ADDRESSING OBSTACLES TO CYBER-ATTRIBUTION: A MODEL BASED ON STATE RESPONSE TO CYBER-ATTACK

CHRISTIAN PAYNE* AND LORRAINE FINLAY**

INTRODUCTION

The international law of *jus ad bellum*—the law surrounding the use of force by states outside of armed conflict—has always been fraught with political complications and potential legal ambiguity.[1] As states have become increasingly dependent on information technology, attack methods targeting information technology infrastructure have begun to receive recognition as a significant issue in the application of jus ad bellum.[2] While milestone cyber-attacks against Estonia in 2007 and the Stuxnet attack against Iran in 2010 are relatively well-known,[3] numerous other states have found themselves victims.[4] Consequently the United States and United Kingdom have both recognized cyber-attack as a major threat to their respective national securities.[5]

Cyber-attacks have properties that make them quite different from existing modes of warfare, and it remains uncertain precisely

* Lecturer, School of Engineering and Information Technology, Murdoch University. Ph.D. 2010, Murdoch University; B.Sc. 1999, Murdoch University.

** Lecturer, School of Law, Murdoch University. LL.M. 2009, New York University; LL.M. 2009, National University of Singapore; LL.B. 2003, University of Western Australia; B.A. 2003, University of Western Australia.

1. *See* Albrecht Randelzhofer, *Article 2(4), in* THE CHARTER OF THE UNITED NATIONS: A COMMENTARY 108–12 (Bruno Simma et al. eds., 1994).

2. *See, e.g.*, TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE, 1–3 (Michael N. Schmitt ed., 2013); Marco Roscini, *World Wide Warfare — Jus Ad Bellum and the Use of Cyber Force, in* 14 MAX PLANCK YEARBOOK OF U.N. LAW 85, 87 (2010).

3. *See* Michael N Schmitt, *The Law of Cyber Warfare: Quo Vadis?*, 25 STAN. L. & POL'Y REV. 269, 269 (2014) (arguing that the attacks on Estonia brought the issue of cyber-attack onto the international agenda); John Richardson, *Stuxnet as Cyberwarfare: Applying the Law of War to the Virtual Battlefield*, 29 J. MARSHALL J. COMPUTER & INFO. L. 1, 8 (2011) (discussing Stuxnet as the first cyber-attack by a state to have caused physical damage).

4. *See* Roscini, *supra* note 2 at 88–90 (2010).

5. WHITE HOUSE, NATIONAL SECURITY STRATEGY 27 (2010) (calling cyber-attack "one of the most serious national security . . . challenges" that the United States faces); WHITE HOUSE, NATIONAL SECURITY STRATEGY 1 (2015) (specifically acknowledging the growing danger of disruptive and destructive cyber-attacks); HM GOVERNMENT, A STRONG BRITAIN IN AN AGE OF UNCERTAINTY: THE NATIONAL SECURITY STRATEGY 11 (2010) (indicating that cyber-attack is one of the four highest priority risks the United Kingdom faces).

how public international law should deal with this new mode of attack.[6] In particular, it is unclear whether existing legal rules are suitably equipped to deal with the novel characteristics of cyber-warfare, or instead, if and to what degree new rules need to be developed.[7] The result is a raft of partly legal and partly technical questions, generally without clear answers. Although there has been much scholarly writing on the subject, state practice varies and there is no international jurisprudence to assist with these specific questions.[8] Indeed, the only extant legal "rules" on the topic are those proposed by scholars.[9] The academic literature has therefore had an unusually large impact on the understanding of the law in this emerging area.

Although scholars have analyzed a range of issues surrounding the law of cyber-attack, the question of how to attribute state responsibility in the event of a cyber-attack has been recognized as a significant hurdle.[10] Indeed, national security law expert Daniel B. Silver has described attribution as the most important practical obstacle to applying the law of jus ad bellum to cyber-attack.[11]

This Article discusses some of the legal issues surrounding cyber-attack generally, and cyber-attribution in particular. Leveraging existing jurisprudence, established international law, and state practice, it proposes a model for cyber-attribution that minimizes the risk of conflict escalation and encourages states to cooperate and pursue peaceful means of resolving disputes arising from cyber-incidents.

The Article begins in Part I by providing a brief overview regarding the nature of cyber-attacks, and particularly some of their unusual properties. Part II follows with a review of the rules of jus ad bellum, especially the prohibition against force and the right of self-defense. Part III then outlines states' lawful options when responding to cyber-attack, focusing on how the rules of jus ad bel-

6. *See* Schmitt, *supra* note 3, at 271–72.

7. *Id.*

8. *See* Lianne J.M. Boer, *'Restating the Law "As It Is"': On the Tallinn Manual and the Use of Force in Cyberspace*, 5 Amsterdam L.F. 4, 5 (2013).

9. *See id.*

10. Scott J. Shackelford & Richard B. Andres, *State Responsibility for Cyber Attacks: Competing Standards for a Growing Problem*, 42 Geo. J. Int'l L. 971, 981 (2011); David E. Graham, *Cyber Threats and the Law of War*, 4 J. Nat'l Sec. L. & Pol'y 87, 92 (2010); 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, 445 (2011).

11. *See* Daniel B. Silver, *Computer Network Attack as a Use of Force Under Article 2(4) of the United Nations Charter*, *in* 76 Computer Network Attack & International Law 73, 78 (Michael N. Schmitt & Brian T. O'Donnell eds., 2002).

lum may apply to this emerging form of conflict. Part IV then discusses the problems that cyber-attribution poses which practically inhibit the ability to apply existing international laws to cyber-attack scenarios. Finally, Part V presents and discusses a model that aims to resolve these problems.

I. The Nature of Cyber-Attack

For the purposes of this Article, cyber-attacks are defined as those whereby states utilize computers and information technology as the primary mechanisms to detrimentally impact the interests of another state. In short, cyber-attack involves information technology employed as a weapon. Although some definitions of the term might include traditional kinetic attacks targeted at computing infrastructure,[12] a more useful definition when concentrating on the unique characteristics of cyber-attacks and the challenges these present focuses on the notional weapon rather than the target.[13] Cyber-attacks involve a variety of attack mechanisms, targets, and consequences, which makes difficult clear identification of a common set of characteristics, compared with traditional kinetic attacks.[14] Combined with the relative newness of cyber-attack, these attributes create the potential for significant legal ambiguity.

Moreover, the complexity and flexibility of reprogrammable computer systems produce a seemingly infinite variety of potential attack methods.[15] The outcomes and constraints of these methods will generally depend upon the precise details of specific systems and their vulnerabilities. Consequently, any attempt at a definitive taxonomy is likely to have a degree of artificiality and be, at best, only temporarily accurate. For this reason, for the purposes of this Article, cyber-attacks are instead classified into one or more of three general categories.

First, direct intrusion attacks are those where the attacker directly interacts with a computer system and gains some form of unauthorized access or control. Second, indirect intrusion attacks involve malicious code or "malware" that the attacker constructs in order to compromise a set of systems.[16] In most cases this code will

12. Oona A. Hathaway et al., *The Law of Cyber-Attack*, 100 Cal. L. Rev. 817, 826 (2012).

13. Silver, *supra* note 11, at 75–76; Michael N. Schmitt, *Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework*, 37 Colum. J. Transnat'l L. 885, 891 (1999).

14. Schmitt, *supra* note 13, at 888.

15. *See generally* Bruce Schneier, Secrets & Lies: Digital Security in a Networked World 354–60 (2000) (explaining the implications of software complexity).

16. Matt Bishop, Computer Security: Art and Science 613–42 (2003).

be self-propagating, delivering a payload in the form of specific actions once it gains access to a system.[17] Third, denial of service attacks involve rendering the system unable to provide its expected services.[18] A common variant of these is the distributed denial of service attack whereby large numbers of computer systems are compromised and used to collectively overwhelm a target with meaningless messages.[19]

Cyber-attack of any type provides an attacker with significant advantages, making it highly asymmetric in nature.[20] To begin with, cyber-warfare is typically inexpensive for the attacker.[21] By one estimate, a devastating cyber-attack on the United States with damage lasting for many years could be deployed for approximately the cost of a single fighter aircraft.[22] Unlike traditional kinetic warfare, a cyber-attack can be launched by a single appropriately skilled individual with a computer, an Internet connection, and a relatively small financial stake.

The vulnerability of states to cyber-attack also varies significantly.[23] States with a greater investment in, and therefore dependence on, information technology have far greater potential exposure than those which are still developing in this area.[24] Nations with extensive information technology infrastructures will often also be economically and militarily strong;[25] notionally powerful states are therefore likely to be far more vulnerable to cyber-attack than other means of coercion such as kinetic warfare or economic action.[26]

Finally, the legal and factual ambiguity associated with cyber-warfare makes this a relatively safe mechanism for a prospective aggressor to employ. A victim state may simply not have sufficient evidence to prove who attacked it, and indeed may not even know

17. *Id.* at 623–24, § 22.7.5.

18. SCHNEIER, *supra* note 15, at 181–86.

19. *Id.* at 184–86.

20. John Dever & James Dever, *Cyber Warfare: Attribution, Preemption, and National Self Defence*, 2 J. L. & CYBER WARFARE 25, 26 (2013).

21. Schmitt, *supra* note 13, at 897–98.

22. *Id.* at 898. According to the figures used by Schmitt in his 1999 article, it would take approximately US$30,000,000 to devastate the United States' information infrastructure for many years while a single F-16 aircraft would cost US$26,000,000. *Id.*

23. *Id.* at 897.

24. *Id.*

25. *See e.g.*, *id.* ("The technological and fiscal wherewithal of the developed states underlies an unprecedented level of military and economic supremacy."). *Cf.* Waxman, *supra* note 10, at 424–25 (security of technology creates a risk for the United States because "its high economic and military dependency on information technology").

26. Dever & Dever, *supra* note 20, at 26.

the identity of the perpetrator.[27] This problem is known as cyber-attribution, the principal focus of this Article.

II. Existing Principles of *Jus Ad Bellum*

The fundamental principles of jus ad bellum are well established. However, their precise application is often still subject to significant debate and disagreement. This Part summarizes these general principles as a basis for later discussion of how this law applies in relation to cyber-attack.

A. *Prohibition on the Use of Force*

The prohibition against the threat or use of force is expressly articulated in the Charter of the United Nations (Charter) Article 2, Paragraph 4, and is considered to be a peremptory norm of international law.[28] Notable instruments that preceded Article 2(4), such as the Kellogg-Briand Pact, were framed as condemning "war" specifically, and thereby allowed military action that could be classified as action other than actual war.[29] Article 2(4) remedied this potential loophole and sought to more thoroughly mitigate the risk of war by framing the prohibition in terms of "force" generally.[30] However, with this approach the Charter framed wrongfulness in expressly instrumental terms, constrained by what is understood to constitute "force." This potential ambiguity has led to debate as to whether the prohibition encompasses armed force or can be extended to economic coercion.[31] The general consensus is that Article 2(4) refers only to armed force, and economic coercion is not included.[32] Nonetheless, some arguments support a wider interpretation.[33] The same question arises with regards to nonmilitary physical force, although again a conservative approach prevails that generally excludes economically coercive acts from the prohibition on force.[34] In contrast, there is a strong consensus

27. Oliver Kessler & Wouter Werner, *Expertise, Uncertainty, and International Law: A Study of the Tallinn Manual on Cyberwarfare*, 26 Leiden J. Int'l L. 793, 799 (2013).

28. Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. 14, ¶ 190 (June 27).

29. Randelzhofer, *supra* note 1, at 110–11.

30. *Id.* at 111.

31. *Id.* at 112.

32. *Id.*; Waxman, *supra* note 10, at 427; Schmitt, *supra* note 13, at 908.

33. Michael N. Schmitt, *Cyber Operations and the Jus Ad Bellum Revisited*, 56 Vill. L. Rev. 569, 574 (2011).

34. Randelzhofer, *supra* note 1, at 113.

that indirect force against another state involving the use of irregulars or rebels will generally breach the prohibition.[35]

The "cognitive shortcut" employed by the drafters of the Charter in terms of force as a coercive mechanism[36] has also led to questions about its applicability to new forms of weaponry.[37] This issue has led to the practical adoption of an approach to evaluating force that focuses primarily on the consequences, rather than the means employed.[38]

B. *The Right to Self-Defense*

1. Article 51 and Armed Attack

The Charter's Article 51 provides that nothing in that document "shall impair the inherent right of . . . [self-defense] if an armed attack occurs."[39] This provision points to the origins of the right to self-defense in customary international law.[40] However, until the advent of the prohibition against war, and ultimately also against the use of force, the right to self-defense was of limited practical legal significance.[41] With the constraints on states that now exist by virtue of the Charter and the subsequent evolution of international norms, the nature and limits of this right have gained significance as one of the few ways that a state may unilaterally, lawfully use force.[42]

As expressed in Article 51, the first constraint on the right to self-defense is that it applies only in the event of an "armed attack."[43] Despite its origins in custom, the Charter states that it does not impair the inherent right to self-defense if an armed attack occurs, arguably implying that any previous customary right to self-defense that may have existed in the absence of such an attack has been

35. *Id.* at 113–14. *But see id.* at 115; Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. 14, ¶ 228 (June 27).

36. Schmitt, *supra* note 33, at 573.

37. Ian Brownlie, International Law and the Use of Force by States 362 (1963); Schmitt, *supra* note 13, at 913.

38. Brownlie, *supra* note 37; Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, ¶ 39 (July 8); Schmitt, *supra* note 33, at 573; Russell Buchan, *Cyber Attacks: Unlawful Uses of Force or Prohibited Interventions?*, 17 J. Conflict & Sec. L. 211, 217 (2012).

39. U.N. Charter art. 51.

40. *See also* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. 14, ¶ 176 (June 27).

41. Albrecht Randelzhofer, *Article 51*, *in* The Charter of the United Nations: A Commentary 661, 662–63 (Bruno Simma et al. eds., 1994).

42. *Id.* at 662–63.

43. U.N. Charter art. 51.

constrained or entirely discarded. Therefore, the right to self-defense is implicated only in the event of an armed attack, and the critical question becomes precisely what such an attack entails.

As there is no established definition of "armed attack" in the Charter or elsewhere in treaty law, its meaning is also determined by custom.[44] Despite statements to the contrary by the International Court of Justice (ICJ),[45] there does not appear to be a consensus on precisely what an armed attack entails.[46] It is generally accepted that such an attack is more narrow than the notion of force as prohibited by Article 2(4).[47] This interpretation restricts the instances where a state is permitted to undertake self-defense to only a subset of all possible circumstances where force has been used against it.[48]

The meaning of the term "armed attack" is therefore critical.[49] As with the prohibition against force, by framing the right to self-defense as permissible only in response to an armed attack, the drafters of the Charter again adopted a means-based approach.[50] However, the threshold required for an armed attack is more ambiguous than for the prohibition against the use of force, but is even more important.

The clearest enunciation of the threshold for an armed attack is found in the *Nicaragua* judgment where the ICJ first distinguished armed attacks as "the most grave forms of the use of force."[51] By itself, this dictum is far from conclusive, given that the use of force is also not an especially certain concept.[52] The ICJ then gave arguably the most concrete judicial test for armed attacks when it stated that they may be distinguished from lesser uses of force

---

44. Nicar. v. U.S, 1986 I.C.J. ¶ 176.

45. *Id.* at 103 ¶ 195.

46. Stanimir A. Alexandrove, Self-Defence Against the Use of Force in International Law 98 (1996).

47. Nicar. v. U.S., 1986 I.C.J. ¶ 191; Randelzhofer, *supra* note 41, at 663, 669. This is not the only position; however, it is the prevailing view. *Id.* at 664–67. A clear and notable exception is the United States, which considers that any illegal use of force enlivens the right to self-defense. Abraham D. Sofaer, *International Law and the Use of Force*, 82 Am. Soc'y Int'l L. Proc. 420, 422 (1988); Tallinn Manual, *supra* note 2.

48. Randelzhofer, *supra* note 41, at 663–64; Nicar. v. U.S., 1986 I.C.J. ¶ 211.

49. Randelzhofer, *supra* note 41, at 1397–428.

50. Schmitt, *supra* note 33, at 587.

51. Nicar. v. U.S., 1986 I.C.J. ¶ 191.

52. As one author wryly noted, "Force is like pornography: the law will recognise [it when it sees it]." Silver, *supra* note 11, at 84; *see also Jacobellis v. Ohio*, 378 U.S. 184 (1964) (Silver's statement alludes to Justice Stewart's concurrence in this case wherein His Honor stipulated that while he would not, and perhaps could not, define what constitutes 'hard-core pornography', he nonetheless stated, "I know it when I see it.").

("mere frontier incident[s]") by their "scale and effects."[53] This wording suggests that the armed force employed must be of a significant magnitude and intensity, perhaps best assessed based on the consequences that result.[54] However, the precise application of this test remains unclear. For example, in a more recent case, the ICJ expressly refused to exclude from the scope of armed attack an action on as small a scale as the mining of a single military vessel.[55]

2. The Elements of a Right to Self-Defense

As the ICJ has repeatedly emphasized, in addition to being a response to an armed attack, an act of self-defense is subject to meeting the elements of necessity and proportionality.[56] These elements are not contained within the text of Article 51, but instead are determined entirely by customary law.[57] An early but authoritative statement of these elements can be found in the diplomatic exchanges made in relation to the *Caroline* incident of 1837.[58] In particular, then-U.S. Secretary of State Daniel Webster formulated necessity to require that the instigative events be "instant, overwhelming, leaving no choice of means, and no moment for deliberation."[59] The effect of this requirement is that self-defense may only be undertaken when there is no realistic alternative.[60]

Necessity includes an aspect of immediacy.[61] While some scholars consider immediacy to be a separate element,[62] a more common, and altogether more logical, view is that it is merely one consideration in a determination whether an act of self-defense is indeed necessary.[63] Although the immediacy aspect has signifi-

---

53. Nicar. v. U.S., 1986 I.C.J. ¶ 195.

54. Avra Constantinou, The Right of Self Defence Under Customary International Law and Article 51 of the U.N. Charter 63–64 (2000).

55. Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. 161, ¶ 72 (Nov. 3).

56. *Id.* at 183 ¶ 43; Nicar. v. U.S., 1986 I.C.J. ¶ 194; Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 245, ¶ 41 (July 1996); Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda), Judgment, 2005 I.C.J. 168 ¶ 147 (Dec. 19).

57. Nicar. v. U.S., 1986 I.C.J. ¶ 176.

58. Tom Ruys, Armed Attack and Article 51 of the UN Charter: Evolutions in Customary Law and Practice 92 (2010).

59. R.Y. Jennings, *The Caroline and McLeod Cases*, 32 Am. J. Int'l L. 82, 89 (1938).

60. Ruys, *supra* note 58, at 95.

61. *See* Nicar. v. U.S., 1986 I.C.J. ¶ 237.

62. *See* Yoram Dinstein, War, Aggression, and Self Defence 209 (2005).

63. *See* Nicar. v. U.S., 1986 I.C.J. ¶ 237; Ruys, *supra* note 58, at 123–24.

cance when considering necessity,[64] a degree of flexibility is required when assessing whether a response is sufficiently prompt.[65]

Webster also articulated the proportionality element to require that the response involve "nothing unreasonable or excessive."[66] Proportionality can be assessed one of two different ways. A qualitative view requires that the response be effectively equal in gravity with regards to casualties, damage, and the nature of the weapons employed.[67] In general, however, a functional approach prevails wherein the focus is on employing a response of an appropriate scale to suitably repel the initial armed attack.[68]

Finally, arising out of the *Oil Platforms* case, a possible additional requirement for engaging in self-defense is that the armed attack be undertaken with the "specific intention of harming" the target that was ultimately attacked.[69] In that case, the ICJ noted that the explosive mine laid could not have been targeted at the specific vessel struck but simply at some target in those general waters.[70] As a result, it found the necessary intent to mount an armed attack lacking.[71] One interpretation of this element, as expressed in the judgment, is that the state which suffered the injury must actually have been the intended target.[72] The intention element is problematic as it requires the injured state to reliably infer what is essentially the state of mind of the attacking state.[73] Despite the clear enunciation found in *Oil Platforms*, this element remains somewhat controversial, and has essentially been rejected by the United States.[74]

---

64. *See* Nicar. v. U.S., 1986 I.C.J. ¶ 237 (wherein the International Court of Justice (ICJ) rejected measures the United States argued were defensive on the grounds that they took place many months after the relevant attack had occurred).

65. Ruys, *supra* note 58, at 100; Dinstein, *supra* note 62, at 210.

66. Jennings, *supra* note 59, at 89.

67. Ruys, *supra* note 58, at 111 (using the word "quantitative" to describe the qualitative proportionality test described above).

68. *Id.* at 112; Randelzhofer, *supra* note 41, at 1426.

69. Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. 161, ¶ 64 (Nov. 3).

70. *Id.*

71. *Id.*

72. *See* Dinstein, *supra* note 62, at 209 (arguing that the correct interpretation of the ICJ's decision, taken in light of the facts of the case, is that an attack which mistakenly targets the interests of a particular state does not enliven that state's right to self-defense); William H. Taft, *Self-Defence and the Oil Platforms Decision*, 29 Yale J. Int'l L. 295, 299 (2004).

73. Constantinou, *supra* note 54, at 62.

74. Taft, *supra* note 72, at 299, 302–03; Tallinn Manual, *supra* note 2, at 56 ¶ 11.

III. Responses to Cyber-Attack Under International Law

This Part reviews the application of existing international law to cyber-attack. After suffering a cyber-attack, one option of a victim state is to respond through a legally oriented process that seeks to establish that the attack was a breach of international law and that the victim state is therefore entitled to reparations. Alternatively, the state may pursue some form of self-help, such as by responding in self-defense. The following Sections discuss these approaches in turn.

A. *Legally Oriented State Responses*

Two processes of law by which a state may respond to a cyber-attack exist through the United Nations.[75] Article 39 of the Charter empowers the Security Council to "determine the existence of any threat to the peace, breach of the peace, or act of aggression" and then to respond through non-forcible or forceful measures.[76] Additionally, Resolution 377(V), known as the Uniting for Peace Resolution, may provide some scope for a response by the General Assembly.[77] However, the legal basis for this latter response is less certain and its scope far more constrained compared with action via the Security Council.[78] In reality, the practical likelihood of either U.N. response following a cyber-attack appears relatively low.[79] There are significant obstacles that stand in the way, and these are primarily political rather than legal.[80] Therefore, these approaches are likely to remain largely theoretical in all but the most exceptional circumstances.

---

75. *See, e.g.*, U.N. Charter arts. 41–42; G.A. Res. 377 (V), U.N. Doc. A/RES/377V (Nov. 3, 1950).

76. U.N. Charter arts. 41–42.

77. G.A. Res. 377 (V), U.N. Doc. A/RES/377V (Nov. 3, 1950).

78. *See generally* Juraj Andrassy, *Uniting for Peace*, 50 Am. J. Int'l L. 563 (1956); Harry Reicher, *The Uniting for Peace Resolution on the Thirtieth Anniversary of Its Passage*, 20 Colum. J. Transnat'l L. 1 (1981) (emphasizing that primary responsibility for maintaining peace lies with the Security Council).

79. Randelzhofer, *supra* note 1, at 119; Schmitt, *supra* note 13, at 902.

80. *See generally* Michael N. Schmitt, *Cyber Operations in International Law: The Use of Force, Collective Security, Self-Defense, and Armed Conflicts*, *in* Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy 151, 161–62 (2010); Yaroslav Radziwill, Cyber-Attacks and the Exploitable Imperfections of International Law 266–301 (2015) (emphasizing that defining a "threat to the peace" is a political challenge).

1. Cyber-Attack as a Use of Force

A more direct legal process that a victim state could pursue involves establishing that the cyber-attack was a breach of international law. States are limited in the ways they can characterize instances of cyber-aggression as a breach of international law. One such characterization is as a use of force in breach of the Charter.[81] The question surrounding this characterization that has most occupied scholars is how to determine that a cyber-attack has exceeded the threshold necessary to be "force."[82]

Unsurprisingly, this has proven to be a difficult question given that, even with regard to kinetic attacks, there is no conclusive definition of force.[83] As discussed above, the origin of this problem is the instrument-based approach to the framing of Article 2(4).[84] The prohibition in the instrument, rather than its consequences, is a layer of indirection, as the consequences are actually the relevant consideration.[85] As Professor Michael Schmitt notes, the "normative shorthand" used within the Charter does not work because a connection no longer exists between the instrument (force) and the consequences the provision seeks to prevent.[86]

Some scholars have advocated the simple test that a cyber-attack must result in physical damage to qualify as force.[87] By far the most influential and highly regarded analysis, however, is that of Professor Schmitt.[88] In his seminal 1999 paper, Schmitt outlined seven criteria to determine whether a given cyber-attack may constitute a use of force.[89] These were recently adopted with relatively minor

---

81. U.N. Charter art. 2, ¶ 4.

82. Schmitt, *supra* note 13; Waxman, *supra* note 10; Roscini, *supra* note 2; Silver, *supra* note 11; Herbert S. Lin, *Offensive Cyber Operations and the Use of Force*, 4 J. Nat'l Sec. L. & Pol'y 63 (2010); Matthew C. Waxman, *Cyber Attacks as "Force" Under UN Charter Article 2(4)*, 87 Int'l L. Stud. Ser. US Naval War Col. 43 (2011).

83. James E. McGhee, *Cyber Redux: The Schmitt Analysis, Tallinn Manual and US Cyber Policy*, 2 J.L. & Cyber Warfare 64, 98 (2013).

84. *See infra* Section II.A.

85. Schmitt, *supra* note 33, at 573; Brownlie, *supra* note 37, at 362 (stating that while it is true that "paragraph 4 applies to forces other than armed forces, it is very doubtful if it applies to economic measures of a coercive nature").

86. Schmitt, *supra* note 33, at 603.

87. Silver, *supra* note 11, at 90; Buchan, *supra* note 38, at 212.

88. *See, e.g.*, McGhee, *supra* note 83, at 67; Waxman, *supra* note 10, at 432; Matthew J. Sklerov, *Solving the Dilemma of State Responses to Cyber Attacks: A Justification for the Use of Active Defences Against States Who Neglect Their Duty to Prevent*, 201 Mil. L. Rev. 1, 56–57 (2009).

89. Professor Schmitt lists six criteria in the body of the text and indicates a possible seventh in a footnote. Schmitt, *supra* note 13, at 914–15 n.81. This seventh factor is adopted in his later work and that of other scholars. Schmitt, *supra* note 33, at 577; Tallinn Manual, *supra* note 2, at 51–52.

changes in the Tallinn Manual,[90] a consensus statement on the law of cyber-warfare by twenty academics and legal practitioners.[91] To the degree that any rules exist in this emerging area of law where there is no specific jurisprudence or clear state practice,[92] Schmitt's criteria on cyber-force are broadly accepted.[93] The criteria proposed by Schmitt are: severity, immediacy, directness, invasiveness, measurability, presumptive legitimacy, and responsibility.[94]

Severity relates to the degree to which attacks involve physical injury or the destruction of property.[95] This is generally considered to be the most influential of the criteria.[96] Immediacy draws a comparison between economic coercion, which is generally not regarded as a use of force, and armed coercion, which is, and suggests that the relative speed with which the latter takes effect may be used to determine if a cyber-attack represents a use of force.[97] The criterion of directness operates similarly, observing that the consequences of armed coercion flow quite directly from the event itself, whereas the effects of economic coercion depend on multiple contributory and supervening external factors.[98]

Invasiveness distinguishes a use of force from economic coercion based upon whether the act associated with the harmful outcome occurs within the target state's territory.[99] Unlike armed force, economic coercion—such as the imposition of sanctions—does not require any territorial intrusion on the part of the coercive state.[100] In his later analysis, Schmitt expands this notion to include the extent to which a target system is secured as a factor influencing the invasiveness of an attack.[101] However, the fundamental connection with the encroachment of territory remains, as Schmitt juxtaposes this aspect with an example of trade sanctions

90. TALLINN MANUAL, *supra* note 2, at 51–52.
91. *Id.*
92. Boer, *supra* note , at 5 (stating that international legal scholars played a significant role in absence of state practice and opinio juris).
93. *E.g.*, McGhee, *supra* note 83, at 67; Waxman, *supra* note 10, at 432; Sklerov, *supra* note 88, at 56–57.
94. Schmitt, *supra* note 13, at 914–15 n.81; Schmitt, *supra* note 33, at 576; TALLINN MANUAL, *supra* note 2, at 51–52.
95. Schmitt, *supra* note 33, at 576.
96. TALLINN MANUAL, *supra* note 2, at 48; Schmitt, *supra* note 33, at 576; Silver, *supra* note 11, at 90–91; Buchan, *supra* note 38, at 212.
97. Schmitt, *supra* note 13, at 914.
98. *Id.*
99. Schmitt, *supra* note 33, at 576.
100. *Id.*
101. *Id.*

being noninvasive.[102] Further, he clarifies the limits of the invasiveness element by noting that espionage is not a use of force, despite its highly invasive nature.[103]

Measurability involves the degree to which the consequences of an event can be ascertained, particularly in a quantitative way, which is largely a function of directness.[104] The criterion of presumptive legitimacy recognizes that, in many cases, cyber-attacks will simply be computerized variants of existing attack methodologies.[105] Schmitt argues that if computer technology is used to undertake some action that would previously not have been prohibited at international law, these computerized operations will also be legitimate.[106] This view follows the general principle that, absent some express prohibition to the contrary, at international law an act will be presumed legitimate.[107]

The final criterion involves the notion that state responsibility can be a factor when evaluating cyber-attack as a use of force.[108] The historically prevailing norm is that only states can exert armed coercion.[109] Building on this, Schmitt reasons that the more closely involved a state is in a particular cyber-operation, the greater basis for other states to consider the act a use of force.[110] In his original analysis, Schmitt dismissed this factor as a practical issue of limited normative value.[111] However, Schmitt's later work recognizes that state involvement in a particular action can have a normative effect beyond the question of attribution.[112] The understanding of this seventh criterion was expanded further by the later recognition that there are actually two separate factors encompassed within it: the degree of state involvement in and the military nature of the cyber-operation.[113] The Tallinn Manual adopts this

102. *Id.*

103. *Id.*

104. *Id.* at 576–77.

105. Susan W. Brenner, *"At Light Speed": Attribution and Response to Cybercrime/Terrorism/Warfare*, 97 J. Crim. L. & Criminology 379, 384 (2007).

106. Schmitt, *supra* note 33, at 577.

107. S.S. Lotus (Fr. v. Turk.), Judgment, 1927 P.C.I.J. (ser. A) No. 10, at 19, 31 (Sept. 7).

108. Schmitt, *supra* note 33, at 577.

109. *See* Schmitt, *supra* note 13, at 915 n.81.

110. *See* Schmitt, *supra* note 33, at 577.

111. *See* Schmitt, *supra* note 13, at 915 n.81. In forming this conclusion, it appears that Schmitt focuses primarily on the technical and factual question of attribution, rather than the normative effect of demonstrated state responsibility.

112. *See* Schmitt, *supra* note 33, at 577.

113. *See* Tallinn Manual, *supra* note 2, at 51.

more refined characterization of the responsibility criterion in Rule 11.[114]

2. Cyber-Attack as a Wrongful Intervention

An alternative theory for establishing a breach of international law involves treating acts of cyber-warfare as violations of the nonintervention principle.[115] While not expressly stated in the Charter, this principle was described by the ICJ as "part and parcel of customary international law."[116] As enunciated by the ICJ in the *Nicaragua* case, the principle proscribes coercive intervention in matters that states should be free to decide by virtue of their sovereignty.[117] For those who might question the Schmitt criteria or insist that a cyber-attack must cause physical damage to qualify as force, the principle of nonintervention represents an alternative limb on which to establish wrongfulness.[118]

If the principle of nonintervention may be used to establish wrongfulness, arguably the threshold question in relation to cyber-force is of limited significance and may be essentially moot. This outcome is somewhat ironic given that the threshold for cyber-force is perhaps the area of cyber-attack law with the greatest clarity and consensus.[119] Nonetheless, one may conclude that a cyber-attack is likely to be a breach of international law under existing general rules without the need for the development of *lex specialis* in this area.

If a cyber-attack is found to have breached international law, a legal response may require that the responsible state "make full reparation for the injury caused."[120] Reparations are likely to be a comparatively appealing option for an injured state in response to a cyber-attack than they would be in relation to a kinetic attack. Cyber-attacks do not involve the element of potentially escalating territorial incursion generally found with kinetic attacks, thus

114. *See id.* The remainder of Rule 11 essentially reflects Schmitt's criteria; *see also* Boer, *supra* note 8, at 7 (highlighting the refined clarity of the Tallin Manual).

115. *See* Buchan, *supra* note 38, at 221.

116. *See* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.) 1986 I.C.J. 14, ¶ 202 (June 27).

117. *See id.* at 106–08 ¶¶ 202–05; *see also* Maziar Jamnejad & Michael Wood, *The Principle of Non-Intervention*, 22 LEIDEN J. INTL. L. 345, 347–48 (2009) (quoting Nicar. v. U.S., 1986 I.C.J., to demonstrate the relationship between coercive intervention and state sovereignty).

118. *See* Buchan, *supra* note 38, at 218–19.

119. *See supra* Section III.A.1.

120. *See* G.A. Res. 56/83, Annex, Responsibility of States for Internationally Wrongful Acts, U.N. Doc. A/RES/56/83, art. 31 (Jan. 28, 2002) [hereinafter ARSIWA].

reducing the importance of repelling an attack through self-defense. Further, cyber-attacks are more likely than kinetic attacks to produce solely or predominately economic damage, making restitution of greater significance. However, in addition to the legal questions of threshold and attribution, seeking reparations will require the injured state to bring the matter before an appropriate tribunal that has both the jurisdiction to hear the case and the power to enforce its decisions. In the international arena, both issues are practically problematic.

B. *Self-Help and Cyber-Self-Defense*

As an alternative to pursuing a legal process in response to the cyber-attack, a victim state may instead employ a form of lawful self-help. That is, rather than becoming the basis for a legal claim, the internationally wrongful act may relieve the victim state of some of its international obligations.[121] The most prominent form of self-help is that of self-defense,[122] the primary focus of this Section. However, states may also employ countermeasures as a response to a cyber-attack.[123]

1. Establishing a Cyber-Attack

Although the properties of an armed attack have been discussed extensively by scholars, its precise nature remains uncertain even when considering kinetic attacks.[124] Definitional ambiguity is far more problematic when considering an armed attack, because the motivation for and consequences of the legal analysis are different.[125] When applied to *cyber*-armed attack, the threshold question only becomes less clear. Unlike a state that has been subject to illegal cyber-force, a state that seeks to establish itself as a victim of a cyber-armed attack very likely intends to respond forcefully in self-defense; otherwise there would be little point in seeking to establish the legal status of the attack with the higher threshold. The stakes in resolving the definitional ambiguity between cyber-force and cyber-armed attack are therefore considerably higher.

---

121. Int'l Law Comm'n, Rep. on the Work of Its Fifty-Third Session, U.N. Doc. A/56/10 at 71 [hereinafter *ARSIWA Commentary*].

122. *Id.* at 74; *see also supra* Section II.B.

123. ARSIWA, *supra* note 120, art. 22.

124. Randelzhofer, *supra* note 41, at 661–76.

125. Schmitt, *supra* note 33, at 604 (suggesting that the prohibition on the use of force has adapted better to the advent of cyber-attack than the notion of armed attack).

In contrast to the broad acceptance of Schmitt's criteria with respect to the use of force, there is no consensus regarding a test for cyber-armed attack.[126] Schmitt has proposed that, to qualify as cyber-armed attack, an incident must result in physical damage,[127] such as the destruction of data designed to be converted into tangible objects.[128] This test does not assist in delineating the boundary between a use of cyber-force and cyber-armed attack; it is very similar to the test for armed force which Schmitt presents earlier in the same paper.[129] Further, some academics view physical damage as the defining element of a use of cyber-force.[130]

Another simple and relatively uncontroversial test advocated by cyber-policy specialist Herbert Lin and others extends the existing notion of cyber-armed attack to encompass cyber-incidents.[131] That is, if the effects of a cyber-attack would be sufficient to qualify the attack as an armed attack if produced by kinetic means, then the cyber-incident will also qualify as a cyber-armed attack.[132] Professor Yoram Dinstein argues that, "From a legal perspective, there is no reason to differentiate between kinetic and electronic means of attack . . . . The crux of the matter is not the medium at hand . . . but the violent consequences of the action taken."[133] Similarly, national security law expert Matthew Waxman suggests there is at least a U.S. consensus that a cyber-attack will be an armed attack where its "features and consequences" closely resemble those of traditional military force.[134] This approach is essentially a continuation of the consequences-based model for interpreting rules framed in terms of the instrument employed.[135] However, in their article on cyber-warfare, John Dever and James Dever reject this approach, arguing that existing concepts are inadequate to prop-

126. *Id.* at 573.

127. *Id.* at 589.

128. *Id.* (giving banking information as an example of such data).

129. *Id.* at 573. Schmitt's test for armed force does not expressly exclude attacks with nonphysical consequences, but that is its natural implication. Schmitt suggests that all uses of armed force will constitute an armed attack. *See id.* at 575. However, he does not impliedly reference the doctrine that all unlawful uses of force may be construed as armed attacks because he structures his analysis in two separate sections, entitled "Uses of Force" and "Armed Attack." *See id.* at 573–78, 587–90.

130. Buchan, *supra* note 38, at 212; Silver, *supra* note 11, at 90.

131. Lin, *supra* note 82, at 63; *see also* Roscini, *supra* note 2, at 115.

132. Lin, *supra* note 82, at 73; *see also* Roscini, *supra* note 2, at 115.

133. Yoram Dinstein, *Computer Network Attacks and Self-Defence*, *in* 76 Computer Network Attack & International Law 99, 103 (Michael N. Schmitt & Brian T. O'Donnell eds., 2002).

134. Waxman, *supra* note 82, at 47.

135. *See supra* Sections II.A, III.A.1.

erly characterize the nuances of cyber-attack.[136] Schmitt notes that opinion is divided as to the nature of cyber-armed attack, particularly where an incident does not cause physical damage but nonetheless produces severe outcomes.[137] Emerging state practice suggests that the scope of armed attack in cyber-scenarios is likely to expand, although where these new boundaries will lie is far from clear.[138]

2. The Elements of Self-Defense

The writings of scholars in this field suggest that the law in relation to the necessity and proportionality elements of self-defense as applied to cyber-attack is relatively straightforward and settled. For example, Dinstein's analysis of this matter deals with necessity in a short paragraph, and is also brief when discussing proportionality.[139] Similarly, Rule 14 of the Tallinn Manual hints at few controversies.[140] However, these issues are not quite as straightforward as they might appear.

First, several issues emerge from the necessity element. The requirement of necessity implies that there is no alternative to a use of force to repel an attack that is either imminent or underway.[141] On this basis, the general absence of physical territorial incursion in the case of cyber-attacks suggests that the necessity element may be fundamentally difficult to establish.

The immediacy aspect of necessity may also be particularly problematic in relation to computer attacks. For these attacks, many basic parameters that are normally readily apparent may remain unclear for some time. For example, the beginning or even the very existence of the attack may not be initially identifiable, as in the case of malware or direct intrusion.[142] In other cases, denial of service attacks may occur intermittently, making it difficult to say whether these constitute a single attack and, if so, whether the attack is still ongoing. These properties significantly complicate the process of deciding when a response in self-defense would be sufficiently immediate and necessary. Obstacles to reliably ascer-

136. Dever & Dever, *supra* note 20, at 28.

137. Schmitt, *supra* note 3, at 282–84.

138. *Id.* at 283–84.

139. Dinstein, *supra* note 133, at 109–10.

140. *See* TALLINN MANUAL, *supra* note 2, at 61–63.

141. Jennings, *supra* note 59, at 89; TALLINN MANUAL, *supra* note 2, at 62 ¶¶ 2–3 ("The key to the necessity analysis in the cyber context is, therefore, the existence, or lack, of alternative courses of action that do not rise to the level of a use of force.")

142. TALLINN MANUAL, *supra* note 2, at 66 ¶ 10; McGhee, *supra* note 83, at 100.

taining attribution and characterizing the type of attack may introduce further delays.[143]

Second, proportionality raises additional concerns. Proportionality limits the force employed to only what is necessary to successfully mount a defense in the circumstances.[144] On this point, notably, scholars appear to be in general agreement that kinetic and cyber- uses of force are completely equivalent and interchangeable as defensive measures.[145] Dinstein makes this claim as a blanket statement, providing neither supporting authorities nor logical argument.[146] The expert authors of the Tallinn Manual also agree with this view: they repeat the claim several times in different contexts throughout the commentary on Rule 14.[147]

While it is generally accepted that the proportionality requirement does not limit an injured state to a response that is qualitatively the same as that employed by the aggressor,[148] it seems improbable and anomalous that the application of this functional model of proportionality to cyber-attack should be so simple and uncontentious. For instance, putting aside the questions of definition and attribution, if Iran had chosen to retaliate against Israel using kinetic force in response to the Stuxnet worm,[149] the proportionality of this response would be regarded by certain sectors of the international community as at least contentious purely because Iran was the first to resort to kinetic weaponry. This is not to suggest that a state which is the victim of a serious cyber-armed attack should be precluded from employing kinetic force in defense. However, contrary to the apparent views of scholars, it seems likely in those circumstances that the proportionality of any response would be legally contentious.

Finally, to assess the proportionality of any response, it is necessary to demarcate precisely what constitutes the initial cyber-armed

143. McGhee, *supra* note 83, at 81; Shackelford & Andres, *supra* note 10, at 998.

144. Jennings, *supra* note 59, at 89; TALLINN MANUAL, *supra* note 2, at 62–63 ¶ 5 (indicating that "[t]he criterion limits the scale, scope, duration, and intensity of the defensive response . . . . [but] [i]t does not restrict the amount of force used to that employed in the armed attack . . . .").

145. *See, e.g.*, Dinstein, *supra* note 133, at 108; TALLINN MANUAL, *supra* note 2, at 59–60 ¶ 3, 60 ¶¶ 5–6.

146. Dinstein, *supra* note 133, at 108.

147. TALLINN MANUAL, *supra* note 2, at 62–63 ¶¶ 3, 5–6.

148. RUYS, *supra* note 58, at 111.

149. For information on the Stuxnet worm, see generally Irving Lachow, *Stuxnet Enigma: Implications for the Future of Cybersecurity*, 11 GEO. J. INT'L AFF. 118 (2010); Richardson, *supra* note 3. Stuxnet is regarded as a use of force, rising to the level of an armed attack. *See* Buchan, *supra* note 38, at 220–21; Schmitt, *supra* note 2, at 45 ¶ 9, 58 ¶ 13.

attack. Dinstein suggests that the "accumulation of events" theory may apply to this problem.[150] However, this theory is far from universally accepted by scholars,[151] and has been rejected numerous times by the U.N. Security Council.[152] Although the ICJ has seemingly accepted the notion that an accumulation of events can constitute an armed attack,[153] the court has been more reluctant to conclude that an accumulation of events may be used when evaluating the proportionality of a given response.[154] The relatively amorphous nature of cyber-attacks further complicates the task of analyzing the magnitude of an initial armed attack to determine a suitably proportionate response.

3. The Significance of Economic Damage

A key difference between instances of kinetic aggression and cyber-attack is the focus on immediate physical damage in relation to the former and the heightened significance of economic damage to the latter.[155] Cyber-attacks can produce physical consequences, but so far these have been extremely rare.[156] Therefore, many cyber-attacks are likely to produce primarily or exclusively economic damage.[157] As noted, cyber-attacks lack the element of potentially escalating territorial incursion by military assets or personnel, diminishing the relevance of traditional state defense

150. Dinstein, *supra* note 133, at 109. The "accumulation of events" theory considers that a series of temporally proximate small-scale attacks may be viewed in aggregate for the purposes of assessing whether they exceed the relevant threshold, particularly in relation to armed attack. *Id.*

151. Ian Brownlie, *The Use of Force in Self-Defence*, 37 BRIT. Y.B. INT'L L. 183, 245 (1961).

152. Derek Bowett, *Reprisals Involving Recourse to Armed Force*, 66 AM. J. INT'L. L. 1, 6–8 (1972). *But see* DINSTEIN, *supra* note 62, at 202, 230–31; RUYS, *supra* note 58, at 174–75; Roberto Ago, *Addendum to the Eighth Report on State Responsibility*, 2 YEARBOOK OF THE INTERNATIONAL LAW COMMISSION 13, 69–70 (1980).

153. *See* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.) 1986 I.C.J. 14, ¶ 231 (June 27); Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. Rep. 161, ¶ 64 (Nov. 3); TALLINN MANUAL, *supra* note 2, at 56 ¶ 8.

154. Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda), Judgment, 2005 I.C.J. Rep. 168 ¶ 147 (Dec. 19).

155. Priyanka R. Dey, *"Use of Force" and "Armed Attack" Thresholds in Cyber Conflict: The Looming Definitional Gaps and the Growing Need for Formal U.N. Response*, 50 TEX. INT'L L.J. 381, 388–92 (2015).

156. Richardson, *supra* note 3, at 8 (noting that Stuxnet is regarded as the first cyber-attack aimed at a state that has caused tangible physical damage). More recently, an entirely digital attack caused substantial damage to a German steel factory, and this has been reported as the second cyber-attack with physical consequences. Kim Zetter, *A Cyber-attack Has Caused Confirmed Physical Damage for the Second Time Ever*, WIRED (Jan. 8, 2015, 5:30 AM), http://www.wired.com/2015/01/german-steel-mill-hack-destruction [https://perma.cc/XC28-MN6F].

157. *See* Hathaway, *supra* note 12, at 822–23, 840.

against such an incursion.[158] As a result, in many cases reparations may be of more value to an injured state than reprisals—even assuming the existence of a clear legal framework for the application of self-defense to cyber-attacks. Indeed, even attacks that intend or cause actual physical damage may be better remedied through reparation than self-defense. For example, it is unclear what form of response would constitute a meaningful act of self-defense to an attack such as Stuxnet. By comparison, reparations in the form of restitution or compensation are likely to be far more satisfactory to the wronged state than a retaliatory response.

4. Countermeasures

As an alternative to self-defense, countermeasures may be a suitable legal mechanism for states to respond to a cyber-attack. The law concerning countermeasures is relatively complex. The clearest statement on the law regarding countermeasures is that contained in the U.N. General Assembly Resolution on Responsibility of States for Internationally Wrongful Acts (ARSIWA).[159]

ARSIWA suggests that if a state directs otherwise illegal actions towards another state that has wronged it, the wrongfulness of those actions will be precluded if they qualify as countermeasures.[160] However, this response is subject to a variety of prerequisites and constraints.[161] First, the countermeasures must be for the purposes of inducing the responsible state to comply with its international obligations[162] and are limited to the nonperformance of an obligation.[163] Second, the countermeasures must be taken in a way so as to not interfere with resolution of the dispute,[164] and as much as possible allow for resumption of the state of affairs that existed previously.[165] Third, countermeasures must comply with peremptory norms of international law, expressly including the

158. *See supra* Subsection III.A.1; Brenner, *supra* note 105, at 404.

159. *See* ARSIWA, *supra* note 120, art. 22, 49–54; *see also ARSIWA Commentary*, *supra* note 121, at 128–39 (providing detailed commentary on countermeasures). The ICJ adopted the major principles outlined therein in the *Gabèíkovo-Nagymaros* case. Gabèíkovo-Nagymaros Project (Hung. v. Slovk.), Judgment, 1997 I.C.J. 7, ¶¶ 82–87 (Sept. 25).

160. ARSIWA, *supra* note 120, art. 22.

161. *Id.* arts. 49–52.

162. *Id.* art. 49(1); *see also* Hung. v. Slovk., 1997 I.C.J. at 7, ¶ 87 (mentioning that purpose of inducing wrongdoing State to comply with its obligations under international law is condition for lawfulness of a countermeasure).

163. ARSIWA, *supra* note 120, art. 49(2).

164. *See id.* art. 52 (describing requirements for countermeasures); *see also* Hung. v. Slovk., 1997 I.C.J. at 7, ¶ 84 (holding that injured State must have called upon State committing wrongful act to discontinue its wrongful conduct or to make reparation for it).

165. ARSIWA, *supra* note 120, art. 49(3).

prohibition against the use of force,[166] and they are subject to a test of proportionality.[167]

The question of how the law of countermeasures applies as a response to a cyber-attack has received comparatively little attention from scholars. For example, author Matthew Sklerov discusses "active defenses," whereby the victim of a cyber-attack responds in kind to an attempt to prevent future attacks or halt one that is ongoing.[168] While Sklerov argues that active defenses are lawful as self-defense, he proposes that countermeasures may be a suitable alternative theory to provide legal justification.[169]

There are several obstacles to understanding how the law of countermeasures would apply in the case of cyber-attack. The details of cyber-attacks are highly technical, and moreover cover a very wide range of possible mechanisms and outcomes. However, as can be seen from even the brief summary above, countermeasures have very specific requirements and constraints.[170] The specificity of the requirements for taking countermeasures mean that the nuances of a given response may be particularly important when ascertaining its legality. The technical subtleties represent a cognitive obstacle to applying this law to a given set of facts, while the wide range of properties make it difficult to identify general rules of application. Further, the classified nature of active defenses potentially limits the opportunity for both courts and scholars to examine the question in the first place.[171] Although not the subject of this Article, the applicability of countermeasures as a response to cyber-attacks is an area where there appears to be opportunity for further research.

IV. The Attribution Obstacle

The above Part summarizes some of the legal issues surrounding responses to cyber-attacks. However, arguably the most significant obstacle is that of cyber-attribution.[172] This Part outlines the

166. *Id.* art. 50.

167. *Id.* art. 51; *see* Hung. v. Slovk., 1997 I.C.J. at 7, ¶¶ 85–87 (holding countermeasure unlawful because it was not proportionate).

168. Sklerov, *supra* note 88.

169. *Id.* at 37.

170. *See* ARSIWA, *supra* note 120, arts. 49–54*; see also* Hung. v. Slovk., 1997 I.C.J. at 7, ¶¶ 82–87 (finding Czechoslovakia countermeasure unlawful after analysis of requirements for justifiable countermeasures).

171. Sklerov, *supra* note 88, at 25; Eric Talbot Jensen, *Computer Attacks on Critical National Infrastructure: A Use of Force Invoking the Right of Self-Defense*, 38 Stan. J. Int'l L. 207, 231 (2002).

172. Silver, *supra* note 11, at 78.

existing law of attribution and discusses why applying this law to cyber-attack is particularly problematic.

A. *The Law of Attribution*

A state's responsibility for a breach of international law is contingent upon the breach being attributable to that state.[173] If an injured state wishes to seek reparations for the results of a cyber-attack by another state, it will need to establish attribution. Alternatively, if the state intends to retaliate in self-defense, attribution is arguably even more important. A hypothetical scenario wherein one state could effectively "frame" another by routing cyber-attacks through systems based within the second state's territory illustrates the importance of attribution and the dangers that cyber-misattribution poses.[174] As will be discussed, misattribution is far more likely in relation to cyber-attacks than to traditional kinetic attacks.[175]

1. Legal Tests for Attribution

ARSIWA Articles 4–11 outline the basic rules of attribution.[176] Fundamentally, a state cannot act except through its agents.[177] It is important not to cast the net too wide when delineating whose acts may be attributable to a state, however. Although actions of private individuals are generally not attributable to a state,[178] ARSIWA Article 8 provides that private conduct will be attributable to a state if those persons act under instruction or the control of that state.[179] The relative availability and cheapness of the technology necessary to mount a cyber-attack makes significant attacks launched by private individuals or corporations far more feasible compared with other types of warfare.[180] The rules concerning attribution of the actions of nonstate actors are therefore of particular relevance to cyber-attack.

---

173. ARSIWA, *supra* note 120, art. 2.

174. Levi Grosswald, *Cyberattack Attribution Matters Under Article 51 of the U.N. Charter*, 36 Brook. J. Int'l L. 1151, 1151–52 (2011).

175. *See infra* Section IV.B.

176. ARSIWA, *supra* note 120, arts. 4–11.

177. German Settlers in Poland, Advisory Opinion, 1923 P.C.I.J. (ser. B) No. 6, at 22 (Sept. 10).

178. *ARSIWA Commentary*, *supra* note 121, at 47.

179. ARSIWA, *supra* note 120, art. 8.

180. *See* Tallinn manual, *supra* note 2, at 31 ¶ 8.

There are two possible standards for the degree of control required to establish attribution.[181] In the *Nicaragua* case, the ICJ held that a state must have had "effective control" over the non-state actor at the point where the alleged breaches occurred.[182] In that case, the ICJ reasoned that general control was insufficient to impute responsibility for specific actions taken by paramilitary groups, even though these groups were at times completely dependent on U.S. support.[183]

In contrast, in the case of *Prosecutor v. Tadić*, the International Criminal Tribunal for the Former Yugoslavia (ICTY) held that, where a hierarchically-organized group is under the "overall control" of a state, this will be sufficient to attribute the group's actions to that state.[184] Further, the ICTY indicated that the degree of control necessary may vary according to the facts of the case, such that not every case will require a high threshold.[185] This is a much broader test than that articulated in *Nicaragua*, and potentially expands the scope of state responsibility significantly.[186]

There is doubt concerning the overall control test as aspects of the ICTY's dictum have been expressly and comprehensively rejected by the ICJ.[187] Nonetheless, the overall control test is frequently referred to in the academic literature on attribution,[188] and represents valuable jurisprudence when considering how the law of attribution could adapt to emerging issues such as cyber-attack.

Further enlarging the legal hurdle of attribution is the standard of proof required to establish this control. Historically there has

---

181. *ARSIWA Commentary*, *supra* note 121, at 47.

182. Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. 14, ¶ 115 (June 27).

183. *Id.* at 64–65.

184. Prosecutor v. Tadić, Case No. IT-94-1-A, Judgment, ¶ 120 (Int'l Crim. Trib. for the Former Yugoslavia, Appeals Chamber, July 15, 1999).

185. *Id.* ¶ 117.

186. *See* Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 43, ¶ 406 (Feb. 26).

187. *Id.* ¶¶ 401–07.

188. *See, e.g.*, Vincent-Joël Proulx, *Babysitting Terrorists: Should States Be Strictly Liable for Failing to Prevent Transborder Attacks?*, 23 Berkeley J. Int'l L. 615, 621 (2005); Jan Arno Hessbruegge, *The Historical Development of the Doctrines of Attribution and Due Diligence in International Law*, 36 N.Y.U. J. Int'l. L. & Pol. 265, 303–06 (2003); Peter Margulies, *Sovereignty and Cyber Attacks: Technology's Challenge to the Law of State Responsibility*, 14 Melb. J. Int'l L. 496, 507 (2013); Shackelford & Andres, *supra* note 10, at 986; Graham, *supra* note 10, at 95; Roscini, *supra* note 2, at 100; Grosswald, *supra* note 174, at 1161; *see also ARSIWA Commentary*, *supra* note 121, at 48.

been a degree of uncertainty regarding standards of proof required for matters brought before the ICJ.[189] The ICJ is expressly required only to "*satisfy itself* . . . that . . . [a] claim is *well founded* in fact and law."[190] In *Nicaragua*, the ICJ referred to "convincing evidence";[191] however, this sheds little light on the required standard.

Overall, the ICJ's judgments suggest that the standard of proof will be commensurate with the seriousness of the allegation.[192] This principle was enunciated in an early case where the ICJ required "decisive legal proof" and "conclusive evidence."[193] Much more recently the more relaxed standard of "balance of probabilities" was applied to a less weighty question concerning the location of a provincial boundary.[194]

Possibly contrary to such a principle, in *Oil Platforms*, the ICJ considered that attributing responsibility for the launch of a missile that had sunk a ship depended upon merely the "balance of evidence."[195] However, it is not clear whether this phrase was intended to express a formal legal standard of proof or merely described a general outcome that remained dependent upon the weight assigned to the evidence submitted.[196]

The apparent lack of a clear rule regarding the standard of proof required[197] culminated in criticism that such a standard should be properly articulated to assist parties appearing before the ICJ.[198] The ICJ responded by specifically addressing the question of standard of proof in the *Bosnian Genocide* case, restating the principle from *Corfu Channel* that "charges of exceptional gravity" require a high standard of proof, and expressly noting that this

189. *See* Anna Riddell & Brendan Plant, Evidence Before the International Court of Justice 129 (2009); Ruth Teitelbaum, *Recent Fact-Finding Developments at the International Court of Justice*, 6 L. & Prac. Int'l Cts. & Tribunals 119, 124, 127–29 (2007).

190. Statute of the International Court of Justice art. 53 (emphasis added).

191. Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.) Judgment, 1986 I.C.J. Rep. 14, ¶ 29 (June 27).

192. *See* Riddell & Plant, *supra* note 189, at 132.

193. Corfu Channel (U.K. v. Alb.) 1949 I.C.J. 4, 16–17 (Apr. 9).

194. Land, Island and Maritime Frontier Dispute (El Sal. v. Hond.: Nicar. Intervening), Judgment, 1992 I.C.J. 351, ¶ 248 (Sept. 11).

195. Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. 161, ¶ 57 (Nov. 3).

196. Teitelbaum, *supra* note 189, at 127–29.

197. *See* Riddell & Plant, *supra* note 189, at 129.

198. *See* Oil Platforms (Iran v. U.S.), Judgment, 2003 I.C.J. 161, ¶ 30–39 (Nov. 3) (Judge Higgins); *id.* ¶¶ 41, 44 (Judge Buergenthal).

principle extends to attribution.[199] Specifically, in *Bosnian Genocide*, the ICJ required that involvement with the crime of genocide be "established beyond any doubt."[200] Therefore, it seems that in the event of a serious cyber-attack, a victim state would likely be required to meet an exacting standard of proof.

2. Bypassing Attribution

Further complicating the question of attribution, since the September 11 terrorist attacks, state practice appears to have expanded to incorporate an alternative notion of responsibility to that articulated in ARSIWA.[201] Professor Vincent-Joël Proulx argues that states may be held indirectly responsible for the acts of private individuals that breach international law, even when there is no causal link between an action of the state and that breach.[202] Although expressed in the context of terrorist attacks, in formulating his approach Proulx nonetheless appears to have contemplated the difficulties of applying traditional attribution where technology such as the Internet is involved.[203]

Proulx's notion of indirect responsibility is somewhat similar to vicarious responsibility,[204] albeit somewhat broader, as indirect responsibility does not require the state to have knowledge of the unlawful activities occurring within its own borders.[205] Essentially, Proulx's proposition is one of strict liability that dispenses with the need to establish attribution. Instead, the accused host state bears the onus of demonstrating why it should not be liable.[206]

B. *Problems with Cyber-Attribution*

1. Technical Attribution

Legal attribution is entirely dependent on the ability to satisfactorily answer highly technical questions concerning the origin of a

199. Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 43, ¶ 209 (Feb. 26) [hereinafter Bosnian Genocide]; *see also* Teitelbaum, *supra* note 189, at 127–28.

200. Bosnian Genocide, 2007 I.C.J. ¶ 422.

201. *See* Proulx, *supra* note 188, at 617.

202. *See id.* at 624.

203. *See id.* at 617.

204. *Id.* at 624 n.43. *See generally* 1 Oppenheim's International Law 501–02 (Robert Jennings & Arthur Watts eds., 9th ed. 1996) (explaining the meaning of vicarious liability).

205. Proulx, *supra* note 188, at 624; Davis Brown, *Use of Force Against Terrorism After September 11th: State Responsibility, Self-Defense and Other Responses*, 11 Cardozo J. Int'l & Comp. L. 1, 13 (2003).

206. Proulx, *supra* note 188, at 656–57.

particular attack.[207] Determination of the relationship between the attacker and responsible state is ultimately necessary. However, this is likely to first entail determination of the true geographic origin of an attack and the identity of those responsible.

Establishing these facts requires overcoming significant technical evidentiary hurdles. Attackers have at their disposal a variety of techniques to hide their identity and location, including use of various cryptographic mechanisms and routing their attacks through multiple compromised systems belonging to third parties.[208] For example, the series of cyber-attacks on Estonia have been reported to originate from at least 177 countries, and as well from within Estonia itself.[209]

Given that cyber-attackers can work in comparatively small groups, or even as individuals, the means to establish their presence are very different from those where nonstate actors such as terrorists or insurgents establish a base in an unaffiliated host state. Further, quite unlike the weapons necessary to inflict traditional kinetic force, cyber-attackers require only commodity computer systems that can be easily, cheaply, and covertly acquired.

Dinstein suggests the challenge of technical attribution is merely a temporary obstacle that will likely be overcome with inevitable technological improvement.[210] However, most scholars expressly or impliedly reject this view. Some note that the architecture of the Internet makes identification inherently difficult.[211] The very serious difficulty of technical attribution is widely acknowledged,[212] and attribution is inherently the most significant practical obstacle to addressing cyber-attack under public international law.[213] Indeed, some academics go as far as to suggest the problem of cyber-attribution may be impossible to solve.[214]

---

207. *See* David A. Wheeler & Gregory N. Larsen, Inst. For Def. Analyses, Techniques For Cyber Attack Attribution 2, 9 (2003).

208. Shackelford & Andres, *supra* note 10, at 981–83; Margulies, *supra* note 188, at 503–04.

209. Schmitt, *supra* note 33, at 569–70.

210. Dinstein cites no source to support this view. Dinstein, *supra* note 133, at 112. Roscini makes a similar statement, but cites only Dinstein for support. Roscini, *supra* note 2, at 97.

211. Graham, *supra* note 10, at 92; Grosswald, *supra* note 174, at 1177.

212. Silver, *supra* note 11, at 78–79 (suggesting that attribution may be the most significant practical obstacle to the development of the law in this area); Lin, *supra* note 82, at 77; Waxman, *supra* note 10, at 443–45.

213. Silver, *supra* note 11, at 78–79.

214. Waxman, *supra* note 82, at 50; Graham, *supra* note 10, at 92; Shackelford & Andres, *supra* note 10, at 981.

Contrary to the minority view expressed by Dinstein,[215] a solution is likely not a matter of waiting for technology to improve. In a sense, computer attackers and defenders are engaged in an arms race. However, the defenders will generally be at least one step behind. Attackers take the initiative in developing new mechanisms, which defenders must then identify, analyze, and defeat. These circumstances will generally favor the attacker, and a sufficiently well-resourced and motivated attacker can be expected to find ways to conceal their involvement and ultimate location.[216] It may be that complete certainty as to the identity of the party responsible for a given attack will never be attainable.[217]

Depending upon the standard of proof required, without some dramatic and fundamental change in the nature of the technology involved, such as the architecture of the Internet, ascertaining even the most basic facts as to the identity and origin of an attacker will probably remain at least very difficult. A high standard of proof is likely to make this task completely infeasible.[218] If a lower standard applies, political motivation may be a more useful indicator of the origin of an attack than technical evidence.[219] However, these difficulties raise the very real risk of one state covertly implicating another innocent state in an attack to damage relations or provoke conflict with the injured state.[220]

2. Legal Attribution

It follows from the extreme difficulty of establishing the basic facts surrounding technical attribution that establishing the facts to meet the requirements for legal attribution is even more problematic. Depending upon the standard required, proving the nuances of the relationship between the attacker and a state—the control exerted by the latter over the former at the relevant times—is likely to be extraordinarily difficult. A high standard of proof is likely to make this task impossible in most cases.

Thus, the ICJ's conservative approaches in *Nicaragua* and *Bosnian Genocide*[221] may simply be inappropriate when dealing with cyber-

215. Dinstein, *supra* note 133, at 112.

216. *See* Schmitt, *supra* note 33, at 570. As an example, the attacks on Estonia have been reported to have originated from at least 177 countries, and from within Estonia itself. *Id.*

217. *See* Waxman, *supra* note 82, at 50.

218. *Id.*

219. Shackelford & Andres, *supra* note 10, at 992.

220. *See* Grosswald, *supra* note 174, at 1151–52.

221. *See generally* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v.

attacks. To illustrate, one analysis of how existing attribution rules would apply to two publicly known cyber-attacks[222] concluded that attribution could not be established for either under the effective control test, and would still be difficult using the overall control test.[223]

Schmitt has therefore advocated a low standard—an indirect responsibility approach whereby a state will be responsible if it fails to take reasonably available measures to stop cyber-attacks originating in its territory.[224] He further suggests that, as long as a state takes reasonable steps to identify the perpetrator of a cyber-armed attack, it may respond forcefully in self-defense, and that it will not matter legally if the retaliating state turns out to be mistaken as long as it acted on the best available information.[225]

Sklerov agrees, arguing that the situation should be judged on the facts at hand and, even if misattribution results, the injured state will have met its international obligations as long as it acts in good faith.[226] Under this model, a state that does not comply with its international duty to prevent the use of its territory for cyber-attack will have assumed the consequent risk.[227] This is a dangerous approach that appears to lack a fundamental understanding of the nature of the threat posed by cyber-attack. All states would face extraordinary practical difficulties in ensuring that no computer system within their territory is ever used for purposes that may constitute a use of force or armed attack, either by an authorized user resident in that country or remotely by an unauthorized party operating extraterritorially. Such an approach would mean that two or more states could be easily and rapidly drawn into a cyber-war based on misattribution, with the ever-present risk of escalation to kinetic warfare.[228]

Although framed in relation to terrorism, rather than cyber-warfare, Proulx's strict liability model could be applied to cyber-attack. By reversing the burden of proof, strict liability dispenses with attribution by requiring that the host state show that it has done what is

---

U.S.), 1986 I.C.J. 14, ¶ 115 (June 27); Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 43, ¶ 422 (Feb. 26).

222. Shackelford & Andres, *supra* note 10, at 991–92.

223. *Id.*

224. Schmitt, *supra* note 33, at 580.

225. *Id.* at 595.

226. Sklerov, *supra* note 88, at 77–78.

227. *Id.*

228. *See* Grosswald, *supra* note 174, at 1151–52.

reasonable in trying to prevent breaches perpetrated by those acting within their territory. Proulx argues that this approach incentivizes states to deal with problems within their own borders.[229] However, setting a low standard for legal attribution, or discarding it entirely, risks the likelihood of an innocent state being wrongly implicated in an attack. This outcome would be particularly problematic if the incident rose to the level of an armed attack and the victim state had responded in self-defense.

V. Resolving the Problem of Cyber-Attribution

A. *Summary of the Standard-for-Attribution Problem*

The technical properties of cyber-attack make attribution extremely problematic, to the extent that even identifying the ultimate location of the parties involved may be extraordinarily difficult.

The established effective control test sets an extremely high bar. Given the technical obstacles, generally proving that a set of cyber-attackers were under the effective control of a state at the relevant time is likely impossible. The overall control test, while much broader in theory, may not lead to a different outcome in the case of cyber-attack.[230] Further, the ICJ in *Tadić* held that this test only applies where a group has a defined, hierarchical structure.[231] The test is suited to quasi-military groups, but its application to small groups of hackers or individuals working in isolation is not clear. The rejection of this test by the ICJ also casts doubt on its legal validity.[232]

Both of these tests would likely present a serious obstacle to the ability of a state to take lawful retaliatory action in response to a cyber-attack. By establishing and upholding such a demanding test as effective control, the ICJ apparently seeks to limit when states can take such action. As a general question of policy, it would seem sensible to limit the opportunities for states to lawfully use force in response to the actions of private individuals. Such a principle applies similarly to retaliation due to cyber-attack, and is of particular importance where that response might include kinetic

229. Proulx, *supra* note 188, at 656.

230. Shackelford & Andres, *supra* note 10, at 991–92.

231. Prosecutor v. Tadić, Case No. IT-94-1-A, Judgment, ¶ 120 (Int'l Crim. Trib. for the Former Yugoslavia, Appeals Chamber, July 15, 1999).

232. Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 43, ¶¶ 401–07 (Feb. 26).

force. Consequently, the technical difficulties concerning attribution are in some respects an argument for maintaining the existing restrictive rules to avoid misdirected retaliation. While problematic, caution should therefore be exercised in contemplating whether to discard this high standard for attribution if there is the prospect of use of kinetic force.

The practical application of the high standard for attribution to cyber-attack is hugely difficult; consequently, scholars argue that states ought to be able to exercise their inherent right to self-defense in relation to a cyber-armed attack either based upon the best information available at the time or bypassing the need for attribution entirely.[233] More cautious scholars note the tremendous risk of escalation and conflict inherent in permitting retaliatory action if strict attribution requirements are not maintained.[234]

Grosswald, for example, analyzes both the jurisprudence-based control tests and Proulx's strict liability model in relation to cyber-attack.[235] He dismisses Proulx's approach as clearly "untenable" in the context of cyber-self-defense due to the risk of misattribution.[236] Further, while the notion of holding a state responsible for activities occurring within its borders may be suitable when dealing with large groups of terrorists or insurgents who have established bases and accumulated weapons, it is less applicable to small, agile, covert groups of cyber-attackers who require only a computer and network connection.

It is neither practical nor desirable to require a state to monitor and control the use of every computer in its jurisdiction simply as a matter of course. Even states that enforce rigorous controls on Internet usage are not able to reliably prevent nontechnical users from accessing specific, easily identifiable websites.[237] Sophisticated cyber-attackers working within a state's territorial borders could easily do so undetectably, even with the kinds of very restrictive Internet controls that would be politically and culturally unpalatable in many countries.

Given the apparent infeasibility of establishing sufficient control and the dangers associated with strict liability, Grosswald argues that the obstacles to cyber-attribution should be resolved by

233. Sklerov, *supra* note 88, at 77–78; Schmitt, *supra* note 33, at 595.

234. Roscini, *supra* note 2, at 100; Grosswald, *supra* note 174, at 1175.

235. Grosswald, *supra* note 174, at 1159–61, 1165–66.

236. *Id.* at 1165–66.

237. *E.g.*, Roman Loyola, *How to Break through the Great Firewall of China on iOS*, Macworld (Oct. 10, 2013), http://www.macworld.com/article/2050501/how-to-break-through-the-great-firewall-of-china-on-ios.html [https://perma.cc/4C44-665G].

encouraging cooperation and collaboration between states.[238] However, Grosswald provides no real indication on how such an outcome could be achieved.[239]

Both the high- and low-standard approaches are therefore problematic. The relative ease with which a state could be implicated in a cyber-attack supports Grosswald's view that it is untenable to bypass the need to attribute responsibility for a cyber-armed attack before responding in self-defense.[240] Further, establishing a presumption of state responsibility for failure to prevent detrimental activities within its borders seems ill-suited to cyber-attack as it would effectively hold the state to an impractically high standard. Dispensing with attribution in this context would lead to a high risk of serious conflict if the position of many scholars is adopted whereby cyber- and kinetic forces are regarded as entirely equivalent and interchangeable for the purposes of evaluating the proportionality of an action taken in self-defense.[241]

On the other hand, requiring a high standard of attribution leads to problems of its own. The apparent unsuitability of the existing rules of international law has led to an effective practical lacuna. If a state cannot prove who is responsible for a cyber-attack it has suffered, retaliation in kind may be the safest option because the target of that retaliation will face similar obstacles. Consequently, the apparent inapplicability of existing rules arguably encourages states to engage in covert, effectively lawless, low-level cyber-exchanges. As this scenario may well produce serious conflict at some point, perpetuating it through impractically onerous requirements for state responsibility is not a sustainable position either.

On its face, Grosswald's view that cyber-attribution can be resolved through "increased state collaboration and sharing of information"[242] seems unrealistic. A state may well cooperate if its territory is unwittingly used to launch a cyber-attack in which it has no involvement. However, if the host state was involved in or supportive of the attack, genuine cooperation with the victim is implausible. In this scenario, the injured state depends upon the cooperation of the perpetrator to establish that party's legal responsibility. Without such cooperation, establishing effective

238. Grosswald, *supra* note 174, at 1155.
239. *Id.* at 1155.
240. *Id.* at 1165–66.
241. Schmitt, *supra* note 2, at 54–55 ¶ 3, 60 ¶¶ 5–6; Dinstein, *supra* note 133, at 108.
242. Grosswald, *supra* note 174, at 1180.

control to even a modest standard of proof is likely to be extremely difficult.

B. *Determining Attribution Requirements Based on State Response*

Employing the correct mechanism to establish cyber-attribution is of great importance. Getting the balance wrong may result in either increased likelihood of escalation of international tension and conflict caused by misdirected retaliation due to flawed attribution or, alternatively, rendering law-abiding states unable to respond to cyber-attack because legal attribution is impossible to establish.[243] Attribution is a prerequisite to any lawful response, and this creates a practical lacuna whereby states are effectively precluded from any such response because of the infeasibility of cyber-attribution.

This Article proposes that these conflicting requirements can be resolved by determining attribution requirements based upon the course of action the victim state chooses to pursue. Grosswald correctly notes that circumventing attribution in the context of self-defense would undermine the purpose of Article 51, and indeed one of the general goals of the Charter, namely to constrain the ability of states to lawfully use force.[244] In the case of cyber-defense, therefore, the status quo should be maintained and states should continue to be required to meet a strict test—such as effective control—when making out attribution for the purposes of self-defense. Irrespective of the general suitability of such a test in response to kinetic attacks,[245] the uncertainty inherent in establishing cyber-attribution demands that the use of force in response is only permitted where there is a high degree of confidence that the force is directed at the correct target. This position is particularly important if cyber- and kinetic responses are considered entirely interchangeable when evaluating proportionality. If a low bar is set for attribution, catastrophic scenarios such as that discussed by Grosswald are plausible,[246] and perhaps likely. The technical difficulties in attribution therefore support maintenance of strict rules to avoid misdirected retaliation.

---

243. Christopher C. Joyner & Catherine Lotrionte, *Information Warfare as International Coercion: Elements of the Legal Framework*, 12 Eur. J. Int'l L. 825, 853 (2001).

244. Grosswald, *supra* note 174, at 1175.

245. *See generally* Ruys, *supra* note 58, at 424–26, 433–43, 447–57 (for state practice in relation to attribution of armed attacks by private actors); Yutaka Arai-Takahashi, *Shifting Boundaries of the Right of Self-Defence — Appraising the Impact of the September 11 Attacks on Jus Ad Bellum*, 36 Int'l L. 1081, 1096–98 (2002).

246. Grosswald, *supra* note 174, at 1151–52.

The above might suggest that strict liability is completely inapplicable to cyber-attack. However, this model still has a role to play. Grosswald's hoped-for increased state collaboration seems idealistic, but is potentially achievable through strict liability. While the risks of strict liability in relation to attribution for self-defense are significant, reversing the burden of proof for attribution when a state seeks to address a breach of international law through legal processes would be a far more successful approach.

Under this model, it would then be in an accused host-state's best interests to cooperate with the injured state to resolve the matter. Although cooperation in relation to cyber-attack scenarios is most likely to occur after the fact, that sequence is appropriate due to the speed at which cyber-attacks can take place, and also because of the difficulty a state is likely to have in ascertaining the existence of a cyber-threat within its territory prior to an incident. In the context of terrorism, Proulx notes that a host state which has sacrificed a degree of sovereignty by allowing foreign forces into its territory will be deemed to have significantly reduced its burden when mounting a defense against strict liability for the actions of private parties.[247]

Following this reasoning, when pursuing cyber-attacks as a wrongful act under a regime of strict liability, genuine cooperation by the host state in assisting the injured state to collect evidence might similarly discharge this obligation.[248] Thus, states which pursue peaceful, legal processes for resolving disputes in relation to cyber-attacks should have the benefit of eschewing the stringent attribution tests normally required. Such a model encourages states to refrain from retaliation in self-defense, and instead pursue the matter through peaceful and measured legal processes. This approach is also likely to be preferred by cyber-attack victims where economic damage is far more prevalent than physical damage or territorial incursion.

Conceptually, effective control concerns positive actions. That is, a state that has effective control over a group is actively involved with activities of that group. Conversely, strict liability is notionally more concerned with omissions. For example, a state that fails to act to prevent a group's use of its territory as a launching pad for a terrorist attack might be held strictly liable for this omission. By instead determining attribution requirements based upon the vic-

247. Proulx, *supra* note 188, at 663.

248. Jennings & Watts, *supra* note 204, at 502 (discussing vicarious responsibility, and ways to fulfill the obligations, attributed to states for actions of non-state actors).

tim state's response, the proposed model dispenses with this distinction. The motivations for adoption of a new application of these concepts in the case of cyber-attack are the technical obstacles to establishment of attribution. Typically, strict liability seeks to avoid a circumstance where an injured party could not have done anything to protect itself. Under the proposed model, strict liability instead serves as a mechanism to elicit cooperation where, without it, a victim would be similarly defenseless, having no practical lawful recourse. Used in this way, strict liability seeks to resolve the scenario where the injured party otherwise could neither protect itself before the fact, nor receive justice after it. Application of strict liability in this way is a logical extension of these existing legal principles in relation to the emerging factual circumstances of cyber-attack.

Conclusion

This Article has outlined some of the major legal issues in relation to cyber-attacks, focusing particularly on the question of attribution. The novel characteristics of cyber-attacks make the existing standards of proof and degrees of control required to establish attribution extremely difficult to determine. As a result, attribution is regarded by many scholars as one of the most significant and immediate practical obstacles to resolving the legal uncertainty surrounding this emerging issue.[249] This Article has analyzed the different approaches to cyber-attribution and proposed a model whereby attribution requirements are linked with the state's response to a cyber-incident. This approach leverages existing international law, discourages states from pursuing retaliatory responses, and incentivizes host states to assist the victims of cyber-attack in identifying the perpetrators. Issues remain regarding compulsory jurisdiction and enforcement of international tribunal decisions on these issues. Furthermore, the applicability of the law on countermeasures to cyber-attack requires further research. However, the approach described represents a viable way to address existing concerns regarding attribution. If the international community adopts this model and it proves successful, that may enable a greater focus on other impediments to effectively dealing with the problem of cyber-attack.

249. *See supra* subsection IV.B.1.