2014 6th International Conference on Cyber Conflict
P.Brangetto, M.Maybaum, J.Stinissen (Eds.)
2014 © NATO CCD COE Publications, Tallinn

The Drawbacks and Dangers of Active Defense

Oona A. Hathaway
Yale Law School
New Haven, CT
oona.hathaway@yale.edu

Abstract: The growing prevalence of cyber-attacks on states, businesses, and individuals has raised new and urgent questions about the legal framework that governs states' capacity to respond such attacks. An issue that has proven particularly vexing is what actions a state may take in response to attacks that fall into the gap between the actions that constitute a prohibited "use of force" under Article 2(4) of the UN Charter and the "armed attacks" to which a state has a right to respond with force in self defense under Article 51. Intrusions that constitution an illegal "use of force" but do not meet the "armed attack" threshold for triggering a legal forceful response—sometimes known as "below the threshold" cyber-operations—are extraordinarily common. Indeed, nearly all cyber-attacks by one state on another fall below the "armed attack" threshold. If states cannot legally use their right to self-defense to respond to such unlawful attacks, what can they do? There is a growing consensus that the answer can be found in countermeasures doctrine. Yet countermeasures doctrine was never intended to be applied to actions that constitute uses of force. There is good reason for this: if forceful countermeasures were allowed, there would be a serious danger that the system restricting illegal use of force would spin out of control. Improper countermeasures are inevitable, and escalation of conflict only a matter of time. This paper outlines the legal principles governing the use of force in international affairs, describes the exceptions to the broad prohibition on the use of military force, outlines the doctrine of countermeasures, and—in its key contribution to the debate—outlines reasons for concern about aggressive countermeasures. The paper concludes by briefly considering non-forceful responses that states may take in response to cyber-attacks.

Keywords: *Active defense, United Nations Charter, international law, self-defense*

Cyber-attacks have become an ever-present threat to states, individuals, and businesses throughout the world.[1] British Petroleum has reported that it faces a barrage of 50,000 attempts at cyber-intrusion a day.[2] The U.S. Pentagon has reported ten million attempts per day.[3] The U.S. National Nuclear Security Administration also records ten million attempts at hacking each day.[4] If only one out of one hundred million attacks succeeds, the national security of the United States is dangerously vulnerable.

These new threats to national security have raised deep questions about the capacity of states to protect themselves. In response, the legal framework that governs the use of force in the cyber context has been slowly taking shape. There is a growing consensus that the standard rules governing use of force in international law apply to this unconventional threat. The Tallinn Manual, now in the midst of revision and expansion, represents an extraordinary collaboration of scholars seeking to outline the specific implications of that law for cyber.[5]

An issue that has proven particularly vexing is the gap between the actions that constitute a prohibited "use of force" under Article 2(4) of the UN Charter and the "armed attacks" to which a state has a right to respond with force in self defense under Article 51. There is a well-known gap between those intrusions that are illegal and those that meet the "armed attack" threshold for triggering a legal forceful response.[6] These "below the threshold" cyber-operations, as Michael Schmitt has dubbed them, are extraordinarily common. Indeed, nearly all cyber-attacks by one state on another fall below the "armed attack" threshold.

If states cannot legally use their right to self-defense to respond to unlawful attacks below the threshold, what can they do? There is a growing consensus that the answer can be found in countermeasures doctrine. States, the argument goes, may respond in kind to an attack as long as they meet the various requirements of countermeasures doctrine—most notably that the countermeasure is proportional to the unlawful behavior that prompted it and is designed to bring the violating state back into compliance.

This paper aims to sound a cautionary note in the face of this growing consensus. It points out that countermeasures doctrine has never been applied in the use of force context and, indeed, commentary on the countermeasures doctrine makes clear that it was not intended to be applied to actions that constitute uses of force. There is, moreover, a good reason for this: if millions of "below-the-threshold" attacks are met with millions of "below-the-threshold" attacks in

1 Portions of this paper are drawn from Oona A. Hathaway et al, *The Law of Cyber-Attack*, 100 Cal. L. Rev. 817 (2012).

2 Michael Tomaso, BP Fight Off Up to 50,000 Cyber0Attacks a Day, CNBC (Mar. 6, 2013).

3 Zachary Fryer-Biggs, U.S. Military Goes on Cyber Offensive, Defense News (Mar. 24, 2012).

4 Jason Koebler, U.S. Nukes Face up to 10 Million Cyber Attacks Daily, U.S. News (Mar. 20, 2012).

5 Tallinn Manual (Michael Schmitt, ed., 2013). Its editor, Michael Schmitt, has also addressed many of the most interesting an important legal challenges relating to the application of the law of jus ad bellum and jus in bello to cyber in his own extensive writings.

6 Harold Koh, while serving as Legal Adviser for the U.S. Department of State, took the position that there was no gap. Koh stated that "the inherent right of self-defense potentially applies against any illegal use of force... There is no threshold for a use of deadly force to qualify as an 'armed attack' that may warrant a forcible response." Michael N. Schmitt, *International Law in Cyberspace: The Koh Speech and Tallinn Manual Juxtaposed*, 54 Harvard Int'l L.J. 21-22 (Dec. 2012). Most scholars disagree with this view, concluding that there is, in fact, a gap between the two. *See id*.; Tom Ruys, 'Armed Attack' and Article 51 of the UN Charter: Evolutions in Customary Law and Practice 139-84 (2010). Randelzhofer shows sympathy for closing the gap between Articles 2(4) and 51 by allowing states to respond to any use of force but expresses doubt about whether that view is consistent with the Charter. A. Randelzhofer, Article 51, in B. Simma et al, eds., The Charter of the United Nations: A Commentary, Vol 1 (2002), at pp. 791-92.

response, there is a serious danger that the system restricting illegal use of force will spin out of control. Improper countermeasures are inevitable, and escalation of conflict only a matter of time.

This paper proceeds in four parts. First, it briefly outlines the legal principles governing the use of force in international affairs. Second, it describes the exceptions to the broad prohibition on the use of military force. Third, it outlines the doctrine of countermeasures. Fourth—in its central contribution to the debate—the paper explains the reasons for concern about aggressive countermeasures. It concludes by briefly considering non-forceful responses that states may take in response to cyber-attacks.

1. GOVERNING LEGAL PRINCIPLES: PROHIBITION ON USE OF FORCE AND INTERVENTION IN INTERNAL AFFAIRS

Article 2(4) of the U.N. Charter provides that member states "shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations."[7] This prohibition is complemented by a customary international law norm of non-intervention, which prohibits states from interfering in the internal affairs of other states.[8] The International Court of Justice ("ICJ") has held that, where the interference takes the form of a use or threat of force, the customary international law norm of non-intervention is coterminous with Article 2(4).[9]

The precise scope of the international prohibition on the threat or use of force has been the subject of intense international and scholarly debate. Weaker states and some scholars have argued that Article 2(4) broadly prohibits not only the use of armed force, but also political and economic coercion. Nonetheless, the consensus is that Article 2(4) prohibits only armed force.[10]

Discussions about cyber-attacks have the potential to reignite debates over the scope of Article 2(4).[11] Because it is much less costly to mount cyber-attacks than to launch conventional

7 U.N. Charter art. 2, para. 4.

8 *See* G.A. Res. 37/10, U.N. Doc. A/RES/37/10 (Nov. 15, 1982); G.A. Res. 25/2625, U.N. Doc. A/RES/25/2625 (Oct. 24, 1970).

9 Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, para. 209 (June 27) ("[A]cts constituting a breach of the customary principle of non-intervention will also, if they directly or indirectly involve the use of force, constitute a breach of the principle of non-use of force in international relations."). It is possible, however, that to the extent cyber-attacks do not constitute a use of force, they may nevertheless violate the customary international law norm of non-intervention, as discussed below.

10 Daniel B. Silver, *Computer Network Attack as a Use of Force Under Article 2(4) of the United Nations Charter*, in Computer Network Attack and International Law 73, 80–82 (Michael N. Schmitt & Brian T. O'Donnell eds., 2002). The principal arguments for the prevailing view are: (1) that Article 2(4) was conceived against a background of efforts to limit unilateral recourse to armed force, not economic and political coercion; (2) that the *travaux preparatoires* show that the San Francisco Conference rejected a proposal that would have extended Article 2(4) to include economic sanctions; and (3) that the ICJ has held that financing armed insurrection does not constitute force, indicating that other economic measures that are even less directly related to armed violence would not constitute prohibited force either. Id. at 81. There remains some ambiguity, however, as to the extent to which Article 2(4) prohibits non-military physical force, such as flooding, forest fires, or pollution. *Id*. at 82–83.

11 *See* Matthew C. Waxman, Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4), 36 YALE J. INT'L L. 421, 458-59 (2011).

attacks, and because highly industrialized states are generally more dependent upon computer networks and are more vulnerable to cyber-attacks, cyber-attacks may prove to be a powerful weapon of the weak. This change in the cost structure of offensive capabilities may both increase the likelihood of cyber-attacks and change the political valence of different interpretations of Article 2(4)'s scope. Stronger states may begin to favor more expansive readings of Article 2(4) that prohibit coercive activities like cyber-attacks.[12]

Cyber-attacks may also violate the customary international law norm of non-intervention, as defined by a growing record of state practice and *opinio juris*. First, states generally do not engage in cyber-attacks openly, but rather try to hide their responsibility by camouflaging attacks through technical means[13] and by perpetrating the attacks through non-state actors with ambiguous relationships to state agencies.[14] As Thomas Franck has observed, "[l]ying about facts . . . is the tribute scofflaw governments pay to international legal obligations they violate."[15] In other words, the very fact that states attempt to hide their cyber-attacks may betray a concern that such attacks may constitute unlawful uses of force. Second, when states acknowledge that they have been victims of cyber-attack, they and their allies tend to denounce and condemn the attacks.[16] Third, in its common approach to cyber-defense, NATO has indicated that cyber-attacks trigger states parties' obligations under Article 4 of the North Atlantic Treaty,[17] which applies only when "the territorial integrity, political independence or security of any of the Parties is threatened."[18] The invocation of this provision strongly suggests that NATO member states believe that cyber-attacks violate the customary norm of non-intervention or a related international law norm.[19] Still, as the next Section explains, the fact that a cyber-attack is unlawful does not necessarily mean that armed force can be used in response.

2. EXCEPTIONS FOR COLLECTIVE SECURITY AND SELF-DEFENSE

Article 2(4)'s blanket prohibition on the non-consensual use or threat of force is subject to two exceptions: actions taken as part of collective security operations and actions taken in self-defense.

12 Walter Sharp has advocated that the United States make precisely this kind of strategic interpretive move, arguing that a broad array of coercive cyber-activities should fall within Article 2(4)'s prohibition. Walter Gary Sharp, Sr., CyberSpace and the Use of Force 129–33 (1999).

13 *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., Fall 2009, at 1, 74–75.

14 *See, e.g.*, Jeffrey Carr, Inside Cyber Warfare 176 (2010), at 29 ("Hacking attacks cloaked in nationalism are not only not prosecuted by Russian authorities, but they are encouraged through their proxies, the Russian youth associations, and the Foundation for Effective Policy.").

15 Thomas M. Franck, *Legitimacy After Kosovo and Iraq, in* International Law and the Use of Force at the Turn of Centuries: Essays in Honour of V. D. Degan 69, 73 (Vesna Crnić-Grotić & Miomir Matulović eds., 2005).

16 *See, e.g.*, Ian Traynor, *Russia Accused of Unleashing Cyberwar to Disable Estonia*, Guardian, May 16, 2007, http://www.guardian.co.uk/world/2007/may/17/topstories3.russia (detailing the reactions by Estonian, EU, and NATO officials to a cyber-attack on Estonia).

17 *NATO Agrees Common Approach to Cyber Defence*, Euractiv.com (Apr. 4, 2008), http://www.euractiv.com/en/infosociety/nato-agrees-common-approach-cyber-defence/article-171377.

18 North Atlantic Treaty, art. 4, Apr. 4, 1949, 63 Stat. 2241, 34 U.N.T.S. 243.

19 NATO has not endorsed the view that cyber-attacks rise to the level of armed attacks justifying self defense. *See NATO Agrees Common Approach to Cyber Defence, supra* note 17.

The first exception falls under Article 39 of the U.N. Charter. Article 39 empowers the Security Council to "determine the existence of any threat to the peace, breach of the peace, or act of aggression, and [to] make recommendations, or decide what measures shall be taken . . . to maintain or restore international peace and security."[20] The Security Council may employ "measures not involving the use of armed force"[21] and authorize "action by air, sea, or land forces."[22] Collective security operations under Article 39 can be politically difficult, however, because they require authorization by the often deadlocked or slow-moving Security Council.

The second exception to Article 2(4) is codified in Article 51, which provides that "[n]othing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs."[23] Lawful self-defense can be harder to define and identify than lawful collective security operations. Indeed, in many armed conflicts, both sides claim to be acting in self-defense, and the international debates tend to focus on factual and political disputes rather than legal doctrine.[24] It is clear, however, that the critical question determining the lawfulness of self-defense is whether or not an armed attack has occurred. A cyber-attack must rise to the level of an armed attack for a state to lawfully respond under Article 51.[25]

In scholarly debates over the application of *jus ad bellum* to cyber-attacks, three leading views have emerged to determine when a cyber-attack constitutes an armed attack that triggers the right of armed self-defense: the instrument-based approach, the target-based approach, and the effects-based approach.[26] Scholarly judgment has largely coalesced around the effect-based approach.[27] In essence, that approach holds that an attack is judge by its effects. For example, Daniel Silver, former General Counsel of the CIA and National Security Agency, argues that the key criterion determining when a cyber-attack constitutes an armed attack is the severity of the harm caused. A cyber-attack justifies self-defense "only if its foreseeable consequence is to cause physical injury or property damage and, even then, only if the severity

20 U.N. Charter art. 39.

21 *Id*. art. 41.

22 *Id*. art. 42.

23 *Id*. art. 51. For example, the White House's recent cyberspace strategy paper includes the right of self-defense as one of the norms that should guide conduct in cyberspace. International Strategy for Cyberspace, White House 5 (May, 2011), [hereinafter White House Cyberspace Strategy] available at http://www.whitehouse.gov/sites/default/files/rss_viewer/ international_strategy_for_cyberspace.pdf. at 10.

24 Christine Gray, International Law and the Use of Force 95–96 (2d ed. 2004).

25 *See, e.g.*, International Strategy for Cyberspace, White House 5 (May, 2011), [hereinafter White House Cyberspace Strategy] *available at* http://www.whitehouse.gov/sites/default/files/rss_viewer/international_strategy_ for_cyberspace.pdf, at 14 ("When warranted, the United States will respond to hostile acts in cyberspace as we would to any other threat to our country. All states possess an inherent right to self-defense, and we recognize that certain hostile acts conducted through cyberspace could compel actions under the commitments we have with our military treaty partners.").

26 Once a state has been the victim of an armed attack, a further question arises as to against whom the state can respond. Where the armed attack is perpetrated by a state, this question is easily answered—self-defense may be directed against the perpetrating state. However, cyber-attacks may be perpetrated by non-state actors or by actors with unclear affiliations with state security agencies. Although some scholars argue that cyber-attacks (and conventional attacks) must be attributable to a perpetrating state in order for the victim state to take defensive action that breaches another state's territory, others—drawing on traditional jurisprudence on self-defense—argue that states possess the right to engage in self-defense directly against non-state actors if certain conditions are met. See Jordan J. Paust, *Self-Defense Targetings of Non-State Actors and Permissibility of U.S. Use of Drones in Pakistan*, 19 J. Transnat'l L. & Pol'y 237, 238–39 (2010) ("The vast majority of writers agree that an armed attack by a non-state actor on a state, its embassies, its military, or other nationals abroad can trigger the right of self-defense addressed in Article 51 of the United Nations Charter, even if selective responsive force directed against a non-state actor occurs within a foreign country.").

27 *See* Hathaway, et al, *supra* note 1.

of those foreseeable consequences resembles the consequences that are associated with armed coercion."[28] Under this test, a cyber-attack on the air traffic control system causing planes to crash would be regarded as an armed attack, because it is foreseeable that such an attack would cause loss of life and substantial property damage. But a cyber-attack on a website or mere penetration of a critical computer system generally would not, unless it caused physical injury or property damage. A cyber-attack on financial systems presents a harder case for this approach—the analysis would depend on whether the attack was found to have caused substantial "property damage." This effects test defines a small core of harmful cyber-attacks that rise to the level of an armed attack.[29] It also focuses the armed attack analysis on a limited set of criteria—particularly severity and foreseeability.[30]

The effects test solves the problem of how to judge the severity of a cyber attack. But it leaves intact the problem of a gap between the uses of force that constitute a violation of Article 2(4) and armed attacks sufficient to give rise to the right to respond with force under Article 51. Indeed, the "armed attack" is linguistically distinct from several other related terms in the U.N. Charter and has been interpreted to be substantively narrower than them.[31] The ICJ has indicated that cross-border incursions that are minor in their "scale and effects" may be classified as mere "frontier incident[s]" rather than "armed attacks."[32] Instead, to be armed attacks sufficient to justify a response under Article 51, attacks must be of sufficient gravity to constitute "most grave forms of the use of force."[33] Where they may not resort to defensive force under Article 51 (because an attack does not arise to the level of an "armed attack"), states may be permitted to respond with retorsions or non-forceful countermeasures within carefully proscribed legal limits.[34]

28 Silver, *supra* note 10, at 90–91. It is important to note that the purpose of the attack is already accounted for in the definition of cyber-attack recommended herein: the attack must have been committed for a political or national security purpose. Therefore a cyber-attack that has unforeseen national security consequences would not be considered a cyber-attack, much less cyber-warfare.

29 *Id*. at 92.

30 The Department of Defense has signaled its approval of this approach. *See* Office of Gen. Counsel, Dep't of Def., An Assessment of International Legal Issues in Information Operations (1999), *reprinted* in Computer Network Attack and International Law 459, 484–85 [hereinafter DOD Memo] (Michael N. Schmitt & Brian T. O'Donnell eds., 2002), at 483 (arguing "the consequences are likely to be more important than the means used," and providing examples of cyber-attacks that would cause civilian deaths and property damage).

31 *See* Yoram Dinstein, Computer Network Attacks and Self-Defense, *in* Computer Network Attack and International Law 99, 100–01 (Michael N. Schmitt & Brian T. O'Donnell eds., 2002).

32 Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 195 (June 27); *cf*. Definition of Aggression, G.A. Res. 29/3314, Annex, art. 2, U.N. Doc. A/RES/29/3314 (Dec. 14, 1974) [hereinafter Definition of Aggression] (determining that "[t]he first use of armed force by a State in contravention of the Charter shall constitute *prima facie* evidence of an act of aggression although the Security Council may . . . conclude that a determination that an act of aggression has been committed would not be justified in the light of other relevant circumstances, including the fact that the acts concerned or their consequences are not of *sufficient gravity*" (emphasis added)) . Scholars generally agree that there is a gap between the prohibition on the use of force and the right of self-defense. *See, e.g*., Dinstein, *supra* note 31, at 99, 100–01.

33 Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 191 (June 27).

34 Retorsions are lawful unfriendly acts made in response to an international law violation by another state; countermeasures are acts that would be unlawful if not done in response to a prior international law violation. U.N. Int'l Law Comm'n Draft Articles on Responsibility of States for Internationally Wrongful Acts, Rep. of the Int'l Law Comm'n, U.N. GAOR, 53d Sess., Supp. No. 10, U.N. Doc. A/56/10 (2001), at 31, 80 [hereinafter Draft Articles]. *See* DOD Memo, *supra* note 30 ("If the provocation is not considered to be an armed attack, a similar response will also presumably not be considered to be an armed attack.").

Until recently, forceful countermeasures were generally regarded as outside the countermeasures regime. As the next section explores, however, that consensus has begun to crumble as a growing number of voices have called for forceful countermeasures for cyber.

3. COUNTERMEASURES

The customary international law of countermeasures governs how states may respond to international law violations that do not rise to the level of an armed attack justifying self-defense—including, implicitly, cyber-attacks. The Draft Articles on State Responsibility define countermeasures as "measures that would otherwise be contrary to the international obligations of an injured State *vis-à-vis* the responsible State, if they were not taken by the former in response to an internationally wrongful act by the latter in order to procure cessation and reparation."[35]

The international law of countermeasures does not define when a cyber-attack is unlawful—indeed the Draft Articles do not directly address cyber-attack at all. The law simply provides that when a state commits an international law violation, an injured state may respond with a countermeasure.[36] As explained above, some cyber-attacks that do not rise to the level of an armed attack nonetheless violate the customary international law norm of non-intervention.[37] These violations of international law may entitle a harmed state to use countermeasures to bring the responsible state into compliance with the law.

The Draft Articles lay out the basic customary international law principles regulating states' resort to countermeasures.[38] The Draft Articles provide that countermeasures must be targeted at the state responsible for the prior wrongful act and must be temporary and instrumentally directed to induce the responsible state to cease its violation.[39] Accordingly, countermeasures cannot be used if the international law violation has ceased. Countermeasures also can never justify the violation of fundamental human rights, humanitarian prohibitions on reprisals, or peremptory international norms, nor can they excuse failure to comply with dispute settlement procedures or to protect the inviolability of diplomats.[40]

35 Draft Articles, *supra* note 34, at 128. Traditionally, these acts were termed "reprisals," but this report follows the Draft Articles in using the more modern term "countermeasures." Reprisals now predominantly refer to forceful belligerent reprisals. *Id*.

36 States thus resort to countermeasures at their own risk. If the use of countermeasures does not comply with the applicable international legal requirements, the state may itself be responsible for an internationally wrongful act. *Id*. at 130.

37 *See* Hathaway et al, *supra* note 1.

38 Countermeasures are distinct from retorsions. Retorsions are acts that are unfriendly but lawful, such as limiting diplomatic relations or withdrawing from voluntary aid programs, and they always remain a lawful means for a State to respond to a cyber-attack or other international legal violation.

39 Draft Articles, *supra* note 34, at 129. Accordingly, the law of countermeasures does not specify how states may respond to international law violations by non-state actors. However, international law violations by non-state actors often lead to international law violations by states. For example, if a non-state actor launches an attack on state A from state B's territory and state B is unwilling or unable to stop it, state B may violate an international law obligation to prevent its territory from being used for cross-border attacks. *See, e.g*., Corfu Channel Case (U.K. v. Albania) (Merits), 1949 I.C.J. 4, 22 (Apr. 9) (holding that states are obligated "not to allow knowingly its territory to be used for acts contrary to the rights of other States"). In the cyber-attack context, a state may commit an international law violation by allowing harmful cyber-attacks to be launched from its territory. *See* Sklerov, *supra* note 13, at 62–72.

40 Draft Articles, *supra* note 34, at 131.

Before resorting to countermeasures, the injured state generally must call upon the responsible state to cease its wrongful conduct, notify it of the decision to employ countermeasures, and offer to negotiate a settlement.[41] However, in some situations, the injured state "may take such urgent countermeasures as are necessary to preserve its rights."[42] Countermeasures need not necessarily be reciprocal, but reciprocal measures are favored over other types because they are more likely to comply with the requirements of necessity and proportionality.[43] Under the customary law of countermeasures, an attacking state that violates its obligation not to intervene in another sovereign state through a harmful cyber-attack may be subject to lawful countermeasures by the injured State.

A rising number of institutions and scholars have left the door open to active countermeasures in response to illegal cyber-attacks. In this view, countermeasures might go beyond "passive defenses," such as firewalls, that aim to repel cyber-attacks, and constitute "active defenses," which attempt to disable the source of an attack.[44] Active defenses—if properly designed to meet the requirements of necessity and proportionality—might be considered a form of "reciprocal countermeasures," in which the injured state ceases obeying the same or a related obligation to the one the responsible state violated (in this case, the obligation of non-intervention).

Before a state may use active defenses as a countermeasure, however, it must determine that an internationally wrongful act caused the state harm and identify the state responsible, as well as abide by other restrictions.[45] The countermeasures must be designed, for example, to induce the wrongdoing state to comply with its obligations. The Draft Articles also have detailed provisions regarding when acts committed by non-state agents may be attributed to a state—for instance, when the state aids and assists the act with knowledge of the circumstances.[46] Countermeasures must also be necessary and proportional. Though there is no requirement that countermeasures are taken in relation to the same or a closely related obligation, the Commentary notes that necessity and proportionality will be more likely to be satisfied if they are.[47]

While countermeasures provide states with a valuable tool for addressing cyber-attacks that do not rise to the level of an armed attack, countermeasures are far from a panacea. Even putting to one side concerns about legality, there are practical challenges to an active countermeasures regime. First and foremost, cyber countermeasures require the identity of the attacker and the computer or network from which the attack originates to be accurately identified. Second, in order for a countermeasure to be effective, the targeted actor must find the countermeasure

41 *Id*. at 135.
42 *Id*.
43 *Id*. at 129.
44 In 2011, the Department of Defense has made clear that it employs such "active cyber defense" to "detect and stop malicious activity before it can affect DoD networks and systems." U.S. Dep't of Def., Department of Defense Strategy for Operating in Cyberspace 2 (July 2011) [hereinafter Dod Strategy], at 7; see Comm. on Offensive Info. Warfare, Nat'l Research Council of the Nat'l Acads., Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities 38 (William A. Owens et al. eds., 2009) [hereinafter NRC REPORT], at 142-49 (outlining possible "active responses" to cyber-attacks); Jay P. Kesan & Carol M. Hayes, Mitigative Counterstriking: Self-Defense and Deterrence in Cyberspace, 25 Harv. J. L. & Tech 415 (2012) (arguing that "permitting mitigative counterstrikes in response to cyberattacks would be more optimal" than the current passive regime). *Cf.* Tallinn Manual; Michael N. Schmitt, "Below the Threshold" Cyber Operations: The Countermeasures Response Option and International Law, 54 V. J. I. L. __ (forthcoming 2014).
45 Draft Articles, *supra* note 34, at 129–34.
46 *Id*. at 65.
47 Commentaries to the draft articles on Responsibility of States for internationally wrongful acts (adopted by the International Law Commission at its 53rd Session) (2001), at 327 [hereinafter ILC Commentaries].

costly—ideally costly enough to cease its unlawful behavior. If the target can easily relocate its operations across national boundaries, as is often possible in the cyber-context, the countermeasure may not impose a significant cost on the actor responsible for the attack. For this reason, countermeasures are likely to be more effective against state actors and less effective against non-state actors. Finally, it can be difficult to design a countermeasure that targets only the actor that perpetuated the legally wrongful attack. In particular, a countermeasure that disables a computer or network may very well cause harm to those who have little or nothing to do with the unlawful attacks. This could have the perverse effect of making the state injured by the original attack a perpetrator of an unlawful attack against those who simply happen to share a network with the actor that generated the original attack or whose computer was being used as a pawn to carry out attacks without their knowledge or acquiescence. Together these challenges can lead a system that relies too heavily on active countermeasures from spinning out of control.

4. THE DRAWBACKS AND DANGERS OF DEVELOPING AN AGGRESSIVE COUNTERMEASURES REGIME

The rising chorus of voices in favor of an active countermeasures regime has thus far not taken full account of the potential drawbacks and dangers of such a regime. In this section, I outline both the legal concerns and policy concerns regarding active countermeasures. My hope is that this will give pause to those advocating an expansive countermeasure regime and encourage some careful thinking in the future about the appropriate limits on active countermeasures.

First, the legal constraints. Those who favor application of countermeasures as a means of addressing the gap between Article 2(4) and 51 often turn to the International Law Associations Draft Articles on State Responsibility as the source of authority on countermeasures. They point, in particular, to Article 49, which outlines the "object and limits" of countermeasures.[48] As described in the previous section, this Article establishes that an injured state may take countermeasures against a State that is responsible for an internationally wrongful act in order to induce the non-complying state to come into compliance.

But often overlooked in this discussion is the Article that follows immediately after Article 49. Article 50—"Obligations not affected by countermeasures"—outlines a series of constraints on countermeasures. Of particular importance to cyber is the first, which provides that "Countermeasures shall not affect . . . the obligations to refrain from the threat or use of force as embodied in the Charter of the United Nations."[49] Furthermore, Article 59 reaffirms that, "These articles are without prejudice to the Charter of the United Nations."[50]

The commentaries on the Draft Articles further reinforce that the Articles apply only to "non-forcible countermeasures."[51] It expressly notes that it "excludes forcible measures from the ambit of permissible countermeasures under chapter II."[52] Moreover, it notes:

48 Draft Articles, *supra* note 34, at art. 49.
49 Draft Articles, *supra* note 34, at art. 50 (a).
50 *Id*. at art. 59.
51 ILC Commentaries, *supra* note 47, at 327.
52 ILC Commentaries, *supra* note 47, at 334.

> The prohibition of forcible countermeasures is spelled out in the Declaration on Principles of International Law concerning Friendly Relations and Cooperation among States in accordance with the Charter of the United Nations, by which the General Assembly of the United Nations proclaimed that "States have a duty to refrain from acts of reprisal involving use of force." The prohibition is also consistent with prevailing doctrine as well as a number of authoritative pronouncements of international judicial and other bodies.[53]

The implications for active countermeasures against cyber-attacks should be obvious. If a cyber-attack constitutes a "use of force" in violation of Article 2(4)—and this is the source of their international wrongfulness—then an active countermeasure that utilizes similar technology to "hack back" is, presumably, also a "use of force." If that is the case, then the ILC Draft Articles and Commentaries would seem to prohibit such countermeasures—at least any countermeasures comparable to the act that prompted the response.

The *Tallinn Manual* experts and Mike Schmitt struggle admirably with these issues.[54] The *Tallinn Manual* experts were unable to decide even how to determine when a cyber-attack constituted an illegal use of force, much less what responses were permissible for those uses of force that fall in the gap between Article 2(4) and 51. Schmitt, writing separately, notes this lack of agreement. He identifies a minority view "that forceful countermeasures reaching the level of use of force are appropriate in response to an internationally wrongful act that constitutes a use of force, but remains below the armed attack threshold,"[55] pointing to a separate opinion by Judge Simma in the *Oil Platforms* case that some read to endorse forceful countermeasures.[56] Read in context, however, the opinion—which was, after all, the opinion of a single judge—does not stand for the proposition that forceful countermeasures are permitted. Instead, it simply makes the commonsense observation that "a State may of course defend itself" even against uses of force that do not amount to an armed attack, but such defense is subject to limits of "necessity, proportionality, and immediacy in a particular strict way."[57]

There is little legal support for the proposition that countermeasures doctrine provides a legal end-run around the prohibition on the use of force in Article 2(4) of the UN Charter. The leading authorities on countermeasures have affirmed that the UN Charter prohibitions are unaffected by the doctrine of lawful countermeasures. A state that counterstrikes or "hacks back" is therefore in violation of Article 2(4) of the UN Charter. It is true that the (now) victim state will not have the legal right to respond with force in self defense under Article 51, but the "hack back" (or "mitigative attack," as one article puts it[58]) is illegal nonetheless. Indeed, as a matter of international law, it is just as illegal as the attack that prompted it.

Is there is a class of cyber-attacks that do not amount to a "use of force" but constitute a violation of a customary norm of non-interference in a sovereign state that would give rise to a right to active cyber-defense? Again, the legal grounds for such a right to active cyber-defense are extremely weak. Those who hold that there is a right to non-interference distinct

53 ILC Commentaries, *supra* note 47, at 334.
54 Schmitt, *supra* note 44, at 16-19; Tallinn Manual, *supra* note 5, r. 48-52.
55 Schmitt, *supra* note 44, at 16.
56 Schmitt, supra note 44, at 16. Oil Platforms (Iran v. U.S.), 2003 I.C.J. 161 (Nov. 6), Separate Opinion of Judge Simma, ¶ 14.
57 *Oil Platforms*, Separate Opinion of Judge Simma ¶ 14.
58 Kesan & Hayes, *supra* note 44, at 469 ("Reflecting attacks back or initiating a new attack could, under the proper circumstances, both be considered mitigative counterattacks.").

from the prohibition on use of force often cite the *Nicaragua* case, where the International Court of Justice explained that the principle of state sovereignty "forbids all States or groups of States to intervene directly or indirectly in the internal or external affairs of other States."[59] A cyber attack could violate the right to non-interference, the argument goes, and therefore constitute internationally wrongful act that would trigger a right to respond with a non-forceful countermeasure (including a similar cyber attack). As yet, however, the norm of non-intervention likely remains too ill defined to support such a claim. It is far from clear that there is, indeed, a norm of non-intervention distinct from the prohibition on use of force in the UN Charter. Even were the norm better defined, cyber-attacks would be a poor fit. According to the *Nicaragua* case, the norm protects states from interference in "matters in which each State is permitted, by the principles of State sovereignty, to decide freely."[60] A cyber attack is generally not intended to "coerce" in this way.

There are important policy reasons for the legal limits on forceful countermeasures. There is a reason that the UN Charter does not permit states to respond with force to every single illegal use of force—in particular, to those uses of force that do not arise to the "most grave" level sufficient to amount to an "armed attack" and trigger Article 51. It is this: The gap between Article 2(4) and Article 51 prevents an endless process of retaliations for small offenses—a process that could, indeed is likely, to spin out of control over time. The gap between 2(4) and 51 puts some play in the joints, requiring states to absorb low-level uses of force without immediately responding in kind.

When considering the wisdom of continuing to observe this force gap, it is important to remember that cyber does not operate in isolation. If the legal principle were established that forceful countermeasures are permitted in cyber, there would be no reason not to apply those same principles outside the cyber context. If a state may respond to a use of force that does not rise to an armed attack with a use of force of its own in cyber, this could effectively eliminate the generally well-accepted gap between "use of force" under Article 2(4) and "armed attack" in Article 51. As a consequence, any use of force could provoke a forceful response. At stake, therefore, is not simply the capacity to respond to cyber-attacks, but the rules that govern the use of force in the international legal system more generally.

Likewise, there are good policy reasons to be wary of endorsing an expansive norm of non-interference that might give rise to a right to engage in active countermeasures. An expansive norm of non-interference could have far-reaching ramifications for other bodies of law. For example, if states have a right to demand non-interference by other states—and have a right to respond with countermeasures against those that do not observe this limit on interference—that might lead to countermeasures for a wide range of extraterritorial activities. Affected activities might include state funding for non-governmental organizations in other countries or extraterritorial application of commercial law (for example, anti-trust law and intellectual property law). It is important that lawyers and policymakers be careful not to create bigger problems in other areas of international law when trying to solve the threshold problem in cyber by engaging in over-interpretation of broadly applicable legal principles.

59 Nicaragua, ¶ 205. The Court continued: "A prohibited intervention must accordingly be one bearing on matters in which each State is permitted, by the principle of State sovereignty, to decide freely. One of these is the choice of a political, economic, social and cultural system, and the formulation of foreign policy. Intervention is wrongful when it uses methods of coercion in regard to such choices, which must remain free ones." Id.

60 Nicaragua, ¶ 205.

CONCLUSION: NON-FORCEFUL RESPONSES TO CYBER-ATTACKS AND A CALL FOR COLLABORATION

The argument made thus far may seem overly rigid and legalistic. Indeed, the prohibition on forceful countermeasures in cyber may appear absurd, effectively blessing illegal uses of force that stay just within the artificial line where a "use of force" crosses over into an "armed attack." But it is important to remember that even if force may not be used in response to an illegal use of force, states are not left powerless in the face of cyber-attacks. States that are subjected to an illegal use of force may respond with economic, diplomatic, or political sanctions—including asset freezes, trade sanctions, withdrawal of cooperation, travel bans, and banking restrictions—none of which are subject to limits under the UN Charter.[61] Customary countermeasures are limited to the suspension of international obligations, must be proportional, generally are "in kind"—involving like action for like action—and cannot be taken by third parties. Economic, diplomatic, and political sanctions are not subject to these same constraints (though they may be subject to independent legal constraints). As a result, sanctions can offer a wider range of options for responding to an unlawful action by a state—particularly an unlawful use of force—than do countermeasures.

States may also respond more directly with non-forceful cyber-measures. These might include some activities that have at times been classified as "active responses" to cyber-attacks—internal notification (notifying users, administrators, and management of the attacked entity), internal response (taking action to defend the system such as blocking certain IP addresses, creating an air gap), and external cooperative responses (including coordinated law enforcement and upstream support to internet service providers).[62] It may also include elements of non-cooperative information gathering and even traceback.

Collaboration between technical experts and international lawyers could be especially fruitful in drawing the line between cyber-responses that constitute uses of force and those that do not. Projecting satellite signals and sound waves into the sovereign space of another country do not constitute "uses of force." Nor does gathering satellite imagery—even very detailed imagery—or reporting activities of international news media, even state-run or state-funded news media, such as the BBC. Some of the more intrusive forms of intelligence gathering are also not restricted by international law, though the precise bounds of the international legal limits on such activities is a point of some contention.[63] The question that technical experts, collaborating with lawyers, could answer is what defensive cyber-measures are functionally similar to these well-accepted activities and which step over the line into use of force.

61 For more on what I call "outcasting," see Oona A. Hathaway & Scott J. Shapiro, Outcasting: Enforcement in Domestic and International Law, 121 Yale L. J. 252 (2011).

62 *See* NRC REPORT, *supra* note 44, at 148-49.

63 Compare 1 Oppenheim, International Law 862 (H. Lauterpacht ed., 8th ed. 1955) (asserting that peacetime intelligence gathering "is not considered wrong morally, politically or legally . . . ."), *and* Geoffrey B. Demarest, Espionage in International Law, 24 DENV. J. INT'L L. & POL'Y 321 (1996) (concluding that "peacetime espionage has always been seen as an issue of domestic law," and therefore not governed by international law), *with* Quincy Wright, Espionage and the Doctrine of Non-Intervention in Internal Affairs, in Essays on Espionage and International Law 3, 12 (Roland J. Stranger ed., 1962) (raising concerns that intelligence gathering may transgress the territorial integrity and political independence of a country, in violation of the UN Charter). It is clear that states may punish captured spies. They do not receive prisoner of war status or any of the immunities due to combatants in an armed conflict.