinated through a Russian online forum that appeared to have been prepped with target lists and details about Georgian Web site vulnerabilities"); and *see* Healey, *supra* note 20 (discussing the NATO DDoS incidents in 1999, and noting that while it was initially thought that the Serbian military directly conducted the attacks, such a claim is "often made about incidents later proven to be conducted by non-states").

[170] TIKK ET AL., *supra* note 27, at 12 ("[M]ajor DDOS attacks observed were globally sourced, suggesting a botnet (or multiple botnets) behind them."). S*ee generally,* JOSE NAZARIO, COOP. CYBER DEF. CTR. OF EXCELLENCE, POLITICALLY MOTIVATED DENIAL OF SERVICE ATTACKS, (last visited Feb. 19, 2018), https://ccdcoe.org/publications/virtualbattlefield/12_NAZARIO%20Politically%20Motivated%20DDoS.pdf (describing the use of the botnets to anonymously conduct DDoS attacks in Estonia in 2007, China in 2008, Georgia in 2008, and in Ukraine in 2008); *see, e.g.*, *HIDDEN COBRA - North Korea's DDoS Botnet Infrastructure*, DEP'T OF HOMELAND SEC. (June 13, 2017), https://www.us-cert.gov/ncas/alerts/TA17-164A https://www.us-cert.gov/ncas/alerts/TA17-164A (providing technical details about how to avoid IP addresses

business or government, not injury or destruction, DDoS operations can be costly.[171] The consensus view among scholars is that the Estonia attacks were not subject to the LOAC because

associated with a malware variant that is used by North Korea to manage its DDoS botnet infrastructure).

[171] *See, e.g.*, Hathaway, *supra* note 28, at 819 (describing the 2010 DDoS attack as one that "took the entire population of Burma off the internet immediately preceding the country's first national election in twenty years"); *id*. at 837 (describing the effects of the DDoS attacks on Estonia as "life-threatening," as the emergency line to call for an ambulance was out of service for an hour); Damien McGuinness, *How a cyber attack transformed Estonia*, BBC NEWS (Apr. 27, 2017), http://www.bbc.com/news/39655415 ("Online services of Estonian banks, media outlets and government bodies were taken down by unprecedented levels of internet traffic."); Schmitt, *supra* note 101, at 89; Matthew J. Skerlov, Solving the Dilemma of State Responses to Cyberattacks: A Justification for the use of Active Defenses Against States who Neglect their Duty to Prevent (Apr. 2009), file:///C:/Users/benja/Downloads/ADA517821.pdf (thesis presented to The Judge Advocate General's School, United States Army, page 5)("[C]yberattacks from Russia crippled the Estonian government and commercial computer networks. These attacks lasted approximately three weeks, disrupted Estonia's ability to govern, harmed Estonia's economy, and damaged their networks so badly that Estonia had to reach out to its NATO allies for help recovering."). *See also* Kim Zetter, *Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid*, WIRED (Mar. 3, 2016, 7:00 AM), https://www.wired.com/2016/03/inside-cunning-

there was no armed conflict with Russia, while the 2008 cyber attacks on Georgia were part of an armed conflict with Russia and were thus subject to the LOAC.[172] So, too, are ongoing DDoS attacks by Russia in its armed conflict with Ukraine.[173] For those DDoS attacks during an armed conflict, whether they constitute LOAC "attacks" or cyber harm, attributing significant attacks could call attention to unlawful interventions by states and perhaps deter some future operations. Improvements in attribution technology, along with commitments from affected states to assign state responsibility for DDoS attacks, could combine to make this category of cyber activity less likely during armed conflict.

---

unprecedented-hack-ukraines-power-grid/ (describing how, even after power was restored to civilians after the cyber attack on a Ukrainian power grid, the damage from the attack required the breakers to be controlled manually, and control centers remained partially incapacitated for more than two months).

[172] *See* TIKK, ET AL., *supra* note 27, at 19–23; TALLINN MANUAL 2.0, *supra* note 5, R. 80(3), at 376; *see* Lubell, *supra* note 4, at 254.

[173] *See* Jan Stinissen, *A Legal Framework for Cyber Operations in Ukraine*, *in* CYBER WAR IN PERSPECTIVE: RUSSIAN AGGRESSION AGAINST UKRAINE 131 (2015) ("During the occupation of Crimea and the armed conflict in Eastern Ukraine, the Law of Armed Conflict applies [and] regulated the conduct of all. . . cyber actors."); TALLINN MANUAL 2.0, *supra* note 5, R. 80(3), at 376; Schmitt & Widmar, *supra* note 4, at 380.

Another form of cyber operation, a semantic attack, involves surreptitiously inputting inaccurate information in a computer system while causing the computer to appear to operate normally while it is failing.[174] Examples in the security realm include an abandoned United States plan in 1999 to provide false target data into the Serbian defense network, thereby interfering with Serbia's capacity to target NATO planes, and a 2007 Israeli semantic operation that compromised the Syrian air-defense system causing Syrian radars to show clear skies at the same time the Israeli Air Force conducted a strike against a nuclear facility in Syria.[175]

The 2010 Stuxnet attack began as a semantic attack but evolved into an operation that disrupted the nuclear facility. Stuxnet has not been officially attributed,[176] and there was no armed

---

[174] *E.g.*, MARTIN C. LIBICKI, WHAT IS INFORMATION WARFARE? 77 (1995).

[175] *See, e.g.*, RICHARD A. CLARKE & ROBERT K. KNAKE, CYBER WAR: THE NEXT THREAT TO NATIONAL SECURITY AND WHAT TO DO ABOUT IT 1–9 (2010).

[176] *See* David E. Sanger, *Obama Order Sped Up Wave of Cyberattacks Against Iran,* N.Y. TIMES (June 1, 2012), http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?mcubz=0 ("[F]orensic investigations into the inner workings of the code. . . . came to no conclusions about who was responsible."); Jon R. Lindsay, *Stuxnet and the Limits of Cyber Warfare*, 22 SEC. STUD. 365, 400 (2013) (discussing how the most persuasive evidence for attributing Stuxnet to the U.S. or Israel is only circumstantial). Besides having the means and motive, there is no direct evidence linking the United States and/or Israel to the Stuxnet attacks. *See, e.g.*, Rid & Buchanan, *supra* note 36, at

conflict between the apparently responsible parties, the United States and perhaps Israel and Iran. When semantic operations coincide with conventional attacks, attribution judgments are less difficult. Aside from conventional attack or armed conflict indicators, however, attribution cannot be accomplished expeditiously because the computer disruption is not knowable until the conventional or kinetic attack occurs. However, in the course of an armed conflict, attributing semantic attacks as part of LOAC compliance could be a potentially useful tool for deterring unlawful cyber activity.

B. Precautions

The principle of precautions in attack requires a commander to take "feasible" precautions to minimize harm to civilians from an attack.[177] The obligation to take feasible precautions is

---

21–22 ("No non-state actor, and indeed few governments, would likely have the capability to test Stuxnet, let alone build and deploy it."); Kim Zetter, *How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History*, WIRED (July 11, 2011, 7:00 AM), https://www.wired.com/2011/07/how-digital-detectives-deciphered-stuxnet/ (noting that despite the talents of private security experts, the strongest language regarding attribution is only that "[t]he sophistication of the code, plus the fraudulent certifications, and [having] Iran at the center of the fallout [makes] it look like Stuxnet could be the work of a government cyber army – maybe even a United States cyberarmy.").

[177] Additional Protocol I, *supra* note 13, at art. 57(2)(a)(i).

manifest in the customary international law obligation to take "constant care" in reducing harm to civilian persons or objects.[178] The traditional role of a precautions analysis lies in target identification and verification, and in assessing collateral harm to civilians that may result from a military operation.[179]

Tallinn Manual 2.0 offers a Rule on precautions, based on API and part of customary law in international and non-international armed conflicts: "During hostilities involving cyber operations, constant care shall be taken to spare the civilian population, individual civilians, and civilian objects."[180] Intended to supplement the distinction and proportionality principles and corresponding Rules,[181] the precautions Rule "requires commanders and all others involved in the operations to be continuously sensitive to the effects of their activities on the civilian population

[178] *Id*. at art. 57(1); INT'L COMM. OF THE RED CROSS, CUSTOMARY INTERNATIONAL HUMANITARIAN LAW VOLUME I R. 15, at 51 (Jean-Marie Henckaerts & Louise Doswald, eds. 2009); INT'L COMM. OF THE RED CROSS, CUSTOMARY INTERNATIONAL HUMANITARIAN LAW VOLUME II 337–39 (Jean-Marie Henckaerts & Louise Doswald, eds. 2005).

[179] *See* Geoffrey S. Corn, *War, Law, and the Oft Overlooked Value of Process as a Precautionary Measure*, 42 PEPPERDINE L. REV. 419, 435–36 (2015); TALLINN MANUAL 2.0, *supra* note 5, R. 115(4), at 479, R. 116(2), at 479–80.

[180] TALLINN MANUAL 2.0, *supra* note 5, R. 114, at 476.

[181] *Id*. R. 114(3), at 477.

and civilian objects, and to seek to avoid any unnecessary effects thereon."[182] The "constant care" admonition "requires situational awareness at all times" in the cyber context.[183] A related Rule on target verification applies to cyber operations that qualify as an "attack"[184] and requires all "feasible precautions"[185] that could include "gathering intelligence . . . to determine the attack's likely effects . . . ."[186] According to the IGE, when target verification is not practically possible, "the decision-maker may have to refrain from conducting an attack" or modify it.[187]

According to the IGE, precautions must extend to "the choice of means or methods of warfare employed in . . . an attack, with a view to avoiding, and in any event to minimising, incidental injury to civilians, loss of civilian life, and damage to or destruction of civilian objects."[188] Recognizing that cyber infrastructure is dual-use, the IGE stressed that commanders "must take all feasible precautions to avoid, or at least minimise, indirect as well as direct collateral

[182] *Id*. R. 114(4).

[183] *Id*. R. 114(5).

[184] *Id*. R. 115(1), at 478.

[185] *Id*. R. 115(4), at 479.

[186] *Id*.

[187] *Id*. R. 115(5).

[188] *Id*. R. 116, at 479–80.

damage."[189] Related Rules describe similar precautions concerning proportionality,[190] choice of targets,[191] warnings,[192] and against the effects of cyber attacks.[193] An example chosen by the IGE focused on a choice of targets and involved disrupting enemy command and control. Given a choice between attacking the dual-use electric grid and the enemy's command and control network directly, the latter must be chosen if it is expected to achieve the desired military advantage because of the significant collateral harm to civilian infrastructure.[194]

The IGE also included a Rule on "passive precautions," those that must be taken by a *defender* based on the effects of cyber attacks.[195] Examples offered include segregating military from civilian cyber infrastructure and civilian systems from the internet, backing up important civilian data, arranging in advance for repairs of systems likely to be harmed, and using anti-virus programs to protect civilian systems.[196]

---

[189] *Id*. R. 116(5), at 480.

[190] *See id*. R. 117, at 481.

[191] *See id*. R. 118, at 481.

[192] *See id*. R. 120, at 485.

[193] *See id*. R. 121, at 487.

[194] *See id*. R. 118(8), at 483.

[195] *Id*. R. 121, at 487–488.

[196] *See id*. R. 121(3), at 488.

API obligates attackers selecting military objectives to choose the one which "may be expected to cause the least danger to civilian lives and to civilian objects."[197] The same article requires attackers to take all feasible precautions in the choice of means and methods to avoid or at least minimize incidental loss of civilian life, injury to civilians, and damage to civilian objects, and refrain from attacks which "may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated."[198] Of course, in general a cyber operation may cause fewer casualties and damage than a kinetic strike and, for that reason, may be preferable and more likely to meet the objectives outlined in the Protocol. However, on a case-by-case basis, where state militaries choose to launch cyber operations against dual-use civilian systems that the military rely on (instead of an internal military cyber target), the impact on civilians may be significant and avoidable. Thus, the *DoD Law of War Manual* recognizes that the "obligation to take feasible precautions may be of greater relevance in cyber operations . . . because this obligation applies to a broader set of activities than those to which other law of war rules apply."[199]

[197] Additional Protocol I, *supra* note 13, at art. 57(3).

[198] *Id*. at art. 57(2)(a)(i–iii).

[199] DOD LAW OF WAR MANUAL, *supra* note 2, § 16.5.3, at 997.

Relatedly, the presumption that war-sustaining objects are proper targets that "contribute[] to military action"[200] could be rebutted if targeting analysis reveals significant civilian losses or insufficient connection to the military effort.[201] At present there are no standards to assess whether an object will in fact have a military use. In at least some instances the cyber response to an intrusion during armed conflict is expected to target computers, systems, or networks similarly situated to those harmed in the incoming operation. Military objectives in cyber can include computers, networks, and other tangible components of cyber infrastructure. In addition, interconnected networks and systems do not lend themselves to clear segregation of civilian and military uses or purposes. As with dual-use targeting, an attribution process attached to the cyber operations that prompt the cyber responses may improve the decisions.

Article 58 of API obligates parties to an armed conflict "to the maximum extent feasible: a) . . . endeavor to remove . . . civilian objects under their control from the vicinity of military objectives."[202] The obligation to take feasible precautions may thus be more fully realized by including an attribution process during armed conflict in order to protect some cyber activities in the civilian sphere, independent of the targeting analysis. To the extent that the *Tallinn 2.0* rules

---

[200] Commander's Handbook, *supra* note 118, at § 8.2 (defining "military objects" only as those objects which, "by their location, purpose, or use, *effectively contribute* to the enemy's war-fighting or war-sustaining capability. . .").

[201] Schmitt & Widmar, *supra* note 4, at 392–93.

[202] Additional Protocol I, *supra* note 13, at art. 58(a).

and commentary reflect customary international law, the precaution principle and its derivative doctrine, along with the requirement that the State have knowledge of the nature of a proposed military objective in targeting, provide harmonious protections alongside an attribution analysis for the cyber components of armed conflict. The attribution analysis precedes the more general LOAC requirements of precautions and discrimination.

C. A proposed LOAC rule

Considering attribution below the armed conflict threshold, the *Tallinn 2.0* IGE agreed "that as a general matter, States must act as reasonable States would in the same or similar circumstances when considering responses to them."[203] As explained in Part I, the IGE opined that

> reasonableness . . . depends on such factors as . . . the reliability, quantum, directness, nature (e.g., technical data, human intelligence), and specificity of the relevant available information when considered in light of the attendant circumstances and the importance of the right involved.[204]

[203] TALLINN MANUAL 2.0, *supra* note 5, at 81.

[204] *Id*. at 81–82.

The IGE found that "as a general matter the graver the underlying breach . . . , the greater the confidence ought to be in the evidence relied upon by a State considering a response[205] . . . . Because there is no international or domestic law on how much evidence suffices for attribution of State responsibility, the attribution bar is set very low by international law. In addition, the legal standards for attribution are malleable to the extent that the evidence of attribution is not required to be and usually is not shared publicly.[206] In short, the law on attribution is anything but robust.

Consider this proposal:

> In conducting military operations during armed conflict, the commander (or other decision maker) must act as a reasonable commander in same or similar circumstances would to attribute the source of a cyber operation before responding with kinetic or cyber weapons, or as soon thereafter as practical. The attribution requirement varies depending on the value of the target and the quality and quantity

[205] In support of its position, the IGE cited Iran v. U.S., 2003 I.C.J. at 33 (separate opinion of Judge Higgins); U.K. v. Alb., 1949 I.C.J. at 17; Bosn. & Herz. v. Serb. & Montenegro, 2007 I.C.J. at ¶¶ 209–10; and Croat. v. Serb., 2015 I.C.J. at ¶ 178.

[206] *See* Brian J. Egan, Remarks on International Law and Stability in Cyberspace at Berkeley Law School (Nov. 10, 2016), https://www.law.berkeley.edu/wp-content/uploads/2016/12/egan-talk-transcript-111016.pdf; TALLINN MANUAL 2.0, *supra* note 5, at 83.

> of available attribution analysis or data. For instance, a state may pursue a very high-value target with less certainty of attribution than in situations involving a target that is of low value. High value can be measured by value to the enemy or the seriousness of the target's actions in relation to the state's own operations.[207]

We know from experience outside the armed conflict setting that attribution is an imperfect process, one that in many cases improves over time through intelligence collection, information sharing, and political or diplomatic discussions. In many cases, even during armed conflict, it may behoove a State to avoid a rush to judgment or immediate counterattack in response to a significant cyber intrusion. Delayed attribution may be more reliable and more authoritative,[208] and a solid evidence-based attribution may enable the States in conflict to avoid international law violations for targeting innocent parties. At the same time and based on limited experience in the *jus ad bellum* realm, legally prescribed standards or criteria for attribution are not likely to be effective.[209]

---

[207] *See* TALLINN MANUAL 2.0, *supra* note 5, at 81–82.

[208] *See* GUITTON, *supra* note 6, at 150–51, 160 (noting that it is unrealistic to expect high-confidence attribution in real-time, if ever, and that reducing the time for attribution does not make sense politically).

[209] *Id*. at 80–81.

Because attribution is usually based on judgment, paying too much attention to standards can cloud the political or policy process required before an attribution judgment is reached.[210]

An analogy to state responsibility and the unlawful intervention rules is instructive in the armed conflict setting in two different respects. The first underscores the importance attached to attribution and state responsibility for cyber intrusions below the armed conflict threshold. As explained above,[211] attribution is *required*, a prerequisite to finding state responsibility for an unlawful intervention by cyber means. The second reminds us that international law has not yet fully adapted to the cyber domain, in the *jus ad bellum* or *jus in bello*. The *jus ad bellum* permits countermeasures in response to an unlawful intervention.[212] Countermeasures may be cyber in nature or not, below the use of force threshold that would be unlawful but for the purpose of stopping the unlawful intervention. However, countermeasures also require prior attribution and notice to the offending State so that it has the opportunity to discontinue its unlawful conduct. The purpose of the countermeasures is to induce compliance with international law.[213] Because state attribution in cyber can be so difficult and time-consuming, countermeasures often are realistically

---

[210] *Id*. at 81. *See also, id.* at 86, 90–92, 98–99 (noting, respectively, that attribution is easily malleable; that establishing the facts can be difficult, but political will can overcome strict standards; and that using criteria to attribute attacks can be used to manipulate evidence).

[211] See *supra* text accompanying note 12.

[212] *See* TALLINN MANUAL 2.0, *supra* note 5, R. 20(1), at 111.

[213] *Id*. R. 21(5), at 118.

unavailable. Countermeasures implemented after the time has passed for encouraging the offending State to stop its intrusions become unlawful punishment.[214] So while the *jus ad bellum* doctrine has not adapted to the realm of cyber conflict, its bedrock principles underscore the importance of attribution before attaching state responsibility for an unlawful act. Although it is unrealistic to expect authoritative attribution in the real-time environment of armed conflict, identifying the source of cyber operations will serve important purposes even after the armed conflict is over.

The importance attached to attribution of cyber incidents below the armed conflict threshold supports the argument for extending an attribution element to the LOAC. Although response targeting analysis and precautions in the LOAC replicate and overlap with some of the value that could be derived from attribution, attribution would attempt to answer the threshold question of who is responsible for the cyber intrusion, setting the stage for more reliable targeting and precautions.

One possible additional component of building in an attribution step for cyber operations in armed conflicts beyond identifying the responsible party is providing some details about the intrusion.[215] The need to protect intelligence sources and methods will continue to foreclose disseminating much of the attribution process and details about intrusions to all but senior officials

---

[214] *Id*. R. 22(5), at 124.

[215] *See* Rid & Buchanan, *supra* note 32, at 26; GUITTON, *supra* note 6, at 47.

and elected leaders. However, particularly when the cyber intrusion causes considerable cyber harm to civilians or civilian infrastructure, publicizing at least some of the details of the operation and responsible parties can enhance the credibility of the victim State and deter adversaries who fear the impact of widespread knowledge of their cyber activities.[216] Communications about attribution are also likely to improve attribution and the collective defenses against cyber attacks.[217]

Public attribution may also change the behavior of the cyber adversaries. The most prominent, recent example is the May 2014 decision by the United States to indict members of the Chinese PLA for computer fraud and abuse, among other crimes. Although the indictment detailed the criminal economic espionage conducted by the PLA Unit, it did not reveal much of the evidence in support of attribution. The not so subtle message was that, if the Chinese did not desist, that information could be released.[218] Not long after the indictments, China and the United States agreed on some parameters for protecting commercial secrets from cyber espionage.[219]

---

[216] Rid & Buchanan, *supra* note 32, at 26–27.

[217] *See id*. at 28; GUITTON, *supra* note 6, at 153–54.

[218] *See* Knake, *supra* note 84, at 5 (noting that the Obama administration named the foreign actors behind some cyberattacks and arguing that making public attribution could be a deterrent).

[219] *See* Barack Obama, President of the U.S., Remarks by President Obama and President Xi of the People's Republic of China in Joint Press Conference (Sept. 25, 2015),

V. CONCLUSIONS

Attribution is a process, not a technical challenge. Conclusive proof of state responsibility is elusively difficult, although circumstantial proof is often available.[220] By far, the biggest constraints States confront in deciding whether and under what circumstances to pursue attribution are political or strategic,[221] including during armed conflict. This chapter has shown, however, that during an armed conflict, attribution can add legitimacy to a lawful conflict. Mistakes can harm innocent civilians, lead to distrust or worse among States, and involve actors and entities in a conflict who were heretofore uninvolved. Given the advanced attribution capabilities now available to sophisticated State adversaries, law reform could remediate the attribution problem by adopting a standard of proof sufficient to establish state responsibility as part of the LOAC.

---

https://obamawhitehouse.archives.gov/the-press-office/2015/09/25/remarks-president-obama-and-president-xi-peoples-republic-china-joint.

[220] *See* GUITTON, *supra* note 6, at 45, 70–71, 185–88; Lin, *supra* note 45, at 77; Rid & Buchanan, *supra* note 32, at 7. Thomas Rid asserts in 2017 that "[i]t is now generally accepted that attributing computer network operations reliably is possible in principle – an assumption that a few years ago was still contested." THOMAS RID, CYBER WAR WILL NOT TAKE PLACE 188 (Oxford Univ. Press 2017) (2013).

[221] *See* Banks, *supra* note 50, at 1511–12; GUITTON, *supra* note 6, at 11.

A process for assigning state responsibility for cyber intrusions may accomplish important objectives in armed conflict. States responsible for harmful cyber operations may be deterred in the future based on the knowledge that they may be held accountable for their unlawful acts. At the same time, military commanders would have more information at their disposal in response targeting during the armed conflict, whether through cyber or non-cyber means. Better targeting guidance could, in turn, enhance compliance with the LOAC, including appropriately tailored precautions. Finally, States may avoid making unlawful mistakes in the armed conflict because of weak or non-existent efforts to attribute incoming attacks. In the aggregate, attribution of cyber attacks in an armed conflict may act as a deterrent to unlawful uses of cyber tools and serve to better protect civilians, particularly if the attributed attacks expose an enemy State's cyber attacks against civilians or civilian infrastructure.[222]

Of at least equal importance is avoiding what Thomas Rid and Ben Buchanan call a "lax de-facto norm of negligible consequences."[223] If States continue to believe that they may engage in cyber operations at will because of the lack of consequences for past bad behavior, such impunity could exacerbate harm to civilians in armed conflict. Although it is commonly assumed that the most industrialized and technically advanced States are the most vulnerable to cyber attacks from otherwise less advanced States, in attribution the reverse is true. Sophisticated States

[222] *See* Rid & Buchanan, *supra* note 32, at 26–28; Knake *supra* note 84, at 5 ("Making public attribution of attacks a routine practice could be a deterrent.").

[223] Rid & Buchanan, *supra* note 32, at 33.

have a greater pool of talent and resources to work toward attribution, and they have the capabilities to gather covert intelligence but hide its sources and methods.[224] Similarly, effective attribution can dispel the adage that, because of the architecture of the Internet, cyber operations are dominated by the attackers. The usual theory is that the defense has to be vigilant and correct all the time while the offense only needs to get it right once. In attribution the offender need make only one mistake, and it could be discovered by effective attribution analysis.[225] Of course, quality attribution requires considerable human and technical resources; the task can be expensive. High quality attribution often also requires considerable time; the results are not instantaneous. Finally, sophisticated adversaries are increasingly able to employ good operational security to obstruct forensic investigation. Investigators often have to rely on the adversaries to make mistakes. They usually do.[226]

Finally, if most cyber intrusions may now or in the foreseeable future be accurately attributed, consider the importance of attribution capabilities beyond the cyber realm, to include autonomous weapons and various artificial intelligence applications in the delivery of weapons of the future. In these frontiers, the knowledge that attribution techniques and processes provide may be essential to better assuring the accountability and lawfulness of weapons in future armed conflicts.

---

[224] *Id*. at 32.

[225] *Id*.

[226] *Id*.