SOVEREIGNTY AND CYBER ATTACKS: TECHNOLOGY'S CHALLENGE TO THE LAW OF STATE RESPONSIBILITY

PETER MARGULIES*

Cyber threats pose fresh challenges to sovereignty and to international law on state responsibility. In addressing kinetic attacks, international law defines state responsibility narrowly. A party asserting that a state is responsible for a kinetic attack must comply with the 'effective control' test adopted by the International Court of Justice in the Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America) *decision or, at the very least, with the 'effective control' test adopted by the International Criminal Tribunal for the Former Yugoslavia in* Prosecutor v Tadić. *Driven by concerns about the risks of escalation, the International Law Commission's ('ILC')* Draft Articles on Responsibility of States for Internationally Wrongful Acts *hardened this narrow approach. The recently published* Tallinn Manual on the International Law Applicable to Cyber Warfare *('*Manual*') tracks the ILC's analysis. While the* Manual *is an exceptionally valuable effort to apply* lex lata *to the fluid cyber realm, caution may not serve international law in this context. Cyber reflects what I call 'attribution asymmetry': cyber threats from private groups assisted by states are both more difficult to trace than kinetic attacks for victims and easier to control for the state providing the assistance. Because of this asymmetry, the international law on state responsibility for kinetic attacks does not adequately address the issue of cyber attacks. A test of virtual control would be more effective, imposing responsibility on a state that has provided financial or other assistance to private groups. The virtual control test would deter states from using private groups to engineer plausible deniability. This heightened deterrence provides a more useful template for the development of international law in the cyber domain.*

CONTENTS

* BA (Colgate University), JD (Columbia University); Professor of Law, Roger Williams University School of Law. I thank Maxine Kutner for her research assistance and anonymous reviewers for comments on a previous draft.

I INTRODUCTION

The cyber age will expose sovereignty to new challenges. Cyber attacks represent new ways of intruding on the sovereign prerogatives of states. As usual, the law has struggled to keep pace with technology. Recent attempts to pinpoint the application of international law to cyberwarfare have made a useful beginning,[1] but have also relied on importing doctrine that does not fit cyberthreats.[2] This commentary suggests that international law on state responsibility for kinetic attacks is inadequate to address state responsibility for cyber attacks. Because of the difficulty in detecting cyber attacks from the outside, coupled with the ease of controlling them from the inside, the test for state responsibility for cyber attacks should be substantially broader than it is in other contexts. Only a broader standard will deter substantial intrusions on sovereignty through the use of cyber-weapons.[3]

The *Tallinn Manual on the International Law Applicable to Cyber Warfare* ('*Manual*'), the most systematic effort to adapt the law of armed conflict ('LOAC') to cyber,[4] takes a cautious stance. The *Manual* relies on the International Law Commission's ('ILC') *Draft Articles on Responsibility of*

---

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

[2] Such doctrinal principles have been imported from cases including: *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America) (Merits)* [1986] ICJ Rep 14 ('*Nicaragua*'); *Application of the* Convention on the Prevention and Punishment of the Crime of Genocide *(Bosnia and Herzegovina v Serbia and Montenegro) (Judgment)* [2007] ICJ Rep 43, 208 [400] ('*Genocide Case*'); *Prosecutor v Tadić (Judgment)* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [131], [145] ('*Tadić*').

[3] For other pieces that argued for a broader standard, but did so on different reasoning and before publication of the *Tallinn Manual on the International Law Applicable to Cyber Warfare* ('*Manual*'), see Michael Gervais, 'Cyber Attacks and the Laws of War' (2012) 30 *Berkeley Journal of International Law* 525, 549–50 (arguing that a victim state may use force against a state that refuses a request by the victim state to take appropriate measures to stop cyber attacks emanating from its territory); Catherine Lotrionte, 'State Sovereignty and Self-Defense in Cyberspace: A Normative Framework for Balancing Legal Rights' (2012) 26 *Emory International Law Review* 825, 890 (suggesting that the victim state should have recourse if the state (the 'territorial state') whose territory was used by non-state groups to stage cyber attacks was 'directly or indirectly involved' in the attacks); 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' (2009) 201 *Military Law Review* 1 (arguing for aggressive action against third-party nations who fail to take precautions against having servers under their sovereign control or on their territory used for attacks against other nations).

[4] See Schmitt, above n 1.

States for Internationally Wrongful Acts ('*Draft Articles*'),[5] which tie state responsibility to showing that a private party is 'acting on the instructions of, or under the direction or control of' a state.[6] Case law[7] also reflects this 'relatively stringent' approach.[8] As the *Manual* notes,[9] the International Court of Justice ('ICJ') has adopted a test of 'effective control' of the state over non-state actors for the purposes of assessing the wrongfulness of state action[10] and holding a state accountable.[11] In a broader formulation, the International Criminal Tribunal for the Former Yugoslavia ('ICTY') has held that state officials may be accountable if they exercised 'overall control' over a group or entity.[12] The *Manual* notes that neither standard would support state responsibility for the 'mere financing and equipping' of private groups.[13] A further narrowing of state responsibility results from LOAC's usual requirement that a party to an armed conflict be either a state or an organised armed group.[14] For individuals and groups without the structure required under international humanitarian law for recognition as organised armed groups, the higher 'effective control' standard applies for attribution of state responsibility.[15]

The caution displayed by the *Manual* camouflages significant risks. On the surface, it might seem self-evident that any restatement of law should start exactly where the *Manual* starts: with a clear account of *lex lata*. Indeed, the most salutary aspect of the *Manual*, and one clearly intended by the authors,[16] is its utility as a starting point for debate and analysis. However, stressing *lex lata* is not without disadvantages. In a fast-moving, fluid realm such as cyber, a cautious approach may fail to harmonise with the rate of change in the field. Reliance on the opinions of certain international tribunals, such as the ICJ, reflects these risks. The ICJ is not always in step with views of the law propounded by scholars and states. For example, its finding that an attack by an organised non-state actor could not justify state action in self-defence under art 51 of the *Charter of the United Nations* ('*UN Charter*')[17] has been widely

---

5 International Law Commission, *Report of the International Law Commission on the Work of Its Fifty-Third Session*, UN GAOR, 56th sess, Supp No 10, UN Doc A/56/10 (2001) ch IV(E) ('*Draft Articles on Responsibility of States for Internationally Wrongful Acts*') ('*Draft Articles*'). Droege relies on the *Draft Articles* in taking a narrow view of attribution for cyber attacks: Cordula Droege, 'Get Off My Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians' (2012) 94(886) *International Review of the Red Cross* 533, 543–4.

6 *Draft Articles*, UN Doc A/56/10, ch IV(E) art 8.

7 Ibid.

8 Schmitt, above n 1, 33 (Rule 6, [11]).

9 Ibid, 32–3 (Rule 6, [10]).

10 *Nicaragua* [1986] ICJ Rep 14, 64 [115].

11 *Genocide Case* [2007] ICJ Rep 43, 208 [399]–[401].

12 *Tadić* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [131], [145].

13 Schmitt, above n 1, 32–3 (Rule 6, [10]).

14 See *Prosecutor v Tadić (Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction)* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-T, 2 October 1995) [70].

15 Schmitt, above n 1, 32–3 (Rule 6, [10]).

16 Ibid 6 (noting that the *Manual*'s drafters 'sought to capture all reasonable positions for inclusion in the *Tallinn Manual*'s Commentary').

17 See *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory (Advisory Opinion)* [2004] ICJ Rep 136, 194 [139].

criticised both within the Court[18] and by scholars.[19] Viewed in this light, the cautious approach of the *Manual* may risk premature irrelevance as state practice overtakes it. Even worse, the *Manual* might ultimately not serve as the springboard for debate that its drafters hoped. Instead, the *Manual* could become a cumbersome anchor that weighs down efforts to adapt to new challenges, as critics have contended of the *Draft Articles* on which the *Manual* relies.[20]

These concerns are particularly apt for the attribution of state responsibility for cyber attacks. The test for attribution shapes the accountability of states for conduct that affects other sovereign nations. Although a state can violate international law without meeting the test,[21] a victim state is sharply restricted in its choice of remedies if attribution to a state is impossible. If the cyber-intrusion by the non-state actor rises to the level of a use of force or an armed attack, self-defence measures are possible against another state only when that state is deemed responsible.[22] A narrow test for attribution therefore permits impunity for states that interfere with other states' sovereignty.

Defenders of narrow attribution assert that their approach keeps conflicts within manageable levels. A test that limits a victim state's recourse to self-defence may prevent escalation of a conflict from relatively modest intrusions to full-scale war. If states cannot retaliate, the spiral of war will not get started or will be much less far-reaching. Avoiding the escalation and needless prolonging of armed conflict is a prime goal of the LOAC. Moreover, as Jinks has noted, greater accountability for states seeking to interfere with others could have counterproductive consequences when measured against other indicia such as international human rights.[23] A broader test might discourage states from

---

18 Ibid 240–1 [3] (Judge Buergenthal).

19 See, eg, Yoram Dinstein, *War, Aggression and Self-Defence* (Cambridge University Press, 4th ed, 2005) 204 (arguing that the decision was unduly narrow because Article 51 of the *Charter of the United Nations* ('*UN Charter*'), which permits self-defence against 'armed attack', does not specify that only states can commit an armed attack); Sean D Murphy, 'Self-Defense and the Israeli Wall Advisory Opinion: An *Ipse Dixit* from the ICJ?' (2005) 99 *American Journal of International Law* 62, 70–2; Ruth Wedgwood, 'The ICJ Advisory Opinion on the Israeli Security Fence and the Limits of Self-Defense' (2005) 99 *American Journal of International Law* 52, 57–9.

20 See David D Caron, 'The ILC *Articles on State Responsibility*: The Paradoxical Relationship between Form and Authority' (2002) 96 *American Journal of International Law* 857, 868. But see Part V(B) below, where I will discuss some reasons to believe that the *Manual*'s influence will be more benign.

21 See Schmitt, above n 1, 33–4 (Rule 6, [13]) (noting that a state can violate international law through conduct that interferes with another state's sovereign prerogatives); *Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda)* (*Judgment*) [2005] ICJ Rep 168, 226–7 ('*Congo*').

22 A victim state may have more limited recourse against the non-state actor operating on the territory of another state (the 'territorial state') that is unwilling or unable to control the non-state entity: see Ashley S Deeks, '"Unwilling or Unable": Toward a Normative Framework for Extraterritorial Self-Defense' (2012) 52 *Virginia Journal of International Law* 483, 499–503. Even when a victim state can take such action against a non-state actor, the principle of *ad bellum* proportionality will reduce the responses that a victim state can legally undertake when attribution to another state is lacking. For example, the general view is that the victim state may target the territorial state's forces only to the extent necessary to allow it to act against the non-state actor: see, eg, Theresa Reinold, 'State Weakness, Irregular Warfare, and the Right to Self-Defense Post-9/11' (2011) 105 *American Journal of International Law* 244.

23 See Derek Jinks, 'State Responsibility for the Acts of Private Armed Groups' (2003) 4 *Chicago Journal of International Law* 83, 91–3.

funding groups seeking to resist a tyrannical regime in another state for fear of inviting the use of force by that regime. Despotic governments could thus leverage a broader test to curb aid to rebels seeking greater freedom.

This cautious view overlooks the risks of cyber prompted by what I call the 'attribution asymmetry'. Cyber is relatively easy to direct, given a sophisticated commander, but very difficult to detect. While it is difficult to direct a group of armed personnel located hundreds or thousands of miles away from the funder of the group, an entity that wishes to control cyber-weapons can control their use from a remote location by requiring groups with state cyber-tools to submit to periodic virtual accounting. On the other hand, unlike conventional kinetic action where effects are manifest within a short time after the weapon is used, cyber-weapons can take months to detect, lying dormant for significant periods or secretly altering data to clandestinely compromise a network's operation.[24] This ability to engage in more precise direction while avoiding detection distinguishes cyber from kinetic weapons.

When coupled with a cautious approach to state responsibility, attribution asymmetry is profoundly destabilising. The narrow view of state responsibility gives the initiative to attackers, sending the message that huge numbers of cyber-intrusions are possible with impunity.[25] Ultimately, this encourages cyber-aggressive states to push the envelope. Indeed, the discretion given to attacking states makes a rapid escalation more, rather than less, likely. Moreover, the current state of the law encourages potential victim states to become attackers themselves, using thinly-veiled assistance to private groups. That process effectively outsources cyberwar, leading to a more polarised threat environment.

On the other hand, a test that would hold a state responsible for any attack that used a network within that state is also unreasonable. A large state like the United States cannot effectively monitor all cyber-traffic or deter all illegal acts over the internet by its own nationals. An overly loose standard for imputing state responsibility may encourage state disregard of rules they viewed as too unwieldy or would start needless wars.[26] Either risk is too severe to take.

To bridge this gap between unduly narrow and broad tests for state responsibility in the cyber domain, this commentary suggests a test focusing on

---

[24] For a discussion of the multistage nature of many cyber attacks, see David D Clark and Susan Landau, 'Untangling Attribution' (2011) 2 *Harvard National Security Journal* 531, 533.

[25] See Peter Margulies, 'Valor's Vices: Against a State Duty to Risk Forces in Armed Conflict' in William Banks (ed), *Counterinsurgency Law: New Directions in Asymmetric Warfare* (Oxford University Press, 2013) 87, 101 (suggesting that 'restrictions on a state's right of self-defense encourage aggression'); Michael W Lewis, 'Drones and the Boundaries of the Battlefield' (2012) 47 *Texas International Law Journal* 293, 312 (arguing that a narrow reading of geographic scope of conflict with a transnational non-state actor such as al-Qaeda 'confers a tremendous strategic advantage' on terrorist groups that violate the law of armed conflict by routinely targeting civilians); Laurie R Blank and Geoffrey S Corn, 'Losing the Forest for the Trees: Syria, Law, and the Pragmatics of Conflict Recognition' (2013) 46 *Vanderbilt Journal of Transnational Law* 693, 720–31 (arguing that a narrow interpretation of criteria such as intensity, duration and organised armed group for defining non-international armed conflicts allows states to use force with impunity against their own people).

[26] Cf Wolff Heintschel von Heinegg, 'Territorial Sovereignty and Neutrality in Cyberspace' (2013) 89 *International Law Studies* 123, 136–7 (noting that international law can only justify state obligation to prevent cyber attacks emanating from within its territory when a state has actual or constructive knowledge of such attacks).

virtual control. This test is designed to encourage states to cooperate in tracing the source of cyber attacks. Under this test, the burden shifts to a state to demonstrate it was not responsible for a cyber attack when the state funds and equips a private entity or individual who subsequently engages in a cyber attack. This test is both fairer and more efficient, since a state that has funded and equipped an individual or entity has far greater access to information than the victim state. If the control state is unwilling to cooperate with the victim state's attribution efforts, the victim may resort to other remedies, including the use of force in self-defence.

This commentary proceeds as follows. Part II discusses types of cyber attacks. Part III describes the technical difficulties of attributing a cyber attack to a specific source. Part IV discusses the law of attribution of state responsibility. It analyses the *Draft Articles*, precedent in transnational tribunals such as the ICTY and the adoption by the *Manual* of the doctrine enunciated by these sources. Part V discusses the flaws in this narrow approach. Introducing the concept of attribution asymmetry, that Part argues that while the narrow test adopted by the *Manual* is appropriate for kinetic force, it fails to fully reckon with the new challenges posed by cyberwarfare. This Part also argues that the *Manual*'s adoption of the ILC approach on attribution of state responsibility risks ossifying the law on state attribution in cyberspace. Part VI discusses the virtual control approach identified in this commentary as an alternative, including a burden-shifting mechanism. Part VII concludes with responses to potential criticisms of the virtual control approach.

II TYPES OF CYBER ATTACKS

Cyber attacks vary widely in nature, scale and scope. Among the most prominent are distributed denial of service ('DDoS') and semantic attacks.[27] This section discusses each in turn.[28]

A *DDoS Attacks*

A DDoS attack uses the power of hundreds or thousands of massed machines to impair the functioning of a particular website. Typically, an attacker will use a virus to take over control of a large number of computers that then form a botnet of 'zombie' machines. The attacker then programs the zombie computers to simultaneously log on to the targeted site. The exponential increase in traffic overwhelms the site's network, often requiring a temporary shutdown.

The best known DDoS attack occurred in Estonia in 2007. An effort by the Estonian government to remove a statue commemorating Russian participation in World War II precipitated a virulent series of DDoS attacks that compromised

---

[27] Oona A Hathaway et al, 'The Law of Cyber-Attack' (2012) 100 *California Law Review* 817, 837–9.

[28] Cyber attacks should be distinguished from so-called cyber espionage, which entails harvesting information from networks without impairing those networks' functionality. Cyber espionage may well violate a state's domestic law, but does not violate international law: see Schmitt, above n 1, 192–5 (Rule 66, [1]–[10]). See also Abraham D Sofaer, David Clark and Whitfield Diffie, 'Cyber Security and International Agreements' in National Research Council, *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for US Policy* (National Academies Press, 2010) 179, 181 (discussing the relationship between cyber espionage and cyber attacks).

the websites of government agencies, political parties, media companies and financial firms for several weeks.[29] At the start of the attack, Estonian officials identified the source of the attacks as Nashi, a Russian nationalist organisation.[30] Nashi, which was reported by journalists as being directed by the Russian government, later claimed responsibility for the attacks.[31]

B *Undermining Operating and Control Systems*

Other attacks are even more serious than DDoS attacks. Syntactic attacks use malicious computer code or malware such as 'worms, viruses, [and] Trojan horses' to compromise computer operating systems.[32] Semantic attacks, in contrast, do not destroy the computer's operating system; instead, they operate more subtly, changing the data generated by monitoring software while maintaining the illusion that the network is fully functional.[33] Semantic attacks aim to undermine control systems, such as the supervisory control and data acquisition ('SCADA') system that regulates many of a machine's moving parts.[34] SCADA systems govern the tolerances of machines such as turbines and centrifuges. Those systems can run at peak level for a limited period of time, after which they develop excess heat and begin to break down.[35] Through semantic attacks, an attacker can alter the data recorded and displayed in SCADA systems. A machine running at peak capacity and approaching the limit of its tolerance can appear to be running at a far slower speed and temperature. Because the machine's operator does not see the correct data, the machine continues running when it should have been stopped and eventually self-destructs. According to published reports, the Stuxnet virus used this method to destroy centrifuges in Iran, setting back Iran's nuclear program.[36]

III THE CHALLENGES OF TECHNICAL ATTRIBUTION

Attributing legal responsibility for cyber attacks to states is made more difficult because it is preceded by a challenging technical step: discerning the

---

29 Sheng Li, 'When Does Internet Denial Trigger the Right of Armed Self-Defense?' (2013) 38 *Yale Journal of International Law* 179, 199.

30 Ibid 203.

31 Ibid.

32 Hathaway et al, above n 27, 828, quoting Vida M Antolin-Jenkins, 'Defining the Parameters of Cyberwar Operations: Looking for Law in All the Wrong Places?' (2005) 51 *Naval Law Review* 132, 139.

33 Hathaway et al, above n 27, 828.

34 See Keith Stouffer, Joe Falco and Karen Kent, 'Guide to Supervisory Control and Date Acquisition (SCADA) and Industrial Control Systems Security: Recommendations of the National Institute of Standards and Technology' (Special Publication No 800-82 (Initial Public Draft), National Institute of Standards of Technology, September 2006) [2.1]. This was the initial public draft; for the final version, published in 2011, see Keith Stouffer, Joe Falco and Karen Scarfone, 'Guide to Industrial Control Systems (ICS) Security: Recommendations of the National Institute of Standards and Technology' (Special Publication No 800-82, National Institute of Standards and Technology, June 2011).

35 Stouffer, Falco and Scarfone, 'Guide to Industrial Control Systems', above n 34 [3.5] (noting the possibility of surreptitious changes to programming that could change 'alarm thresholds' designed to signal risky operations and thereby 'result in damage to equipment (if tolerances are exceeded)').

36 See David E Sanger, *Confront and Conceal: Obama's Secret Wars and Surprising Use of American Power* (Crown, 2012) 198–200 (although not referring to the virus by name).

actual source of the attacks. In conventional international armed conflicts involving kinetic attacks, this is not a problem. State forces typically distinguish their weapons and personnel with clear markings that identify their provenance. However, technical attribution is more difficult with cyber, even when states are involved.[37]

Identifying the source of harm is crucial for the allocation of legal consequences. As Glennon has noted, it is 'attributability … [the] ability to say "who did it" … that makes law work. When a transgressor can be identified, penalties can be assessed and retaliation and deterrence are possible — and so is legal regulation'.[38] Unfortunately, attribution in cyberspace can be as challenging as it is vital. On the internet, information typically travels in packets, which are discrete units of data outfitted with delivery instructions such as destination addresses.[39] Dedicated machines (routers) convey these packets.[40] Packets carry a source internet protocol ('IP') address, but that information is not especially useful to those seeking to verify a packet's source. Because a router's primary function is to relay a packet to a destination, routers typically do not seek to confirm that a source address is genuine.[41] Indeed, the architects of the internet viewed such confirmation as clashing with the router's role.[42]

These structural features give an advantage to a sender who wishes to conceal a particular packet's source. A recipient or a third party investigating a cyber attack may be able to discern the owner of a particular computer that happened to send a message. However, a sophisticated sender can readily make this information far less useful for identifying the source of an attack. Attacks often involve several stages. In these multistage attacks, the attacker will commandeer one computer with code that converts that computer into a platform for attacking a second computer.[43] Senders can 'spoof' other IP addresses so that a computer that originally sent the message is disguised as another machine. In DDoS attacks, the attacker may harness hundreds or thousands of computers, making identification of the original sending machine extraordinarily difficult. For the

---

37 Hathaway et al, above n 27, 856.

38 Michael J Glennon, 'The Road Ahead: Gaps, Leaks and Drips' (2013) 89 *International Law Studies* 362, 380. See also Clark and Landau, above n 24, 532 (noting that '[a]ttribution is central to *deterrence*' (emphasis in original)).

39 See Clark and Landau, above n 24, 534; Melissa E Hathaway and John E Savage, 'Stewardship of Cyberspace: Duties for Internet Service Providers' (Paper presented at Cyber Dialogue 2012: What is Stewardship in Cyberspace?, University of Toronto, 18–19 March 2012) 3 <http://www.cyberdialogue.citizenlab.org/wp-content/uploads/2012/2012papers/CyberDialogue2012_hathaway-savage.pdf>; Adam Candeub and Daniel McCartney, 'Law and the Open Internet' (2012) 64 *Federal Communications Law Journal* 493, 497. For a discussion of the difference between tracing the proximate versus ultimate source of cyber intrusion, see Herbert S Lin, 'Offensive Cyber Operations and the Use of Force' (2010) 4 *Journal of National Security Law & Policy* 63, 77–8.

40 Clark and Landau, above n 24, 534.

41 Ibid 534–5.

42 For further discussions on the role of routers, see Erik M Mudrinich, 'Cyber 3.0: The Department of Defense Strategy for Operating in Cyberspace and the Attribution Problem' (2012) 68 *Air Force Law Review* 167, 177 (explaining that the internet's architecture is designed first and foremost to facilitate the conveyance of packets to a destination, 'whether or not the … packets are recognized by the network'). See also Duncan B Hollis, 'An e-SOS for Cyberspace' (2011) 52 *Harvard International Law Journal* 374, 378 (noting that 'the very architecture of the Internet enables hackers to maintain anonymity if they so desire').

43 Clark and Landau, above n 24, 533. See also Lin, above n 39, 78–9.

same reason, the physical location of a sending machine can be unrevealing.[44] A sender may be located in country W, but may act on behalf of country X and route a malicious packet through servers in country Y in the course of attacking country Z.[45]

That said, we may be making progress on technical indicia of attribution.[46] In other realms of forensics, analysts look for signatures. Details of the crime scene can tell an expert a great deal about the method, modus operandi and 'playbook' of a violent criminal. Criminals use particular weapons and ways of stalking a victim. They may also use particular methods for trying to keep their work secret. Although digital forensics often lacks the physical details that aid in a conventional forensic investigation, many behavioural and technological cues can aid in attribution. For example, just as people in certain regions have identifiable accents and cultural baselines, digital actors may have certain internet platforms that they favour. Particular kinds of spoofing may be favoured by hackers from a certain country or region. These cues have become vital in police work and will undoubtedly become increasingly important in the digital domain.

IV CYBER ATTACKS, KINETIC ATTACKS AND INTERNATIONAL LAW ON STATE RESPONSIBILITY

Like some cyber attacks, attribution is a multistage process. Even if we have resolved the difficult technical question of forensic attribution, we must then consider the legal question of when an attack traced to a particular source within a state can be attributed to the state itself. That legal attribution is vital for international law, which governs state responsibility for harm to other states. With the guidance provided by international law, we can determine when states are responsible for harms committed by ostensibly private actors and what remedies are available to the victim state.

---

44 Clark and Landau, above n 24, 548 (noting that locational mapping of internet protocol addresses is 'approximate' and therefore yields 'plausible deniability' for alleged source).

45 See Schmitt, above n 1, 33 (Rule 6, [12]) (posing a hypothetical in which a group in State A enlists computers in State B for an attack on State C's network); Todd C Huntley, 'Controlling the Use of Force in Cyber Space: The Application of the Law of Armed Conflict during a Time of Fundamental Change in the Nature of Warfare' (2010) 60 *Naval Law Review* 1, 12 (noting that identifying 'general geographic location' of machines conveying malware may not be helpful for attribution, since packets conveying malware may have been routed through that location from somewhere else).

46 See Department of Defense, 'Department of Defense Cyberspace Policy Report: A Report to Congress pursuant to the *National Defense Authorization Act* for Fiscal Year 2011, Section 934' (Policy Report, November 2011) 4, <http://www.defense.gov/home/features/2011/0411_cyberstrategy/docs/NDAA%20Section%20934%20Report_For%20webpage.pdf> (noting that the Department of Defense 'seeks to increase … attribution capabilities by … developing new ways to trace the physical source of an attack, and seeking to assess the identity of the attacker via behavior-based algorithms'); Matthew C Waxman, 'Self-Defensive Force against Cyber Attacks: Legal, Strategic and Political Dimensions' (2013) 89 *International Law Studies* 109, 119 (asserting that 'attribution challenges may be overstated, especially for the United States and its premier intelligence and cyber-forensic capabilities'); Shane McGee, Randy V Sabett and Anand Shah, 'Adequate Attribution: A Framework for Developing a National Policy for Private Sector Use of Active Defense' (2013) 8 *Journal of Business & Technology Law* 1, 28–32 (discussing approach to attribution that combines technical analysis with inferences drawn from social and political analysis and intelligence gathering).

A *International Law: The US Position, the* Draft Articles *and the Cases*

Most scholars and states agree that international law governs state responsibility in the cyber domain. For example, the US has noted that 'the development of norms for state conduct in cyberspace does not require a reinvention of customary international law'.[47] According to the US, '[l]ong-standing international norms guiding state behavior — in times of peace and conflict — also apply in cyberspace'.[48] However, the US has also observed that the interpretation of international law must accommodate the 'unique attributes of networked technology'.[49] According to the US, those special traits might require 'additional understandings' to 'supplement' traditional international norms.[50]

To avoid interference with sovereign prerogatives, international law has interpreted state responsibility narrowly in the domain of conventional or kinetic attacks.[51] One vital aspect of sovereignty is the state's reliance on officials chosen according to the state's own rules. To vindicate that reliance, international law holds that only decisions by officials can bind the state.[52] Sovereign states therefore bear responsibility only for acts and omissions that a reasonable observer can trace to state officials. The ILC summarised this principle, noting that 'the general rule is that the only conduct attributed to the State … is that of its organs of government, or of others who have acted under the direction, instigation or control of those organs, ie, as agents of the State'.[53] According to the ILC, a party seeking to attribute the actions of a private group to a state must show the 'existence of a real link between the … group

---

47 White House, 'International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World' (May 2011) 9 <http://www.whitehouse.gov/sites/default/files/rss_viewer/international_strategy_for_cyberspace.pdf>; 'US Cyber Operations Policy' (Presidential Policy Directive/PPD-20, October 2012) 4, reproduced in 'Obama Tells Intelligence Chief to Draw Up Cyber Target List — Full Document Text', *The Guardian* (online), 8 June 2012 <http://www.guardian.co.uk/world/interactive/2013/jun/07/obama-cyber-directive-full-text> (stating policy that the US shall conduct defensive and offensive cyber operations in a fashion that is 'consistent with its obligations under international law'). Cf Lotrionte, above n 3, 834–6 (describing US policy).

48 White House, above n 47, 9.

49 Ibid.

50 Ibid.

51 See *Draft Articles*, UN Doc A/56/10, ch IV(E).

52 See Reuven Young, 'Defining Terrorism: The Evolution of Terrorism as a Legal Concept in International Law and Its Influence on Definitions in Domestic Legislation' (2006) 29 *Boston College International and Comparative Law Review* 23, 62 (noting fairness of linking state responsibility to acts or omissions of 'individuals sufficiently connected to a state').

53 *Draft Articles*, UN Doc A/56/10, ch IV(E) 80 (commentary to Chapter 2: Attribution of conduct to a State, [2]).

performing the act and the State machinery'.[54] Any other rule would disrupt the state's chosen form of government. No government could act effectively if the random or isolated acts of its citizens were automatically attributable to the state. A state labouring under this burden would spend most of its time responding to the acts of individuals instead of forging state policy. To guard against this paralysis of sovereign prerogatives, international law holds that 'the conduct of private persons is not as such attributable to the State'.[55]

Reflecting this wariness about individuals binding the state, art 8 of the *Draft Articles* takes a narrow view of state responsibility for private actors, noting that conduct of a person or entity is attributable to the state when a 'group … is in fact acting on the instructions of, or under the direction or control of, that State in carrying out the conduct'.[56] State 'instructions' are express orders to engage in a particular operation in a particular way.[57] The two situations noted by the *Draft Articles* involving state 'direction or control' are more 'general'.[58] However, both direction and control still require evidence of a high level of state involvement in a particular operation carried out by a private party. For example, according to the ILC, a state that has not instructed a group to engage in an attack or some other violation of international law is answerable for the group's conduct 'only if it directed or controlled the specific operation and the conduct complained of was an integral part of that operation'.[59]

Transnational tribunals have tended to interpret the terms 'direction or control' narrowly. Consider, for example, the ICJ's decision in *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)* ('*Nicaragua*'),[60] which limited US responsibility for abuses committed by a rebel group, the Contras, funded and equipped by the US; or the ICJ's decision in *Application of the* Convention on the Prevention and Punishment of the Crime of Genocide *(Bosnia and Herzegovina v Serbia and Montenegro)*,[61] which declined to find Serbia guilty of genocide for massacres committed by a private paramilitary group. In both cases, the ICJ couched these terms in its formulation of what it called an 'effective control' test. While to American ears 'effective control' may connote practical control, the ICJ's use of the term

---

54 Ibid ch IV(E) 104 (commentary to art 8, [1]). From its inception, international law has sought to discourage the conflation of individual and state agendas. The Enlightenment publicist de Vattel opined that international law refused to recognise certain acts, including acts of violence against another state, as those of lawful belligerents: Emmer de Vattel, *The Law of Nations or the Principles of Natural Law* (Liberty Fund, 2008 ed) vol 3, 542–3 [trans of: *Le Droit des gens, ou, principles de la loi naturelle appliqués à la conduit et aux affaires des nations et des souverains* (first published 1758)]. Cf Kenneth Watkin, 'Warriors without Rights? Combatants, Unprivileged Belligerents, and the Struggle over Legitimacy' (Occasional Paper No 2, Program on Humanitarian Policy and Conflict Research, Harvard University, 2005) 13–16 (discussing the traditional requirement of law of war that individuals participating in hostilities possess 'right authority', defined as state consent and approval, in order to be regarded as 'combatants').

55 *Draft Articles*, UN Doc A/56/10, ch IV(E) 81 (commentary to ch 2: Attribution of Conduct to a State, [3]).

56 Ibid ch IV(E) art 8.

57 See Ibid ch IV(E) 104 (commentary to art 8, [1]).

58 Ibid.

59 Ibid (commentary to art 8, [3]).

60 *Nicaragua* [1986] ICJ Rep 14.

61 *Genocide Case* [2007] ICJ Rep 42.

requires something closer to specific, comprehensive control. The clearest case is one in which the state has specifically instructed private groups to engage in a cyber attack.[62] Arming and training a private group engaged in violence may give rise to state responsibility for a use of force that violates art 2(4) of the *UN Charter*.[63] However, mere funding is insufficient,[64] even when a state has provided 'heavy subsidies' to the group.[65]

The ICTY in *Prosecutor v Tadić* ('*Tadić*') announced a broader standard hinging on 'overall control' by state officials.[66] However, the 'overall control' test announced in *Tadić* is also quite demanding. As the ICTY noted, 'overall control' must be more than the 'mere financing and equipping of such forces'.[67] Instead, it entails 'coordinating or helping in the general planning of [the group's] military activity'.[68]

B *The* Manual*'s Take*

The *Manual* largely imports this restrictive language from the ILC and the case law. It takes the *Draft Articles* as a touchstone and cites both the 'effective' and 'overall' control tests.[69] Tellingly, it does not cite the language from *Tadić* cited above, which describes general helping behaviour as meeting the overall control test.[70] This language hinted at a broader, more flexible standard for state responsibility. Instead, the *Manual*'s drafters included other language that was more rigid in tone, if not substance, in which the ICTY opined that a finding of state responsibility required official 'participation in the planning and supervision of military operations'.[71] On this view, a state would not share responsibility under international criminal law for harm a private group causes in cyber-activities unless the state did more than finance and equip the group.

The *Manual* acknowledges that 'providing an organized [armed] group with malware and the training necessary to use it to carry out cyber attacks' would constitute a use of force by the helping state that would violate art 2(4) of the *UN Charter*.[72] However, the *Manual* follows the ICJ in observing that the provision of financial aid, without more, does not constitute a use of force.[73] Moreover, the *Manual* draws the same conclusion regarding a state's provision of sanctuary or safe haven to a private group engaged in cyber attacks.[74]

---

62 See *Genocide Case* [2007] ICJ Rep 42, 208 [400], 210 [406].

63 *Nicaragua* [1986] ICJ Rep 14, 118–19 [228]. See also Schmitt, above n 1, 46 (Rule 11, [3]) (discussing the standard).

64 *Draft Articles*, UN Doc A/56/10, ch IV(E) 105–6 (commentary to art 8, [4]); *Nicaragua* [1986] ICJ Rep 14, 118–19 [228].

65 *Nicaragua* [1986] ICJ Rep 14, 62 [109].

66 See *Tadić* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [131], [145].

67 Ibid [145].

68 Ibid [131].

69 Schmitt, above n 1, 32–4 (Rule 6, [9]–[13]).

70 See *Tadić* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [131].

71 Ibid [145], quoted in Schmitt, above n 1, 33 (Rule 6, [10]).

72 Schmitt, above n 1, 46 (Rule 11, [4]), citing *Nicaragua* [1986] ICJ Rep 14, 118–19 [228].

73 Schmitt, above n 1, 46 (Rule 11, [3]), citing *Nicaragua* [1986] ICJ Rep 14, 118–19 [228].

74 Schmitt, above n 1, 46–7 (Rule 11 [5]).

The caution of the *Manual* in imposing responsibility on states is most striking in the one scenario it addresses where traditional LOAC principles would support an inference of state responsibility: the use of state infrastructure for an attack. Prior to the cyber age, a reasonable observer would have drawn a far stronger inference of state involvement on the theory that it would have been difficult for a non-state actor to gain access to state weapons without that state's authorisation.[75] The *Manual* departs from this common sense intuition, asserting that the staging of an attack using government infrastructure is merely 'an indication' that the state is 'associated with the operation'.[76]

Another question concerns whether cyber aid by one state to rebels in another state transforms a non-international armed conflict ('NIAC') into an international armed conflict ('IAC'). Under international law, mere tangible aid, such as the provision of weapons, would not internationalise the conflict; a finding of overall state control of a private group's activities is necessary.[77] The *Manual* imports this traditional view into cyberspace, finding that a state's knowing provision of 'cyber attack tools' to rebel forces would not change a NIAC into an IAC.[78] However, the *Manual* concedes that this threshold would be crossed if a state provided operational intelligence, such as information on another state's specific 'cyber vulnerabilities'.[79]

V THE DIFFERENCE CYBER MAKES: CRITIQUING THE *MANUAL*

The *Manual*'s range and richness mask some flaws, particularly in the handling of state aid to non-state actors engaging in cyber attacks. First, the *Manual* fails to address at least one feature of the *Nicaragua* decision that gestures toward a broader definition of state control. Secondly, the *Manual*, by relying so heavily on the *Draft Articles*, fails to address the full impact of the attacks of 11 September 2001 and also risks the same deleterious impact that the *Draft Articles* have had on state practice. Thirdly, in one important realm, the *Manual* relies on overly fine distinctions that fail to provide sound guidance. Fourthly, the *Manual* is actually more cautious than traditional LOAC rules for kinetic warfare on the attribution of attacks from state infrastructure. Fifthly, the *Manual* fails to reckon with the special attributes of cyber that make analogies to kinetic force treacherous. This section addresses each point in turn.

A *Neglecting a Broader Reading of International Case Law*

The *Manual* fails to acknowledge aspects of the *Nicaragua* and *Tadić* decisions that might have supported a broader reading of state responsibility. One fact highlighted in the *Nicaragua* decision hinted that the tribunal was actually treating state funding of a private group as shifting the burden to a state to demonstrate lack of control. The ICJ noted that while aid ended in October 1984, the Contras continued to fight.[80] Although the ICJ did not suggest that it

---

75 Ibid 35 (Rule 7 [3]).
76 Ibid 34 (Rule 7).
77 *Tadić* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [145].
78 Schmitt, above n 1, 81 (Rule 22, [6]).
79 Ibid.
80 *Nicaragua* [1986] ICJ Rep 14, 62 [110].

was employing a burden-shifting approach, its reliance on the Contras' continued resistance after the termination of US aid is consistent with the tribunal viewing aid as a burden-shifting device, with the presumption of state control rebutted by resistance after the aid's cessation. In *Tadić*, although the ICTY cited a range of shared elements between the Republic of Yugoslavia and a genocidal paramilitary group, the tribunal's textual discussion of those links highlighted the Republic of Yugoslavia's payment of the group's salaries.[81] The *Manual* could have cited these passages to temper its reliance on the pat test announced in the *Draft Articles*.[82]

B *The* Draft Articles*' Siren Song*

The *Manual*'s reliance on the *Draft Articles* also triggers scepticism because the *Draft Articles* shut off development of the law of state responsibility.[83] As Caron has noted, the ILC's codification effort risked making the law unduly rigid.[84] The comprehensive scope and authoritative tone of the *Draft Articles* may have triggered more deference than the ILC's work-product merited, shutting down evolution based on state practice.[85] Importing the ILC's recommendations into the cyber realm, with its exceptionally fluid character, has compounded this risk of ossification.[86]

That disparity between the ILC's codified work-product and the exigencies of world affairs became particularly glaring after September 11. The *Draft Articles*' narrow criteria sent a troubling policy message: in interpreting state responsibility so restrictively, the *Draft Articles* unduly discounted the need for due diligence in monitoring non-state actors' use of state aid. In the wake of the September 11 attacks, international law paid greater attention to states' obligations to disrupt networks of violent non-state actors operating from their territory, including groups actually receiving state assistance.[87] Because of the sea change occasioned by the September 11 attacks, the *Manual*'s reliance on the ILC's approach is particularly risky.

---

81 *Tadić* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-A, 15 July 1999) [150]. Cf at [150] n 180 (noting the presence of other factors including shared communications, logistics and tactics).

82 See *Draft Articles*, UN Doc A/56/10, ch IV(E) art 8.

83 Caron, above n 20, 868.

84 Ibid 860 (observing that '[c]odification brings clarity to the law … [b]ut it can also inject unwelcome rigidity').

85 See ibid 861 (describing codification of the law of state responsibility as 'a risky proposition').

86 Ibid 859 (asserting that, 'even as an area of doctrine is codified, the world it was intended to address may move on to a new form that tests the structure of the previous order').

87 See David E Graham, 'Cyber Threats and the Law of War' (2010) 4 *Journal of National Security Law & Policy* 87, 96 (arguing that 'evolving consensus regarding the establishment of a new standard for imputed state responsibility solidified following … September 11'). See also *Congo* [2005] ICJ Rep 168, [160]–[162] (finding that the state did not exert control over operations of the private group, so that responsibility for the group's attacks could not be attributed to the state, and also holding that the state's aid to the group violated international law). But see Jinks, above n 23, 90–3 (warning of risks associated with heightened state responsibility for private actors). Cf Vincent-Joël Proulx, 'Babysitting Terrorists: Should States Be Strictly Liable for Failing to Prevent Transborder Attacks?' (2005) 23 *Berkeley Journal of International Law* 615, 637–41 (suggesting that even a state's indirect responsibility for terrorist attacks can trigger a right of the victim state to use force to prevent further attacks).

The *Manual*'s approach is sufficiently different from the *Draft Articles* to make this concern less compelling. While the ILC laboured arduously for decades, the *Manual* was completed within a very tight time frame, making the scope of its analysis all the more impressive. Because international experts working on the *Manual* did their job so efficiently, state officials were not forced to sit on their hands, glumly awaiting the experts' handiwork. Just as importantly, the experts drafting the *Manual* signalled repeatedly that they did not expect their product to be the last word on the subject. The efficiency of the Tallinn project and the dialogue in its commentary about divergent views will limit ossification. However, ossification is always a concern. Moreover, the *Manual*'s analysis of state responsibility buys into the ILC approach without the qualifications that characterise the *Manual*'s commentary on other areas, such as the definition of a use of force or armed attack in the cyber realm.[88] Therefore, at least as regards the law of state responsibility for cyber-intrusions, some of the fears that commentators have voiced about the *Draft Articles* are also appropriate in evaluating the *Manual*. More debate in the *Manual*'s comments on principles of state responsibility would have dissipated this concern.

C *The Wages of Unduly Fine Distinctions*

The *Manual* relies on overly fine distinctions in its discussion of when a state role in a NIAC can internationalise that conflict, making it a conflict between nations. For example, consider the *Manual*'s distinction, discussed above, between an assisting state's provision of 'cyber attack tools' to rebels — which would not internationalise a conflict — and its provision of 'specific intelligence on cyber vulnerabilities' — which would create an international armed conflict between the assisting and victim states.[89]

This is a welcome adaptation of international law to cyber, but the *Manual*'s advice is less clear than it should be. In practice, there is very little difference between provision of cyber-tools to groups and instruction in how to use those tools to exploit the vulnerabilities of an adversary. Provision of tools will inevitably meld into instruction on their use. Parsing this distinction, state officials assisting rebels in another state might use hypotheticals that studiously avoided identifying information about an adversary's specific internet infrastructure but nonetheless furnished rebels with all they needed to know. The *Manual*'s drafters should have avoided fine distinctions that promote such disingenuous tactics.

The *Manual* makes a similarly fine distinction on when shutting down a victim state's cyber-network would internationalise a conflict. According to the *Manual*, action by an assisting state that would 'shut down State B's cyber-communications capabilities' would only internationalise a conflict if State

---

[88] See Schmitt, above n 1, 57–8 (Rule 13, [13]) (noting disagreement among the *Manual*'s drafters on whether the Stuxnet episode constituted an armed attack). See also Gary D Brown, 'Why Iran Didn't Admit Stuxnet Was an Attack' (2011) 63 *Joint Force Quarterly* 70, 71 (stating that Stuxnet clearly amounted to an attack). Cf Sean Watts, 'Low-Intensity Computer Network Attack and Self-Defense' (2011) 87 *International Law Studies* 59, 75 (observing that state practice may ultimately come to reflect the view that even low-intensity cyber attacks such as DDoS exploits trigger a state's right of self-defence).

[89] See Schmitt, above n 1, 81 (Rule 22, [6]).

B 'relies' on the system for military use.[90] Action directed at civilian uses of the network would not internationalise the conflict. This distinction seems highly artificial. There may be significant interoperability between military and civilian communications capabilities. For example, a state's government may include different branches, such as a legislative and an executive branch, that communicate in part through a network. These civilian communications may shape the strategic or tactical directives passed down through the military chain of command.[91] Shutting down such civilian communications will affect the military, whether or not the military 'relies' on the system within the *Manual*'s definition. A rule that distinguishes between civilian networks and those that the military relies on merely invites confusion on the part of both attackers and defenders.[92]

D *The Difference Cyber Makes*

The *Manual*'s resort to fine distinctions in particular areas, such as the difference between the provision of cyber-tools and the provision of operational data, contrasts with the Manual's overarching project of 'normalising' cyber within the LOAC framework. This section will argue that the normalisation goal obscures vital differences between cyber and other kinds of attacks. However, before reaching this point, one should note another possible source of the *Manual*'s embrace of the *Draft Articles*' approach to state responsibility. The *Manual*'s drafters worried that difficulties in technical attribution of cyber attacks would lead to mistaken judgments by states. In the one area — responsibility for attacks originating with state infrastructure — where the normalisation goal clashed with concern about technical attribution, the *Manual* mistakenly prioritises the latter concern. This section first critiques the *Manual*'s approach to the state infrastructure point and then pivots to a critique of the *Manual*'s more typical approach to state responsibility: analogising cyber to kinetic attacks.

1 *Critiquing the* Manual*'s Caution on Use of State Infrastructure for Cyber Attacks*

The clarity that restatements promise was imperative on one crucial issue: responsibility for cyber attacks staged using state infrastructure. Viewed *ex ante*, avoiding unnecessary conflicts requires that states receive the strongest possible message on the need for due diligence in ensuring that state infrastructure not be commandeered by private groups committing cyber attacks. Instead, the *Manual*'s drafters opted for a message that is thin gruel: according to the *Manual*, the initiation of an attack using government infrastructure is merely 'an indication' that the state is 'associated with the operation'.[93] As the *Manual*'s

---

[90] Ibid 81 (Rule 22, [7]).

[91] Rule 39 of the *Manual* discusses the issue of overlapping civilian and military networks: see ibid 134 [1].

[92] The *Manual*'s drafters agreed that an attack on a civilian network that prompted civilian casualties would be an armed attack and thus constitute an independent basis for an international armed conflict: ibid 82–3 [12] (discussing criteria for defining an 'armed attack').

[93] Ibid 34 (Rule 7).

authors acknowledge, traditional principles would have drawn a far more robust inference of state involvement based on the position that a non-state actor could not gain access to state weapons without that state's consent.[94] The *Manual*'s authors assert that cyber is different, because it is more likely that government infrastructure could have been taken over by non-state actors without state authorisation.[95] This is right as far as it goes. However, in fashioning a rule, the central goal should be imposing the costs of compliance on the party best able to efficiently bear those costs. States are best able to monitor their own networks, detect unauthorised activity and explain breaches to others, including international bodies and victims of attacks launched from state infrastructure.[96] Treating use of a state's network for a cyber attacks as merely 'an indication' of state responsibility allows states to cut corners on each of these central tasks. That is precisely the wrong message to impart.

2 *Distinguishing Cyber from Kinetic Attacks: Attribution Asymmetry*

In other areas, the *Manual*'s normalisation goal and its concerns about technical attribution dovetail to result in adherence to the *Draft Articles* approach. This approach fails to recognise that cyber attacks pose particular threats that undermine analogies to the law of state responsibility for kinetic attacks. Cyber is not necessarily unique and the *Manual*'s drafters are absolutely correct that cyber can fit into the framework of LOAC. However, the special character of cyber attacks calls for some tweaking in the law of state responsibility. Cyber's differences fall under what I call attribution asymmetry.

The asymmetry arises because of material differences between cyber and kinetic attacks. Cyber attacks, particularly those of the most sophisticated variety, are difficult to detect. Paradoxically, cyber attacks are also easy to direct for individuals or entities who initiate them. Moreover, cyber attacks require far less in the way of personnel. This subsection addresses each point in turn.

First, kinetic means are readily detectable. By definition, kinetic means have an immediate impact, entailing physical harm to persons and/or property. Because a kinetic attack typically involves visible harm and triggers a range of visible and audible rescue and response operations, an attacking state will not be able to keep the impact of a kinetic attack secret for long. Indeed, while the planning for a kinetic attack may well be conducted in secret in the hopes of surprising an adversary, the reasonable attacker will plan an attack with the understanding that any element of surprise dissipates rapidly once the operation is under way.

Cyber attacks are different. As noted when we discussed technical attribution, cyber attacks can proceed in stages over time, with the initial stages being latent. For example, a cyber attacker could use a covert agent[97] with a thumb drive to

---

94 Ibid 35 [3]. One of the *Manual*'s drafters has reaffirmed the relevance of this presumption to cyber attacks: see Heintschel von Heinegg, above n 26

95 Schmitt, above n 1, 35 (Rule 7, [3]). See also Droege, above n 5, 543–4 (arguing against a presumption of state responsibility in use of state infrastructure for a cyber attack).

96 See Lotrionte, above n 3, 915 (discussing the informational advantage enjoyed by states).

97 For more analysis of the different domestic legal requirements in the US for covert action as compared with traditional military action and how the two domains have converged after September 11, see Robert Chesney, 'Military-Intelligence Convergence and the Law of the Title 10/Title 50 Debate' (2012) 5 *Journal of National Security Law & Policy* 539, 544–70.

install malware on a network that would become fully operational only when a triggering event occurred, such as a certain volume of trades on a national stock exchange. The whole point of this kind of cyber attack is to install software secretly in one phase that will lead to damage later.

Moreover, compared with conventional and symmetrical uses of kinetic force, cyber attacks require far less in the way of personnel. Conventional massed forces are ineffective without thousands of people participating in a given attack, providing logistical support and maintaining weapons systems. Asymmetric attacks by violent non-state actors do not require the same numbers because small numbers of attackers can inflict significant civilian casualties. Cyber attacks require fewer attackers still. A resourceful and creative individual with modest help from a small group can, with the right software, take over a significant number of computers and command them to forward phishing messages or malware.[98] The determined hacker can use tacit support from others, including the larger hacker community, that would be very difficult to obtain for the use of kinetic force. The ability to keep participation a secret in the cyber realm aids recruiting, just as the difficulty of keeping a secret in the kinetic world inhibits participation in that realm.

Asymmetries also exist between cyber and kinetic means on the ease of supervision. Supervising kinetic means employed by non-state actors is difficult, especially at a distance. As the US Supreme Court has noted, a state has only limited ability to monitor events in a foreign country.[99] Because of this difficulty, it may well seem unfair to require that a state that funds a non-state group located overseas accept responsibility for the harm caused by the group's use of kinetic means.

However, geographic proximity has little relevance to the cyber domain. In the cyber realm, one party to a conflict can affect operations in a state thousands of miles away[100] and do so without the threat of detection that characterises kinetic operations. Moreover, a state has a far greater capacity to control such groups. It therefore seems entirely fair to hold the state to a higher level of responsibility.

Cyber is also relatively easy to direct, given a sophisticated operator. Indeed, the US already requires that internet service providers ('ISPs') maintain and produce data about customers such as file-sharing sites that offer third parties access to copyrighted material.[101] The *Digital Millennium Copyright Act* ('*DMCA*') provides for a safe harbour for ISPs that make expeditious efforts upon receipt of information regarding infringing activity to stop or prevent that

---

98 See Mudrinich, above n 42, 169–70.

99 See *Holder v Humanitarian Law Project*, 130 S Ct 2705, 2727 (2010) (observing that 'national security and foreign policy concerns arise in connection with efforts to confront evolving threats … where information can be difficult to obtain and the impact of certain conduct difficult to assess'). See also Peter Margulies, 'Advising Terrorism: Material Support, Safe Harbors, and Freedom of Speech' (2012) 63 *Hastings Law Journal* 455.

100 See Chris C Demchak and Peter Dombrowski, 'Rise of a Cybered Westphalian Age' (2011) 5(1) *Strategic Studies Quarterly* 32, 33 (observing that the Stuxnet virus demonstrated cyber's 'ability to deliver a potentially killing blow without being anywhere near the target').

101 See Clark and Landau, above n 24, 552; *Recording Industry Association of America Inc v Verizon Internet Services Inc*, 351 F 3d 1229 (DC Cir, 2003) ('*Verizon*') (analysing the process for identifying websites).

activity.[102] Consistent with the regime, a state could identify particular ISPs that grantees were required to use for the storage of data, including temporary functions such as the cache of a web page or more permanent functions such as the ongoing storage of websites or the hosting of information-locating or retrieval tools.[103] A state could mandate comparable due diligence elsewhere in the cyber domain. For example, states promoting due diligence have particular leverage over their own grantees. A state could mandate that grantees disable 'anti-attribution' devices that conceal IP addresses and permit users to visit websites using pseudonyms.[104]

E *Summary*

In sum, the *Manual*'s account of state responsibility and cyber attacks raises more questions than it answers. The *Draft Articles* may not be an accurate account of the law after September 11. Moreover, even before September 11, factual wrinkles in the leading cases of *Nicaragua* and *Tadić* may have cut against an unduly narrow reading of the effective and overall control tests. In addition, because of attribution asymmetry, the law governing kinetic armed conflicts may contribute little guidance to the law applicable to state responsibility for cyber attacks. Even if we accept the effective and overall control tests as *lex lata* regarding the kinetic realm, the status of those tests in cyberspace is open to serious question. There may be no reliable *lex lata* to govern this fluid domain. Filling the gap is an urgent need.

VI AN ALTERNATIVE TO THE *MANUAL*: THE VIRTUAL CONTROL APPROACH

Because of attribution asymmetry, the traditional tests of overall and effective control do not work well in the cyber domain. Another test is needed. This commentary calls that alternative the test of virtual control.

A *Virtual Control Defined*

Virtual control recognises that the difficulty of detection, the diminished need for personnel and the ease of monitoring by a diligent state require burden shifting when a state funds and equips or knowingly provides sanctuary to a private entity that subsequently engages in a cyber attack against another state. Under this approach, the victimised state can demand further information from the state in virtual control. Information produced by the state in virtual control may show that it was in fact uninvolved in the attack or was unable to control the individual or entity responsible for the attack. Alternatively, that information may show the virtual controller's responsibility. If the control state is unwilling to provide the information, the victim may pursue other remedies, including the

---

[102] *Digital Millennium Copyright Act*, 17 USC §§ 512(c)(1)(A)–(C) (1998) ('*DMCA*'). For a critique of other *DMCA* provisions concerning liability for prohibited activities, see Zoe Argento, 'What the *Digital Millennium Copyright Act* Can Learn from Medical Marijuana: Fixing the Antitrafficking Provisions by Basing Liability on the Likelihood of Harm' (2012) 35 *Columbia Journal of Law & the Arts* 503, 506–11 (discussing *DMCA* provisions).

[103] *Verizon*, 351 F 3d 1229, 1234 (DC Cir, 2003) (noting conditions imposed by US law for internet service providers transmitting material copyrighted by others).

[104] For a discussion on such 'anti-attribution' devices, see Clark and Landau, above n 24, 545–6.

use of force in self-defence if a cyber attack has crossed the threshold into an armed attack under art 51 of the *UN Charter*.[105]

Furthermore, a virtual control approach would shift the burden in the one area where the *Manual* is more cautious than kinetic-based LOAC — the use of state infrastructure for an attack. Launching an attack from state infrastructure should create a rebuttable presumption that the state owning the infrastructure is responsible. That responsibility acts as an incentive for a state, which is the cheapest cost avoider, to invest in the best, most current anti-hacking measures, including software that can trace the sources of cyber-intrusions. A state that does so will have an opportunity to ascertain who is to blame in the event of an attempt by non-state actors to take over its infrastructure. That deflecting of blame, based on acceptable forensic evidence, would rebut the presumption of state responsibility, leaving state responsibility governed by the overall control test. However, if the state could not demonstrate the presence of adequate digital controls, its failure to meet its burden would result in a risk of sanctions from international bodies as well as the possibility of the use of force in self-defence by the victim state.

Another question concerns whether cyber-aid by one state to rebels in another state constitutes an IAC. Under international law, tangible aid such as weapons for kinetic attacks would not internationalise the conflict.[106] However, although the *Manual* is as conservative on the existence of an IAC as it is on state attribution per se, there are reasons for regarding the interference connoted by the provision of cyber-tools as sufficiently serious to justify finding an IAC.

Here, attribution asymmetry provides insight. Cyber-weapons can disable a financial system or a state's critical infrastructure. Providing a rebel group with tools to do this, even without instructing them to do so or providing specific advice, is sufficiently serious to internationalise a conflict. Certain cyber-tools are after all far more interactive in nature than conventional weapons. Instruction in weaponry typically involves pointing the weapon in a particular direction or activating a precision-guidance system. Cyber-tools, in contrast, can be provided only with guidance on how to use those weapons to interact with pre-existing cyber-architecture. Troubleshooting will involve not merely maintaining those tools in good working order, but consulting with the group on particular problems that one would expect to find when navigating through another state's cyber-system. At this point, the difference between 'the provision of cyber attack tools' and 'providing specific intelligence on cyber vulnerabilities'[107] becomes vanishingly small. The virtual control approach recognises this functional link between attack tools and specific operations.

B *Implementing the Virtual Control Approach*

To see how the virtual control approach would work in practice, consider the following example. Suppose that Utopia was the victim of a cyber attack that required the reinstallation of operating systems on tens of thousands of machines,

---

105 Cf Terry D Gill and Paul A L Ducheine, 'Anticipatory Self-Defense in the Cyber Context' (2013) 89 *International Law Studies* 438, 452–8 (discussing criteria for using force in self-defence against anticipated cyber attacks).

106 Schmitt, above n 1, 81 (Rule 22, [5]–[6]).

107 Ibid 81 (Rule 22, [6]).

including those in Utopia's central bank. The attack resulted in millions of dollars in losses, measured by the lost profits on bank transactions, the need to purchase new software and the thousands of hours logged by Utopia's information technology personnel. Assume that the loss of functionality in Utopia's machines was an effect that rose to the armed attack level.[108] After a sophisticated digital forensics investigation, Utopian officials concluded that the attack originated from an IP address assigned to the Ruritanian Resistance Group ('RRG'). The RRG had attempted to spoof another IP address. Initial intelligence reports suggested that the RRG received funding and software from Ruritania. Ruritania's assistance to the RRG therefore met the 'virtual control' standard outlined here.

Under the proposal described, Utopia would first have to request cooperation from Ruritania. Ruritania would bear the burden of proof on either of two possible defences. First, it could demonstrate that it did not fund or equip the RRG. Secondly, it could show that although it had provided such assistance, it had employed due diligence that had prevented the RRG from using this aid in its cyber-intrusion on Utopia.

Suppose Ruritania was unable to meet its burden on either defence. In that event, assuming that Utopia's attribution process was reasonably reliable, Utopia could treat the cyber-intrusion as an armed attack and respond in self-defence. If Ruritania sought relief in the ICJ or the United Nations Security Council, it would bear the burden of proving that it was not responsible for the initial attack on Utopia.

This framework encourages states to better monitor and control the activities of recipients of their aid and also aligns incentives with access to information. Both fairness and efficiency counsel imposing the burden of proof on the party with superior access to information.[109] That party can produce information with less effort and runs a lower risk of having wrongdoing attributed incorrectly. States funding private groups have better access to information than victim states. The virtual control approach aligns the law of state responsibility for cyber attacks with these factors.

VII POTENTIAL OBJECTIONS

A *Invading Privacy and Curbing Dissent*

One objection to the burden-shifting aspect of the proposal would be that if states adopted it, they would have an incentive to monitor private communications of their citizens to a greater degree to more readily discharge their burden of proof. Some American legislation, such as the *DMCA*, has

---

108 As the *Manual* commentary notes, there is disagreement about whether such a cyber intrusion is an 'armed attack' under art 51 of the *UN Charter*: ibid 108–9 [10]–[11].

109 See Alina Das, 'Immigration Detention: Information Gaps and Institutional Barriers to Reform' (2013) 80 *University of Chicago Law Review* 137, 156 (asserting that in criminal cases, 'the placement of a high burden of proof upon the government strengthens the government's incentive to acquire and use information to meet its burden').

already triggered concerns about undue curbing of internet expression.[110] Recent revelations about the US National Security Agency compiling call record data sharpen this concern.[111] On balance, however, the privacy concerns associated with the burden-shifting proposal are overplayed. First, the proposal would only incentivise state monitoring of those who received funding from the state for cyber-related work. That would minimise the number of individuals and entities affected. Secondly, grantors already often require regular statements or other documentation from grantees. This documentation is necessary for any grantor to satisfy its due diligence obligations to its own contributors and constituents. The proposal would require only a marginal increase. Nothing in the proposal would require broader monitoring of a state's nationals or curbs on lawful online speech.

B *Risking Escalation*

The burden-shifting approach also arguably encourages the escalation of disputes. However, this view unduly discounts two important factors. First, it offers an incomplete account of incentives that trigger escalation. Secondly, it fails to acknowledge other legal constraints on the victim state's ability to respond to cyber attacks.

The biggest incentive for escalation of a dispute is the unwillingness of a victim state to appear helpless in the face of an attack. That image of helplessness has negative repercussions for a victim state regime, both domestically and internationally. Most regimes will strive to find some basis for retaliation, perhaps stretching international law norms in the process. Viewed *ex ante*, allowing a broader range of responses by a victim state can address this dynamic by deterring cyber attacks in the first instance. That deterrent effect is a central rationale for the right of self-defence. Here, as elsewhere, the best way to avoid escalation of a dispute is to ensure that no dispute even arises.

When a conflict occurs over a cyber attack, a victim state is still subject to a range of legal restrictions that limit escalation. For example, the principle of *ad bellum* proportionality provides some limit, even when the necessity principle

---

110 Cf Roger Hurwitz, 'Taking Care: Four Takes on the Cyber Steward' (Paper presented at Cyber Dialogue 2012: What is Stewardship in Cyberspace?, University of Toronto, 18–19 March 2012) 3–4 <http://www.cyberdialogue.citizenlab.org/wp-content/uploads/2012/2012papers/CyberDialogue2012_hurwitz.pdf> (discussing concerns about privacy and free expression and selective application in US policy); Nathan Alexander Sales, 'Regulating Cybersecurity' (2013) 107 *Northwestern University Law Review* (forthcoming) (discussing models and rationales for domestic US cyber-security regulation).

111 Ascertaining the legality of the National Security Agency ('NSA') programs is beyond the scope of this commentary, although defenders of the legality of the programs make strong arguments. See Benjamin Wittes, 'The Minimization and Targeting Procedures: An Analysis' on *Lawfare* (23 June 2013) <http://www.lawfareblog.com/2013/06/the-miminization-and-targeting-procedures-an-analysis> (suggesting that the NSA program that obtains the content of phone conversations involving non-US persons located overseas includes reasonable safeguards on the obtaining of data involving protected groups and is therefore consistent with US domestic law).

permits force.[112] *Ad bellum* proportionality limits the 'scale, scope, duration, and intensity' of the victim's use of force to that required to definitively address the threat posed by the armed attack.[113] If a cyber response will end the threat, international law would bar the use of kinetic force.[114] While the principle of *ad bellum* proportionality would not rule out the use of kinetic force — for example, when the control state had sophisticated cyber-defences that made it impervious to attack — it would limit the use of kinetic force in situations where a cyber response would be sufficient.

VIII CONCLUSION

Cyber attack is a challenging subject that creates a dilemma for authors of a restatement. Creating new rules divorced from LOAC would unmoor cyber from current constraints. This would create the potential for expanded conflict and unnecessary suffering, undermining LOAC's objectives. On the other hand, unduly rigid limits based on current law can stifle state practice and inhibit the development of LOAC as applied to cyberspace. In pondering difficult questions associated with the attribution of state responsibility, the drafters of the *Manual* have erred on the side of the former concern, anchoring cyber capabilities to current law. However, their caution about new rules may inhibit adaptation of the law to cyber's special challenges.

The *Manual*'s nod to the *Draft Articles* is an understandable but costly gesture. The *Draft Articles* systematically discounted the threat posed by state support of violent non-state actors and offered little in the way of guidance for addressing transnational terrorism. If that is true for the kinetic realm, it is even more accurate for the cyber domain. The tests of effective and overall control, although justified in the kinetic realm, fail to address what this article calls the

---

112 See Schmitt, above n 1, 62–3 (Rule 14, [5]). See also William Banks, 'The Role of Counterterrorism Law in Shaping *ad Bellum* Norms for Cyber Warfare' (2013) 89 *International Law Studies* 157, 175–93 (discussing the role for evolving counterterrorism norms in regulating cyber attacks). Cf Eric Talbot Jensen, 'Cyber Attacks: Proportionality and Precautions in Attack' (2013) 89 *International Law Studies* 198, 204–9. Jensen notes that the conduct of hostilities is also subject to the *in bello* proportionality rule, which bars harm to civilians that is, in the words of art 51(5)(b) of *Additional Protocol I* to the *Geneva Conventions*, 'excessive in relation to the concrete and direct military advantage anticipated' by a commander ordering a specific attack: at 204, quoting *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts*, opened for signature 8 June 1977, 1125 UNTS 3 (entered into force 7 December 1978). See also Noam Lubell, 'Lawful Targets in Cyber Operations: Does the Principle of Distinction Apply?' (2013) 89 *International Law Studies* 252, 264–74 (noting application of the *in bello* principle of distinction, which permits targeting only of military objectives).

113 Schmitt, above n 1, 62–3 (Rule 14, [5]). To limit escalation, Professor Schmitt has also suggested that under international law states that are victimized by cyber intrusions can respond with appropriate countermeasures. See Michael N. Schmitt, '"Below the Threshold" Cyber Operations: The Countermeasures Response Option and International Law' (2014) 54 *Virginia Journal of International Law* (forthcoming) <http://ssrn.com/abstract=2353898>. While this article is a valuable inquiry into countermeasures' potential in the cyber arena, the limits on countermeasures may not provide an adequate deterrent to cyber intrusions. Cf ibid 22 (noting that states with greater cyber capabilities cannot engage in countermeasures on behalf of victim states with more modest capabilities).

114 Cf Schmitt, above n 1, 62–3 (Rule 14, [5]) (noting that international law on the use of force 'limits the scale, scope, duration, and intensity of the defensive response to that required to end the situation' that created the threat).

attribution asymmetry of cyber. Cyber attacks are often more difficult to detect than kinetic onslaughts, yet far easier to control at the source, precisely because their deployment is not linked to territorial imperatives. That combination of the difficulty of external monitoring and the ease of internal restraint requires a broader test, which this commentary has called virtual control.

Under the virtual control test, a victim state that has demonstrated that another nation funded or equipped a non-state actor can hold the second state responsible for the non-state actor's cyber attacks, unless the second state rebuts the presumption of responsibility. The second state may rebut that presumption through cooperation in the victim state's attribution efforts. Cooperation will shift the burden of attribution back to the victim state and keep the standard at the overall control standard endorsed by the ICTY in *Tadić*.

Sovereignty may look different with a broader test for attribution of state responsibility. States may need to do more to regulate groups that they assist. That may well have implications for some aspects of state governance. However, the consequence need not be rigid state control over the internet or elimination of privacy. Instead, greater diligence in state funding and monitoring of grantees should solve the problem. States typically reserve the right to monitor and audit grantees today, so the measures added by this commentary's proposal should occur only at the margins.

Of course, states that wish to preserve plausible deniability for cyber attacks will have reason to reject this proposal. However, that prospect merely confirms the proposal's value. In the brave new world of cyber attacks, plausible deniability is a luxury international law cannot afford.