# The Law of Attribution: Rules for Attributing the Source of a Cyber-Attack

## Introduction

Long after the conclusion of the 2016 presidential election in the United States, the story of Russian hacking has lived on. Public reports of Russian interference with the election first arose on June 14, 2016, when the Washington Post reported that Russian agents had compromised the Democratic National Committee's information systems, leaking their internal reports and emails to the public.[1] After subsequent investigations, the Department of Homeland Security and Director of National Intelligence James Clapper announced on October 7, 2016, that the U.S. intelligence community was "confident that the Russian Government directed the recent compromises."[2] Intelligence leaks to the New York Times and Washington Post in December later confirmed that the instances of Russian hacking were acts intentionally launched to sway the outcome of the election towards Trump.[3] Though seventeen American agencies agree that Russia is responsible for hacking the DNC and Clinton campaign,[4] then president-elect Trump continued to deny the fact of Russian interference, [5] only acknowledging that possibility

[1] *See* Ellen Nakashima, *Russian Government Hackers Penetrated DNC, Stole Opposition Research on Trump*, Wash. Post (Jun. 14, 2016), https://www.washingtonpost.com/world/national-security/russian-government-hackers-penetrated-dnc-stole-opposition-research-on-trump/2016/06/14/cf006cb4-316e-11e6-8ff7-7b6c1998b7a0_story.html?utm_term=.046a4916d0d6 (last visited Jan. 13, 2017).

[2] *See Joint Statement from the Department Of Homeland Security and Office of the Director of National Intelligence on Election Security*, DHS Press Office (Oct. 7, 2016), https://www.dhs.gov/news/2016/10/07/joint-statement-department-homeland-security-and-office-director-national.

[3] *See* Adam Entous, Ellen Nakashima & Greg Miller, *Secret CIA Assessment Says Russia was Trying to Help Trump Win White House*, WASH. POST (Dec. 9, 2016), https://www.washingtonpost.com/world/national-security/obama-orders-review-of-russian-hacking-during-presidential-campaign/2016/12/09/31d6b300-be2a-11e6-94ac-3d324840106c_story.html (last visited Jan. 13, 2017); David E. Sanger & Scott Shane, *Russian Hackers Acted to Aid Trump in Election, U.S. Says*, N.Y. TIMES (Dec. 9, 2016), https://www.nytimes.com/2016/12/09/us/obama-russia-election-hack.html (last visited Jan. 13, 2017).

[4] *See* Domenico Montaro & Brian Naylor, *On Intelligence and Russian Hacking, Are Trump and his Team Missing the Point?*, NPR (Jan. 6, 2017, 11:12 AM), http://www.npr.org/2017/01/06/508520414/on-intelligence-and-election-hacking-trump-and-his-team-continue-to-miss-the-poi (last visited Jan. 7, 2017).

[5] During the second presidential debate, Trump dismissed the idea of Russia being responsible for the hack of DNC. He continued making such statements in December after he had won the election, saying in an interview that reports

for the first time on January 11, 2017.[6] Even Russian presidential spokesman, Dmitry Peskov, responded by declaring that the United States "should either stop talking about [Russia being responsible for the DNC hack] or produce some proof at last."[7]

Although the Office of the Director of National Intelligence has since publicly published its most detailed report concluding that Russia was responsible for the DNC hack, the twenty-five page report says little about the evidence the agencies have establishing Russia's involvement in the hacks.[8] Even though U.S. intelligence agencies may have legitimate reasons for withholding the basis for their attribution, [9] absent the presentation of their evidence, the subsequent space of uncertainty has allowed many across the political spectrum to question the validity of the claim put forth by U.S. intelligence agencies.[10] Continued doubt about such attribution has served to frustrate the possibility of more forward-looking discussions on how to respond to such cyber-attacks, and muddles the picture for future policy decisions.

---

of Russian hacking were "ridiculous" and that U.S. intelligence had "no idea" if Russia was behind the hacking. *See* Justin Fishel & Veronica Stacqualursi, *A Timeline of Russia's Hacking into US Political Organizations Before the Election*, ABC NEWS (Dec. 15, 2016, 1:01 PM), http://abcnews.go.com/Politics/timeline-russias-hacking-us-political-organizations-ahead-election/story?id=44140526 (last visited Jan. 13, 2017).

[6] *See* David Nakamura & Abby Phillip, *Trump Acknowledges Russian Involvement in Meddling in U.S. Elections*, WASH. POST (Jan. 11, 2017), https://www.washingtonpost.com/politics/trump-cites-kremlin-statement-to-deny-reports-of-russia-ties-asks-if-we-are-living-in-nazi-germany/2017/01/11/a710f2b4-d777-11e6-b8b2-cb5164beba6b_story.html (last visited Jan. 13, 2017).

[7] *See* Laura Smith-Spark, *Russia Challenges US to Prove Campaign Hacking Claims or Shut Up*, CNN (Dec. 16, 2016, 4:49PM), http://edition.cnn.com/2016/12/16/europe/russia-us-hacking-claims-peskov/index.html (last visited Dec. 20, 2016).

[8] *See* David A. Graham, *An Intelligence Report that Will Change No One's Mind*, ATLANTIC (Jan. 6, 2017), https://www.theatlantic.com/politics/archive/2017/01/odni-report-on-russian-hacking/512465/ (last visited Jan. 17, 2016).

[9] It's entirely possible, if not probable, that much of the evidence they have acquired may be derived from covert intelligence operations, and the agencies may not have a method of revealing such evidence without revealing the corresponding covert operations. Such a problem is discussed *infra* 51-56.

[10] Sam Biddle, *Here's the Public Evidence Russia Hacked the DNC – It's Not Enough*, Intercept (Dec. 14, 2016, 8:30 AM), https://theintercept.com/2016/12/14/heres-the-public-evidence-russia-hacked-the-dnc-its-not-enough/ (last visited Jan. 14, 2017); Catherine Herridge & Pamela K. Browne, *'Guccifer' Casts Doubt on Obama Administration's Russia Hacking Claims*, Fox News (Jan. 4, 2017), http://www.foxnews.com/politics/2017/01/04/guccifer-casts-doubt-on-obama-administrations-russia-hacking-claims.html (last visited Jan. 13, 2017).

This situation captures the seriousness of the threats facing a country's cybersecurity, and the equally important task of creating a legal structure for attributing attacks to those who are responsible. Cyber-attacks[11] have the potential to cause significant and wide-ranging harm across a number of critical arenas. These attacks include targeted attacks against nuclear infrastructure (Stuxnet),[12] attacks against commercial entities (the Sony hack),[13] attacks against government infrastructure (the Estonia DDOS attack),[14] and attacks against the infrastructure of the internet itself (the Mirai botnet attack).[15] The threat posed by these attacks even prompted Director of National Intelligence James Clapper to note that in 2013, cyber-attacks surpassed terrorism on the United States' list of national threats.[16] And such cyber-attacks show no sign of abating, as the recent DNC hack demonstrates. While the persistence of cyber-attacks may be due, in part, to their relatively low cost,[17] it may also largely result from the difficulty in tracing these attacks to their source. As a result, cyber-attacks provide a perfect venue for actors to engage in malicious activity without fear of attribution or retribution, allowing them to strike with impunity.

---

[11] By cyber-attack, I refer to the definition used by Oona Hathaway as "any action taken to undermine the functions of a computer network for a political or national security purpose." *The Law of Cyber-Attack*, 100 Cal. L. Rev. 817, 826 (2012).

[12] *See* Kim Zetter, *An Unprecedented Look at Stuxnet, the World's First Digital Weapon*, WIRE (Nov. 3, 2014, 6:30AM), https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/ (last visited Dec. 20, 2016).

[13] *See* Andrea Peterson, *The Sony Pictures Hack, Explained*, WASHINGTON POST (Dec. 18, 2014) https://www.washingtonpost.com/news/the-switch/wp/2014/12/18/the-sony-pictures-hack-explained/?utm_term=.de9b70817680 (last visited Dec. 20, 2016).

[14] *See* Stephen Herzog, *Revisiting the Estonian Cyber Attacks: Digital Threats and Multinational Responses*, 4 J. OF STRATEGIC SECURITY 49 (2011).

[15] *See* Lily Hay Newman, *The Web-Shaking Mirai Botnet is Splintering–But Also Evolving*, WIRED (Nov. 15, 2016, 7:00AM), https://www.wired.com/2016/11/web-shaking-mirai-botnet-splintering-also-evolving/ (last visited Dec. 20, 2016).

[16] *See* Aaron Boyd, *DNI Clapper: Cyber Bigger Threat Than Terrorism*, Federal Times (Feb. 4, 2016), http://www.federaltimes.com/story/government/it/management/2016/02/04/irs-hardware-failure/79811920/ (last visited Jan. 22, 2017).

[17] *See* W. Earl Boerbert, *A Survey of Challenges in Attribution*, PROCEEDINGS OF A WORKSHOP ON DETERRING CYBERATTACKS: INFORMING STRATEGIES AND DEVELOPING OPTIONS FOR U.S. POLICY, 43 (2010), available at https://www.nap.edu/read/12997/chapter/5 ("The amount of information on the Internet about malicious functionality is so large that a relatively low level of technical competence is required to exploit it.").

The issue of attribution has long been a problem in the realm of cybersecurity. While architectural anonymity has been one of the defining hallmarks and strengths of the internet, it also is the source of this confounding problem. Though most prior scholarship has focused on these technological barriers to attribution, this paper seeks to examine this problem anew by focusing on how the law, not technology, can resolve the problem of attribution. Though attribution has long been thought of as a technical problem, the technical barrier to attribution presents a much narrower problem than the one generally sought by legal attribution. Technological attribution zooms in on the narrower question of whether or not it is possible to guarantee an attribution of an attack to individual(s) purely through technological means.[18] But as legal scholars and practitioners of law know, questions of responsibility are rarely decided solely through a single technological tool or form of evidence, and judgments of responsibility are often not based upon smoking-gun declarations of guilt. Judgments of law are often based on heavy accumulations of evidence, either direct or circumstantial, that in their totality paint a picture of responsibility for malicious behavior.[19] And the very same logic applies even in the context of cybersecurity and attribution. The real question, then, is how systems of law can create a framework by which rules of evidence and procedure can satisfy parties in producing

[18] Although "attribution" as a term can more generally refer to discovering the cause behind an action, I use the term attribution here to refer to the process of identifying the actor behind a cyber-attack. *See* David A. Wheeler & Gregory N. Larsen, Inst. for Def. Analysis, *Techniques for Cyber Attack Attribution*, 1 (2003), http://www.dtic.mil/cgibin/GetTRDoc?AD=ADA468859.

[19] *See Desert Palace, Inc. v. Costa*, 539 U.S. 90, 99–100 (2003) (stating that the Court has "often acknowledged the utility of circumstantial evidence in discrimination cases" and that "[t]he adequacy of circumstantial evidence also extends beyond civil cases; [the Supreme Court] has never questioned the sufficiency of circumstantial evidence in support of a criminal conviction."); *Siegert v. Gilley*, 500 U.S. 226, 236 (1991) (Kennedy, J., concurring) ("I would reject, however, the Court of Appeals' statement that the plaintiff must present direct, as opposed to circumstantial evidence. Circumstantial evidence may be as probative as testimonial evidence."); *Holland v. United States*, 348 U.S. 121, 140 (1954) ("Circumstantial evidence in this respect is intrinsically no different from testimonial evidence.").

legal judgments of responsibility that can legally declare a party to be the cause of a cyber-attack.[20]

While this cybersecurity problem emerges at the intersection of policy and technology, it is also a particularly appropriate problem for the law to resolve. If, fundamentally, law concerns the system to adjudicate disputes, then the question of attributing a cyber-attack raises precisely such a dispute that can be subject to resolution via law. A legal process also bestows the outcome with greater legitimacy and formalizes such resolution with greater institutional weight. And in a more contentious and politicized environment where all reports are held under suspicion of partisan bias, a conclusion derived from legal process is more difficult to dismiss as mere "fake news."[21] And once the culprits of cyber-attackers are known, their tactics and methodologies can be studied, retaliation can be threatened, countermeasures can rectify past incursions, and norms for appropriate behavior can be established and entrenched. But the inability to determine the source of attack frustrates each and every one of these tools. Attribution allows the law to emerge after answering a key requisite question: who is responsible for wrongdoing? A legal framework for attribution is therefore a critical stepping stone to creating a regime to restrict and redress the harms of cyber-attack.

Thus, this paper proceeds to envision a law of attribution in several parts. It begins first in Part I by reviewing the problem of attribution, discussing the threats posed by recent cyber-attacks, the problematic lack of accountability for such attacks, and the general technological

[20] Other scholars have called for the creation of new legal frameworks to address the issues that arise in cyber-attack. Duncan B. Hollis, for example, called for the creation of an "International Law for Information Operations," *Why States Need an International Law for Information Operations Symposium: Crimes, War Crimes, and the War on Terror*, 11 Lewis & Clark L. Rev. 1023 (2007). As Hollis himself states, however, his article "does not aim to offer any specific content for an [International Law for Information Operations], but rather seeks to address the threshold question of why states need an ILIO in the first place." *Id.* at 1029.

[21] *See, e.g.*, Nicholas Loffredo, *'Fake News' Cries Follow Discovery of Russian Malware at Vermont Utility*, NEWSWEEK (Dec. 31, 2016, 5:22 PM), http://www.newsweek.com/fake-news-cries-discovery-russian-malware-vermont-utility-537567 (last visited Jan. 27, 2017).

barriers that scholars and policymakers generally have understood to prevent the attributing a cyber-attack. Part I then rebuts that longstanding position by asserting that the technological question of attribution is much narrower than that required by law, and how attribution may instead reflect a legal question that may more readily be resolved. Part II of this paper then presents a vision for what an international law of attribution might look like. First, it will briefly touch upon the setting and significant details such as the type of forum, adjudication, and key procedural rules such as the burden of proof and standard for assessing state responsibility for the behavior of non-state actors. Part II will suggest procedural and legal rules not only to imagine what a law of attribution would look like, but how such a law will bear an appropriate and reasoned relationship to its substance, the glue that binds the process of law to its legitimacy. Part III of this paper then will then touch upon the most difficult element of a law of attribution, the possible incentives that states would have to join or participate in such a process. While such a question may reflect a much broader general question about the nature of international relations and the issue of state cooperation and compliance, this paper will survey various past examples of international tribunals or modes of international adjudication that might serve as possible models for the law of attribution.

## Part I: The Problem of Attribution

How do you stop an adversary when you don't even know who they are? The inability to identify the source of a cyber-attack allows actors to employ such attacks with impunity, frustrating efforts at creating international laws or treaties to regulate this harmful behavior. Even in cases where formal law is not the answer–where cyber-attacks might best dealt with in ad-hoc state-to-state interactions–states would still need to attribute an attack in order to employ any

informal means of sanctioning the aggressor and their behavior. Thus, the attribution problem is crucial, because attribution is the key prerequisite to any attempt at imposing rules or restrictions on malicious cyber-attacks. As others have noted, "Attribution of a cyber attack to a state is a, if not *the*, key element in building a functioning regime."[22]

The current international regime does little to expressly regulate or control states conduct in the realm of cyber-hacking. No international laws or treaties expressly regulate the use of cyber-attacks.[23] And while scholars point to the potential application of the law of armed conflict, such law has notably not been invoked thus far to respond to cyber-attacks.[24] Given the general uncertainty in international affairs, states may understandably be risk-averse, and hesitate in employing such innovative interpretations of international law when it comes to legal and diplomatic action against other states. The absence of attribution therefore limits institutional and legal solutions, perpetuating cybersecurity's status as an essentially international Wild West, with continued prospects of escalation and uncertainty about the scope and magnitude of future cyber-attacks.[25]

---

[22] Scott J. Shackelford, *From Nuclear War to Net War: Analogizing Cyber Attacks in International Law*, 27 BERKELEY J. OF INT'L L. 191, 232 (2009).

[23] The recently released Tallinn Manual 2.0, for example, surveys the realm of all relevant "specialized regimes of international law and cyberspace," and includes discussion of international human rights law, diplomatic and consular law, law of the sea, air law, space law, and international telecommunications law. None of these categories explicitly set out a regulatory regime for cyber-attacks, cyber-hacking, or cyber espionage. TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS (Michael N. Schmitt ed., 2017) [hereinafter TALLINN MANUAL 2.0]. In fact, the Tallinn manual directly acknowledges that some cyber operations, such as cyber espionage, fall under no per se regulations in international law. *Id.* at 168; *see also* Ashley Deeks, *An International Legal Framework for Surveillance*, 55 VIRGINIA J. OF INT'L L. 291, 300 (2015) ("[M]ost scholars agree that international law either fails to regulate spying or affirmatively permits it.").

[24] *See, e.g.*, Oona Hathaway, *The Law of Cyber-Attack* at 817 (noting that "existing international legal frameworks offer only embryonic or piecemeal protection").

[25] *See, e.g.*, John Ribeiro, *Obama Aims to Avoid a 'Cycle of Escalation' in Cyberattacks by Countries*, PC WORLD (Sep. 6, 2016, 3:08PM), http://www.pcworld.idg.com.au/article/606336/obama-aims-avoid-cycle-escalation-cyberattacks-by-countries/ (last visited Dec. 20, 2016); Alex Kreilein, *Amid Growing U.S. Cybersecurity Threat, A Critical Lack of Trained Experts*, DENVER POST (Sep. 24, 2016, 5:51PM), http://www.denverpost.com/2016/09/24/amid-growing-u-s-cybersecurity-threat-a-critical-lack-of-trained-experts/ (last visited Dec. 20, 2016); Jamie Condliffe, *Security Experts Agree: The NSA Was Hacked*, MIT Tech. Rev. (Aug. 18, 2016), https://www.technologyreview.com/s/602201/security-experts-agree-the-nsa-was-hacked/ (last visited

From the perspective of international relations theory more generally, attribution also provides the linchpin to the creation of international law. It would be easy to see why attribution of cyber-aggressors is needed for liberals to impose institutions of law, since the identification of an aggressor is needed before an institution can impose punishments of any form in order to shape state preferences. But even realists would recognize the necessity of attribution for states to maintain order, even in the absence of an overarching international law. For realists, the traditional mantra is that there is no central authority above states, and that states are always seeking power and to advance their self-interest.[26] While this understanding of international relations would pose an initial hurdle to international cooperation or international law, one counterargument is made through reciprocity. Derived from game theory, advocates of reciprocity point to the fact that rational, self-interested actors who are given a choice between cooperation or defection would optimally choose to cooperate given repeat iterations of the game. This occurs because players punish or reward the others' behaviors in future "games" (or interactions) based off the decisions made in prior iterations. Thus, even assuming the realist framework for state behavior, reciprocity allows international laws to form in the process of cooperation, since international relations often involves repeat interactions between states that form the repeat "iterations" of the international relations game.

Reciprocity, however, assumes that states can accurately punish or reward each others' behavior; although countermeasures may present such a response, the proper use of countermeasures is inextricably tied to proper attribution.[27] Not only is attribution a basic

---

Dec. 20, 2016); Tom Risen, *Iran's Growing Cybersecurity Threat*, U.S. NEWS (Dec. 15, 2014, 11:15AM), http://www.usnews.com/news/articles/2014/12/15/irans-growing-cybersecurity-threat (last visited Dec. 20, 2016).

[26] *See, e.g.*, JOHN J. MEARSHEIMER, THE TRAGEDY OF GREAT POWER POLITICS (2001).

[27] *See, e.g.*, TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE, 29 (Michael N. Schmitt ed., 2013) [hereinafter "TALLINN MANUAL 1.0"] ("A State bears international responsibility for a cyber operation *attributable* to it." (emphasis added)); Lee Ferran, *The NSA is Likely 'Hacking Back' Russia's Cyber Squads*, ABC News (Jul. 30, 2016), http://abcnews.go.com/International/nsa-hacking-back-russias-cyber-

requirement for a state to sanction the responsible malicious actor–proper attribution is also essential to a state claiming legitimacy to its use of sanctions or countermeasures. Law serves not only to determine the outcome of a conflict–the law also serves to legitimize that outcome-determination to others. This is especially true in the realm of international law and international relations, where states lack an overarching authority to compel compliance via force, and instead must cooperate through norms established and legitimized by customary international law.[28] As noted previously, attribution is an essential and necessary condition to further legal action. But in order to take the appropriate legal response (be it countermeasures, diplomatic answers, or responses of any other kind), a state need not only identify the source of an attack – states also need to legitimize their attribution of an attack to other state actors, in order to justify whatever recourse or countermeasure that is to follow. Thus, attribution serves a twofold function in a reciprocity regime by identifying the wrongdoer and legitimizing formal or informal sanctioning behavior to third parties. Consequently, the attribution question is the pivotal first step to any system of law limiting the use of cyber-attacks.

**Why is Attribution So Difficult?**

The difficulty in tracing the source of a cyber-attack has long plagued discussions of cybersecurity, and in much of current scholarship, it has been long accepted wisdom that the technological architecture of the internet makes attribution an exceedingly difficult problem.[29]

squads/story?id=41010651 (last visited Oct. 31, 2016) (mentioning attribution six times in the context of US countermeasures).

[28] Jack L. Goldsmith, & Eric A. Posner, *A Theory of Customary International Law*, 66 U. CHI. L. REV. 1113 (1999).

[29] The *New Yorker* famously published a 1993 drawing of two dogs sitting at a computer, with one telling the other, "On the Internet, nobody knows you're a dog." Glenn Fleishman, *Cartoon Captures Spirit of the Internet*, N.Y. Times (Dec. 14, 2000), http://web.archive.org/web/20141030135629/http://www.nytimes.com/2000/12/14/technology/14DOGG.html. *See also* P.W. SINGER & ALLAN FRIEDMAN, CYBERSECURITY AND CYBERWAR: WHAT EVERYONE NEEDS TO KNOW, 73 (2014) ("Perhaps the most difficult problem is that of attribution."); Aaron P. Brecher, Note, *Cyberattacks and the*

It's a problem that has led scholars and experts to devote countless works to discussing the issue of attribution,[30] and its persistence as a challenge is what led P.W. Singer and Allan Friedman describe attribution as "[p]erhaps the most difficult problem" in the cyber arena.[31]

Attributing cyber-attacks to their source is difficult for a number of reasons. The first reason is primarily due to the structural design of the internet and how information is transmitted across its networks. This entails a brief discussion of the structure of the internet and how it all works.[32] When you wish to do something through the internet–let's say search for a video of Corgi puppies on YouTube–your computer needs to find a way to communicate with the machine hosting YouTube's content, and have that machine send the content of Corgis rollicking around to your machine. How does this happen? First, every machine is assigned an Internet Protocol (IP) number that serves as its "address" This address is usually assigned by an Internet

---

*Covert Action Statute: Toward a Domestic Legal Framework for Offensive Cyberoperations*, 111 Mich. L. Rev. 423, 423 (2012) (saying that cyber attacks "can be nearly impossible to attribute definitively to their sources"); W. Earl Boebert, *A Survey of Challenges in Attribution*, PROCEEDINGS OF A WORKSHOP ON DETERRING CYBERATTACKS: INFORMING STRATEGIES AND DEVELOPING OPTIONS FOR U.S. POLICY, 51–2 (2010); Herbert S. Lin, *Offensive Cyber Operations and the Use of Force*, 4 J. Nat'l Sec. L. & Pol'y 63, 77 (2010) (describing attribution as a problem that "[n]o one has come close to solving"); Stephen Dycus, *Congress's Role in Cyber Warfare*, 4 J. NAT'L SEC. L. & POL'Y 155, 156 (2010) ("[T]he apparent ease with which a cyber attack may be carried out without attribution could make it impossible to fight back at all.").

[30] *See, e.g.*, Thomas Rid & Ben Buchanan, *Attributing Cyber Attacks*, 38 J. OF STRATEGIC STUDIES 4 (2014); Clement Guitton & Elaine Korzak, *The Sophistication Criterion for Attribution*, 158 RUSI J. 62 (2013); Nicholas Tsagourias, *Cyber Attacks, Self-Defence and the Problem of Attribution*, 17 J. OF CONFLICT AND SEC. LAW 229 (2012); Erik M. Mudrinich, *Cyber 3.0: The Department of Defense Strategy for Operating in Cyberspace and the Attribution Problem*, 68 A.F.L. Rev. 167 (2012); David D. Clark & Susan Landau, *Untangling Attribution*, Harvard Nat. Sec. J. (Mar. 16, 2011), http://harvardnsj.org/wp-content/uploads/2011/03/Vol.-2_Clark-Landau_Final-Version.pdf; Jeffrey Hunker, Bob Hutchinson & Jonathan Margulies, *Role and Challenges for Sufficient Cyber-Attack Attribution*, INSTITUTION FOR INFORMATION INFRASTRUCTURE PROTECTION (2008); Susan W. Brenner, *At Light Speed: Attribution and Response to Cybercrime/Terrorism/Warfare*, 97 J. OF CRIM. L. AND CRIMINOLOGY 379 (2007); Lily Hay Newman, *Hacker Lexicon: What is the Attribution Problem?*, WIRED (Dec. 24, 2016, 7:00AM), https://www.wired.com/2016/12/hacker-lexicon-attribution-problem/.

[31] SINGER & FRIEDMAN at 73.

[32] While this may seem rudimentary to those familiar with computer science or the infrastructure of the internet, this paper aims to answer a technological question by proposing a legal solution, meaning much of its audience may be legal or policy professionals with less familiarity with the design and function of the internet. Thus, this paper attempts to present a fairly lay-person friendly description of the internet to communicate the technological issues at play in attribution. Moreover, such explanations are important in dispelling the mysticism surrounding cyber-technology, in order to emphasize the ordinariness of the problems at issue and how legal regimes possess the tools capable of resolving them.

service provider or network, and your computer will generally start out knowing the address of your local internet router, which will then relay your request to the wider internet.

To get your request to the machine with Corgis, your machine will need to know the address of that machine. How does the machine find this out? From the person's perspective, she or he types in "www.youtube.com" in the search bar. On the machine end, these recognizable names are translated to the machine address, or IP number, through the Domain Name System, which can be thought of as a global and decentralized directory that matches website names to IP numbers. So now that your machine figures out the address of the website that holds the bountiful bundles of puppy videos, how does the data from your computer (the request to retrieve content from YouTube) get to YouTube, and how does YouTube send that data back? To simplify a fair amount, the request (the text you enter, or the action of clicking a website link) is translated into data (numbers) at the HyperText Transfer Protocol (HTTP) layer, which then passes the data to the transport and network layers. At the transport layer, the data is broken down into packet-sized chunks of data that each individually contain their destination address, like little envelopes sent through the mail. These packets are transmitted to various servers nodes in the overall network on the way to their destination (think thousands of possible layover destinations on a long trip[33]), until they finally reach the final destination and are reassembled into the original request for data from YouTube. On the machine with YouTube content, the process then repeats itself in the opposite direction as YouTube sends its information back to the user.

[33] Furthermore, the path that a packet of information takes will change every time, given the sheer number of different nodes that can be taken, and the fact that packets and the transportation layer are designed to take the fastest route, which changes at any given time based on the overall traffic that is currently traveling through a system. This represents the fundamentally decentralized nature of the system, and why it is difficult to accomplish attribution by imposing various "checkpoints" in the internet, given the countless other routes that information might otherwise take.

"Aha!" you might interject, "doesn't this mean that the source address is known the whole time? Can't we use that to trace the source of cyber-attacks or attempts to hack someone?" Not quite. The only reason YouTube knows where to send its response is that the original request intentionally includes its address so YouTube can send the data back. Other types of activity– such as uploading a video to YouTube–don't need to embed a return address in the information that you send over. The request for information from YouTube would be like sending a pen-pal a letter, with your address written in that letter in order to receive their response. The upload to YouTube would be like sending someone a gift, which could be done completely anonymously. This current structure of our internet thus does not require the original source of a data transmission for our machines to participate in online activity. The packets of data that we send through the internet only need to know their destination, not their source. Unlike the post office, a return address is not needed, since any data that fails to go through is lost, since you simply can attempt another request again and again until it gets through.

The structural anonymity of the internet is also compounded by a second obstacle to attribution, which is the fact that users can employ a number of techniques and program applications to hide the trail of their online activity. To the extent that any user's IP address is logged in any activity that they perform on the internet, users have the option of using proxy servers[34] or onion-routing tools such as Tor to mask their IP addresses when acting online.[35] Think back to the post office analogy. How might you mask the origin of your envelope? You could hand it to a friend, and ask them to send it out through a different post office than the local one closest to you. You can also "spoof" your original address by writing down a fake return

[34] *See* Larry Greenemeier, *Seeking Address: Why Cyber Attacks are so Difficult to Trace Back to Hackers*, Scientific American (Jun. 11, 2011), https://www.scientificamerican.com/article/tracking-cyber-hackers/.
[35] *See* Joan Feigenbaum, Aaron Johnson & Paul Syverson, *A Model of Onion Routing with Provable Anonymity*, 4886 Financial Cryptography and Data Security, [insert brief explanation of onion-routing]

address. One experiment concludes that nearly one third of internet users could spoof their source IP address without detection.[36]

But can't this structure of the internet be changed, designed such that you *have* to include an origin address for every bit of data transmitted, and include mechanisms authenticating these sources so they couldn't be faked? "Why not redesign the internet so that it does require each packet of data to retain its authentic source address?" the curious congressman might ask. Didn't Larry tell us that the design of the internet is malleable and subject to alteration or modification?[37] The problem is that redesigning the entire internet to trace the original source of every packet of data would be a massive undertaking that would require an overhaul of the entire system, and such a drastic change would also massively reduce the efficiency, speed, and reliability that we have come to rely upon with our modern internet. Think again about the post office comparison.[38] How would you redesign that structure to authenticate the source of every

[36] Robert Beverly, Arthur Berger, Young Hyun & K. Claffy, *Understanding the Efficacy of Deployed Internet Source Address Validation Filtering*, 1 (2009), https://www.akamai.com/cn/zh/multimedia/documents/technical-publication/understanding-the-efficacy-of-deployed-internet-source-address-validation-filtering-technical-publication.pdf.

[37] As Lawrence Lessig points out, the creation of "cookies" in 1994 "introduced a protocol to make it possible for a web server to deposit a small bit of data on your computer when you accessed that server. That small bit of data—the "cookie"—made it possible for the server to recognize you when you traveled to a different page. . . . A small change in the protocol for client-server interaction now makes it possible for websites to monitor and track those who use the site." LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE 2.0, 48 (2006). While Lessig notes that the "traceability of IP addresses and cookies is the default on the Internet now," he also acknowledges that "steps can be taken to avoid this traceability," *Id.* at 49, discussed *supra* n. 24.

[38] A more technical explanation involves a brief explanation of cryptography and its role in authenticating internet connections. How do you authenticate something (such as the authenticity of a packet's origin address) on the internet? Normally, you authenticate yourself by virtue of what you are (visually being recognizable as a particular person), producing something only you have (e.g. your driver's license) or producing something only you know (such as a password). To replicate this on the internet, we use cryptography, which is a way of encoding data such that it is only intelligible to a recipient who has the proper "key" to unlock its information. Because a digital master "key" would be easy to compromise (the moment it gets copied, it can instantly be reproduced infinitely and disseminated widely), modern encryption uses a [technical term for public-private key system] which means that each encryption is its own unique "lock" that has a correspondingly unique "key." Particular details aside, the main takeaway is this – enacting this form of encryption takes time and resources. While certain secure servers may perform an initial encryption to make sure

package that travels through the system? To do this, you would have to place heightened security at every point of entry, requiring greater resources and slowing down the entry of information.

Since the original source address be spoofed at the first point of entry–it could be spoofed, modified, and faked at any "layover" or network server along the way–any structural redesign of the internet would require the same onerous verification to take place at every node in the transportation route. This means that any increased source-verification, and its subsequent slowing of the data transmission process, will have an exponential effect on the overall speed of the internet. Remember that information is not sent in a single package across the internet–to increase the speed and efficiency of the process, data is broken up into packets that are later reassembled at their destination. To do something over the internet, then, you don't need just one packet of data to get through. You need most, if not all of them, to make it to the end. Say that a video contains one thousand packets of data. If a source-verification process made packets take twice as long to pass through, then that effect would be multiplied a thousand times. And this assumes a straight-shot transmission of data from one point to another, when the data in reality will have to cross through multiple servers on the way to its final destination. Since new information can enter at any intervening server, and since the source of any information can be spoofed at any intervening server on the way to a destination, the whole process will have to repeat itself, for every packet, at every server up to the destination. The exponential increase in cost and time would render the system unworkable.[39]

Third, even if the internet could arduously be redesigned to embed the source internet protocol or IP address of every bit of data sent over the internet, these addresses would

[39] Nor would trashing the packet system be feasible. The whole reason the packet system was adopted was that it presented a far more efficient and robust information transmission system that was less vulnerable to disruptions. The only existing alternative is the circuit system–which is the technology we associate with phone lines used in the 20th century. *See* LAWRENCE LESSIG, THE FUTURE OF IDEAS, 31 (2001).

accomplish the goal of merely identifying the source machine of an attack, and not a person, creating another degree of attenuation between attack and the attacker. There are innumerable situations where attackers may steal or compromise another person's device,[40] or exploit public devices or networks used by multiple persons (such as a library computer, or in the wireless network of a coffee shop). The Mirai Botnet attack, for example, was an instance in which malicious agents exploited thousands of other devices which were co-opted into becoming the instruments of the attack.[41]

Fourth, even if all the technological problems are overcome and a particular person is identified as having launched a cyber-attack, there remains the question of whether or not a state can be held responsible for that individual's actions. In other words, cyber-attacks also raise the question of when states can be held responsible for the wrongdoing of non-state actors. While this legal conundrum is most frequently discussed in the context of terrorists or corporations,[42] the issue is just as salient when it comes to hackers and cyber-attackers, who generally lack a uniform or flag to identify them as acting in the name of any particular state. Note that this is not a technological barrier to attribution, but a legal one.[43] This particular concern foreshadows the need to create a legal solution to the problems posed by attribution.

[40] This technique is used to create "zombie" computers or "botnets" that are then used to launch attacks, often from an army of such devices. *See* Greenemeier, *supra* note 27.
[41] *See* Robinson Meyer, *How a Bunch of Hacked DVR Machines Took Down Twitter and Reddit*, Atlantic (Oct. 21, 2016), http://www.theatlantic.com/technology/archive/2016/10/how-a-bunch-of-hacked-dvr-machines-took-down-twitter-and-reddit/505073/ (last visited Jan. 27, 2017).
[42] *See, e.g.*, Andrew Clapham, *Human Rights Obligations for Non-State-Actors: Where are We Now?*, *in* DOING PEACE THE RIGHTS WAY: ESSAYS IN INTERNATIONAL LAW AND RELATIONS IN HONOR OF LOUISE ARBOUR (Fannie Lafontaine & François Larocque eds., 2015); Oona A. Hathaway, Emily Chertoff, Lara Dominguez, Zachary Manfredi & Peter Tzeng, *Ensuring Responsibility: Common Article 1 and State Responsibility for Non-State Actors*, 95 TEXAS L. REV. (2017) (forthcoming); Derek Jinks, *State Responsibility for the Acts of Private Armed Groups*, 4 CHI. J. OF INT'L L. 83 (2003).
[43] *See* Shackelford, *supra*, note 19 at 233.

These are the numerous hurdles, technological and legal, that have often been cited as the barrier to the creation of a legal regime for regulating cyber-attacks.[44] While previous scholarship has often viewed the technological problem of attribution as an intractable difficulty best left to the engineers, recent scholarship has begun to recognize that the attribution problem may not be the impossible task it has been portrayed to be.[45] Yet, while these scholars have pointed of the possibility of a political solution to the attribution puzzle,[46] these pieces fall shy of proposing an actual legal or political framework to resolve the attribution problem once and for all.

**The Attribution Problem is a Red Herring**

Despite the numerous technological barriers to attribution, the technological problem is a red herring. These technical obstacles only prevent us from the very narrow conclusion of when we might be absolutely certain that an agent was responsible for a cyber-attack. The law, however, almost never operates on the impossibly high standard of absolute certainty. Even United States criminal law, with its famously high burden of proof in favor of the defendant, demands only that there be no *reasonable* doubt before a conviction, as opposed to demanding that there be no doubt at all.[47] Upon reexamination, the attribution question is, at its core, a question of responsibility. And responsibility is a fundamentally legal question, one that the law has frequently answered, even in cases without absolute causal certainty. Thus, this paper answers the attribution problem with two main points:

---

[44] *See, e.g.*, Matthew C. Waxman, *Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4)*, 36 YALE J. OF INT'L L. 421, 443-44 (2011).
[45] *See* Rid & Buchanan, *Attributing Cyber Attacks*, 6 (explaining how actual attribution is more common and nuanced phenomenon than previously thought, and that the attribution issue is more of a political, rather than purely technological question).
[46] *Id.*
[47] *See* James Q. Whitman, *The Origins of "Reasonable Doubt,"* Faculty Scholarship Series, 8 (2005).

*First*: despite the barriers to attribution, computer scientists have developed a whole range of tools to trace cyber-attacks, and empirically, large-scale attacks tend to leave behind enough footprints (or circumstantial evidence) to lead us to their source.

*Second*: the law does not demand guaranteed certainty, but a sufficient degree of certainty that someone is responsible; the question of what counts as a sufficient degree of certainty is a purely legal question that we can answer.

Once these two points are established, the question is no longer *whether* cyber-attacks can be attributed, but *how* we might configure our system of law to do so, developing rules of evidence, procedure, burdens of proof, etc.

On the first point, the emphasis on the technological nature of attribution has naturally attracted much interest from those with a technical perspective, and computer scientists have responded in turn by developing a whole suite of tools to attribute cyber-attacks or intrusions.[48] While none of these methods may individually present silver bullet solutions, they each offer forensic techniques that can shed some light on any particular case, and that cumulatively present the very real possibility of a confident degree of attribution. Returning once more to the post office analogy, in the same way that anonymous envelopes could be traced through forensic evidence (searching for fingerprints, identifying handwriting, etc.), there are ways to use circumstantial evidence to attribute the transmission of digital information and subsequent cyber-attacks. This is especially true of the cyber-attacks that this paper is concerned with, namely major cyber-attacks that are likely to trigger or demand state responses. By virtue of their larger

[48] *See, e.g.*, David A. Wheeler & Gregory N. Larsen, *Techniques for Cyber Attack Attribution*, Institute for Defense Analyses (2003) (listing techniques such as store logs and traceback inquiries, input debugging, modifying transmitted messages, transmitting separate messages, reconfiguring and observing networks, querying hosts, inserting host monitoring functions, stream matching, honey pots, forward-deployed Intrusion Detection Systems, and network ingress filtering); *see also* Haining Wang, Cheng Jin & Kang G. Shin, *Defense Against Spoofed IP Traffic Using Hop-Count Filtering*, 15 IEEE/ACM Transactions on Networking 40 (2007) (describing a technical method of addressing the "spoofing" technique described *supra* on p. 12).

scope or scale, such attacks tend to be more likely to leave tracks behind. Bigger operations also require greater resources, limiting the field of potential adversaries capable of launching such cyber-attacks.

In fact, this was precisely what happened with three recent major cyber-attacks: the Stuxnet attack, the Sony attack, and the recent DNC hack. This paper reviews each attack in turn to describe how accumulations of forensic and circumstantial evidence led to the attribution of these attacks, thus demonstrating that the technological problem of attribution is overblown.

**Stuxnet**

Starting in 2009, Iran's uranium centrifuges began failing, and nobody understood why.[49] Nearly 1,000 of Iran's 6,000 centrifuges were destroyed over the course of a year.[50] In the summer of 2010, a computer security firm in Belarus was hired to troubleshoot Iranian computers that mysteriously kept crashing–and in this investigation, the firm stumbled upon a series of files that would later become known as the Stuxnet virus.[51] The Stuxnet virus was recognized as the "world's first digital weapon."[52] It was a complex malware designed to infiltrate secure Iranian nuclear facilities, infect the industrial controllers that operated the nuclear centrifuges, and destroy such centrifuges by manipulating the pressure levels and rotor speeds inside those centrifuges.[53] The virus was intentionally designed to cause such havoc

[49] Kim Zetter, *How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History*, WIRED (Jul. 11, 2011, 7:00AM), https://www.wired.com/2011/07/how-digital-detectives-deciphered-stuxnet/.
[50] Ellen Nakashima & Joby Warrick, *Stuxnet was the Work of U.S. and Israeli Experts, Officials Say*, WASH. POST (Jun. 2, 2012), https://www.washingtonpost.com/world/national-security/stuxnet-was-work-of-us-and-israeli-experts-officials-say/2012/06/01/gJQAlnEy6U_story.html.
[51] Zetter, *How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History*.
[52] KIM ZETTER, COUNTDOWN TO ZERO DAY: STUXNET AND THE LAUNCH OF THE WORLD'S FIRST DIGITAL WEAPON (2015).
[53] Ralph Langner, *To Kill a Centrifuge*, The Langner Group, 4-12 (Nov. 2013), http://www.langner.com/en/wp-content/uploads/2013/11/To-kill-a-centrifuge.pdf#http://www.langner.com/en/wp-content/uploads/2013/11/To-kill-a-centrifuge.pdf.

slowly and gradually, rendering detection less likely; it even included a function that manipulated Iranian sensors to pretend that the manipulated functions were working as normal.[54]

Despite its significant attempt to cover its track, experts concluded that Stuxnet was a joint United States and Israeli production.[55] Contextual cues, such as the target state and the targeted data or device, often narrows down the list of possible adversaries. In Stuxnet's case, that information alone was nearly dispositive, since so few would have the motivation and the means of targeting Iran's nuclear centrifuges.[56] Furthermore, the scale of an attack often reveals information about an attacker. Although advanced persistent threats are some of the scarier forms of cyber-attack, their strength also becomes their weakness, since only a few states would have the intelligence and resources to develop such a threat. This was another giveaway from the Stuxnet attack–the fact that the code had four zero-day exploits[57] (which would have been worth millions to private hackers in terms of its resale value)[58] again implied that there was serious firepower behind the attack, almost guaranteeing that such an attack came from a state. Finally, little telltale clues can often identify the source of an attack. Through Stuxnet's code, investigators were able to discover the main target of its attack based off names and ID numbers that referenced Siemens devices–the industrial centrifuge controllers that were the target of manipulation. Given the narrowness of the target, and the immense resources that went into it, it was easy to deduce the states behind the attack.

---

[54] *Id.* at 9, 15.
[55] *See, e.g.*, Nakashima & Warrick, *Stuxnet was the Work of U.S. and Israeli Experts, Officials Say*; Nate Anderson, *Confirmed: US and Israel Created Stuxnet, Lost Control of It*, ArsTechnica (Jun. 1, 2012, 6:00AM);
[56] Although the informants in the *Zero Day* documentary did claim that Stuxnet should have gone undiscovered but-for the ramped up aggression of Israeli
[57] A zero day exploit is "a cyber attack exploiting a vulnerability that has not been disclosed publicly." Leyla Bilge & Tudor Dumitras, *Before We Knew It: An Empirical Study of Zero-Day Attacks in the Real World*, CCS '12 Proceedings of the 2012 ACM Conference on Computer and Communications Security (Oct. 2012).
[58] ZERO DAYS (Alex Gibney 2016).

**Sony Attack**

In October 2014, hackers raided the computer network of Sony Pictures.[59] The hackers downloaded nearly the entirety of Sony Pictures' records, including internal communications, scripts, and even unreleased movies, and the hackers proceeded to dump these all online while erasing them from Sony's computers.[60] This attack affected over 3,000 computers and 800 servers,[61] and it was famously known for canceling the theatrical release of "The Interview," the comedy film where Seth Rogen and James Franco assassinate Kim Jong Un.[62]

Only twenty-five days after the attack, the FBI attributed it to North Korea. FBI director Comey announced that he had "very high confidence" that the attack came from North Korea,[63] and NSA Director Michael Rogers similarly said that he was "confident" that "this was North Korea."[64] But how exactly did they reach this conclusion, and reach it with such confidence? Again, the attribution of the attack was made easier through context. Although this attack targeted a private actor, instead of public one (as in the Stuxnet attack), Sony officials were well aware that the Interview could antagonize North Korea, whose regime "had been widely blamed for a series of cyber attacks" in the past.[65] These reports were confirmed by two consultants, who each warned Sony executives that North Korea would likely employ its hackers to wreak havoc.[66] The North Korean Ministry of Foreign Affairs even published a statement, prior to The

---

[59] Andrea Peterson, *The Sony Pictures Hack, Explained*, Wash. Post (Dec. 18, 2014), https://www.washingtonpost.com/news/the-switch/wp/2014/12/18/the-sony-pictures-hack-explained/?utm_term=.ddf01aeda7ae.
[60] Peter Elkind, *Inside the Hack of the Century: Part I*, Fortune (Jun. 25, 2015, 6:00 AM), http://fortune.com/sony-hack-part-1/.
[61] *See* Steve Kroft, *The Attack on Sony*, CBS News (Apr. 12, 2015), http://www.cbsnews.com/news/north-korean-cyberattack-on-sony-60-minutes/.
[62] Peterson, *The Sony Pictures Hack, Explained*.
[63] Peter Elkind, *Inside the Hack of the Century: Part III*, Fortune (Jun. 27, 2015, 8:00 AM), http://fortune.com/sony-hack-final-part/.
[64] *See* Sam Frizell, *NSA Director on Sony Hack: 'The Entire World is Watching'*, Time (Jan. 8, 2015), http://time.com/3660757/nsa-michael-rogers-sony-hack/.
[65] Elkind, *Inside the Hack of the Century: Part I*.
[66] *Id.*

Interview's release, declaring that North Korea would take a "decisive and merciless countermeasure" if Sony released the movie.[67]

So there was means and motive.[68] There was also forensic evidence. FBI officials noted similarities to the DarkSeoul attack, a previous cyber-attack that North Korea launched against South Korean banks.[69] They also discovered evidence that the malware was produced on computers with Korean language settings.[70] Moreover, the data revealed a trail of internet staging points for the attack that similarly pointed towards North Korea.[71] Finally, the FBI cited intelligence from "sensitive sources and methods"[72]–in other words, the United States had evidence collected from spying on North Korea.[73]

**DNC Hack**

The DNC Hack, still fresh in recent memory, offers the latest example of a major attack that has been attributed to a state actor. As in the Sony attack, the U.S. intelligence community has concluded with "high confidence" that the DNC hack came from Russia.[74] Although this determination also relied on classified intelligence information, several private cybersecurity

[67] *See* Michael Cieply & Brooks Barnes, *Sony Cyberattack, First a Nuisance, Swiftly Grew Into a Firestorm*, N.Y. Times (Dec. 30, 2014), https://www.nytimes.com/2014/12/31/business/media/sony-attack-first-a-nuisance-swiftly-grew-into-a-firestorm-.html.
[68] Means, motive, and opportunity are a common way of describing some of the elements of criminal law. See, for example, motive described in relation to intent by Walter Wheeler Cook, *Act, Intention, and Motive in the Criminal Law*, 26 YALE L.J. 645 (1917). For a translation of the terms means, motive, and opportunity to the context of cyber attacks, see Elizabeth Van Ruitenbeek, Ken Keefe, William H. Sanders & Carol Muehrcke, *Characterizing the Behavior of Cyber Adversaries*: *The Means, Motive, and Opportunity of Cyberattacks*, 2010 International Conference on Dependable Systems and Networks Supplemental, https://www.perform.illinois.edu/Papers/USAN_papers/10VAN01.pdf.
[69] Elkind, *Inside the Hack of the Century: Part III*.
[70] *Id*.
[71] *Id*.
[72] *Id*.
[73] David E. Sanger & Martin Fackler, *N.S.A. Breached North Korea Networks Before Sony Attack, Officials Say*, N.Y. Times (Jan. 18, 2015), https://www.nytimes.com/2015/01/19/world/asia/nsa-tapped-into-north-korean-networks-before-sony-attack-officials-say.html.
[74] Massimo Calabresi & Pratheen Rebala, *Here's the Evidence Russia Hacked the Democratic National Committee*, Time (Dec. 13, 2016), http://time.com/4600177/election-hack-russia-hillary-clinton-donald-trump/.

firms were consulted in the investigation, and offer public evidence tracing the attack to Russia.[75] They noted, for example, that the DNC hackers used exfiltration tools and coding identical to ones used by a group of Russian hackers known to work for the Russian FSB (Russia's successor to the KGB).[76] These analysts also linked the DNC hack to the same IP address used to conduct an attack against the German Parliament in 2015.[77] Security experts noted a signature in Russia's Cyrillic alphabet left behind as a digital signature.[78] And even more subtly, security analysts noted that the DNC hackers stopped operations on Russian holidays, and that their work hours aligned with a Russian time zone.[79]

Of course, such circumstantial evidence isn't completely conclusive,[80] and it is possible that some of the information could be planted. But systems of law have long been able to allocate punishment and responsibility, even when responsibility is derived solely from circumstantial evidence.[81] In the case of the DNC hack, while it is possible that someone planted clues like the Cyrillic signature as a red herring, it is far less likely that the hacker groups coordinated their operations entirely within Russian time zones and holidays as part of their ploy, since such efforts would have high coordination costs, and would require an unusual degree of

---

[75] *Id.*
[76] *Id.*
[77] *Id.*
[78] *See* Josh Meyer, *Why Experts are Sure Russia Hacked the DNC Emails*, NBC News (Jul. 26, 2016), http://www.nbcnews.com/news/us-news/why-experts-think-russia-hacked-dnc-emails-n616486 (last visited Oct. 31, 2016).
[79] *Id.*
[80] One author, for example, acknowledges that the evidence that Russia was involved in the hack was good, but comments that "'good' doesn't necessarily mean good enough to indict Russia's head of state for sabotaging our democracy. *See* Sam Biddle, *Here's the Public Evidence Russia Hacked the DNC – It's Not Enough*, Intercept (Dec. 14, 2016, 11:30 AM), https://theintercept.com/2016/12/14/heres-the-public-evidence-russia-hacked-the-dnc-its-not-enough/.The question of when such evidence "good enough" to indict a state is precisely the kind of legal dispute that a law of attribution is needed to resolve.
[81] *See, e.g.*, *People v Benzinger*, 36 N.Y.2d 29, 31-32 (1974); *People v Cleague*, 22 N.Y.2d 363, 367 (1968); M. Alex Johnson, *'Circumstantial' – the Scarlet C?*, NBC News, http://www.nbcnews.com/id/3340617/ns/us_news-crime_and_courts/t/circumstantial-scarlet-c/#.UkHZcSqF9rc (last visited Oct. 31, 2016).

sophistication. Ultimately, just as in criminal cases, sufficient evidence can accumulate to identify the source of an attack.

The problem, then, is not in identifying the source of an attack–the problem is in convincing other states that a source has correctly been identified. A state that wishes to employ countermeasures needs to convince other states of the accuracy of its attribution in order to establish the legitimacy of its attack.[82] This issue may arise for two main reasons: 1) attribution may be based on data collected through state espionage or intelligence gathering efforts that states may wish to keep secret[83] and 2) when states have plausible factual bases for attributing an attack, they may not want to disclose such evidence, since cyber-attackers could learn from those mistakes and avoid leaving the same fingerprints in the future.

While these efforts were ultimately based on an accumulation of circumstantial evidence, circumstantial evidence provides a sufficient degree of confidence to support legal judgments in many areas of law. After all, the question of attribution is largely identifying the actor *responsible* for an attack, and responsibility (and what defines responsibility) is a question that is well within the domain of law. It is also one that the law has addressed on a number of occasions, even in contexts that attenuate or obfuscate the link between actor and harm. In torts, for instance, the doctrines of strict liability and *res ipsa loquitur* demonstrate that the dispositive question may not always be "who did it" (a question that is often already answered through context) but "how do we hold this person or entity accountable." And the use of different liability standards in different contexts reflects the law's flexibility in creating appropriate

[82] While countermeasures themselves might be covert, the presumption is that even a covert act ought to be legally justifiable, since the attribution of a countermeasure is always a significant risk, given the discussion of attribution earlier.

[83] *See, e.g.*, David E. Sanger & Martin Fackler, *N.S.A. Breached North Korean Networks Before Sony Attack, Officials Say*, N.Y. Times (Jan. 18, 2015), http://www.nytimes.com/2015/01/19/world/asia/nsa-tapped-into-north-korean-networks-before-sony-attack-officials-say.html (last visited Oct. 31, 2016).

frameworks to resolve such conflicts. When designing our law of attribution, then, these concerns will involve some inquiry into the general standards of proof and causation invoked in other areas of law where courts have employed legal tools to establish a sufficient degree of confidence to assign responsibility to an actor.

## Part II: The Law of Attribution

How do you begin to imagine a system of rules and procedures–a system of law–from the ground up? Fortunately, prior systems of law and procedure provide abundant material to draw upon, presenting numerous institutional features and designs to consider in outlining such a structure. A system of law, of course, is built with certain ends in mind. So this exercise begins with the general goals driving a law of attribution to provide a starting point, and from there, it will proceed logically across the questions that such a system will need to answer in order to produce a satisfactory ruling of law.

An international law of attribution must address several questions when designing its structure and parts. First, this section will address the question of whether a trans-substantive set of rules for attribution is possible, and the related question of what ends this law of attribution will be used for. These answers lay the foundation for the general outline the system's overall structure and framework, which will address design choices such as an adversarial model versus an inquisitorial system, and other key aspects of institutional setting. Then, this section will discuss the key procedural rules that would define the boundaries of substantive law. This includes rules such as the burden of proof, the standard for assessing state responsibility for the behavior of non-state actors, and rules for evidence and managing sensitive intelligence that might be produced to support a claim of attribution. While procedural in nature, such rules have

tremendous influence over the potential outcome of cases, and an appropriate process must be developed to ensure that the process of law bears an appropriate and reasoned relationship to the substance of the law, the glue that binds the process of law to its legitimacy.

**A Trans-Substantive Law of Attribution**

First: is it possible to develop a trans-substantive law of attribution whose rules will apply, regardless of the legal or political action justified by the attribution? Put another way, are the procedural rules and requirements for attribution contingent upon the subsequent legal sanction that might be imposed on those who are attributed with causing a cyber-attack? We can easily imagine, for instance, that laws for attribution could change their standards of strictness or flexibility based on the seriousness of the sanction imposed upon the state to whom an attack is attributed. To answer the question of trans-substantivity, we might first conceive of the various possible legal sanctions, and consider whether or not those conditions alone are sufficient to change what we think the procedural rules or process for attribution should be.

Speaking broadly, there may be several purposes behind a law of attribution–that is, several types of subsequent sanctions or responses that might be justified by a legal claim of attribution. First, after attributing an attack, negative economic punishment might be placed upon the state responsible for the cyber-attack, such as that of an economic sanction. Second, a state attributed with launching an attack might be denied positive benefits, through denying them participation in future international treaties or agreements. Third, attribution might justify a hack-back countermeasure. Fourth, attribution might justify a military response. These possible responses to attribution might further be divided along two categories: unilateral action or multilateral action.

| | **Unilateral** | **Multilateral** |
|---|---|---|
| **Denial of Diplomatic Access/Agreements** | Refusal to engage in trade agreement, treaty, or other bi-party agreement that targeted state might seek, denial of diplomatic access | Denial of membership in broader trade agreement or treaty |
| **Negative Economic Punishment** | Economic sanctions, rescinding current trade agreements | Multilateral sanction regime |
| **Cyber Countermeasure** | "Hack-Back" | Jointly produced cyber strikes, e.g. Stuxnet |
| **Military Countermeasure** | Military invasion, targeted military strike, remote bombardment, drone strike, special ops etc. | Coalition-based Military Force |

While these options do present a host of practical and policy responses that states might pursue after an attributed cyber-attack, for the purposes of creating rules of attribution, these responses can be considered along two main axes of salience when it comes to their influence on how we design our rules of attribution: 1) whether the action is unilateral or multilateral, and 2) how "serious" the punishment is.

The first question–whether attribution is used to launch a unilateral or multilateral response–actually has a fairly narrow effect on the overall theory for a law of attribution. This is largely because the purpose behind a law of attribution is generally consistent across both unilateral and multilateral responses–attribution justifies a punishment in the eyes of the international community. Whether or not a state wishes to punish a cyber-aggressor with its own unilateral action, or the action of a multilateral coalition, the goal of attribution is to legitimize that behavior in the eyes of third-parties in the international community.

The one exception is in cases where multilateral commitment is not guaranteed, and an aggrieved state needs to convince others not only that retribution is justified, but also that other states ought to participate in the retribution. These cases may tilt the theory of a law of attribution to more stringent requirements, since states might demand higher confidence of attribution before committing their own resources to responding to a cyber-attack that did not afflict them directly. As a result, there may be a confidence gap between the aggrieved state and states that might participate in the multilateral response.

There are two responses to the confidence gap concern: 1) states who suffer the attack directly have an extremely high interest in correctly identifying the source of the attack (to maintain credibility, to ensure signal deterrence capabilities for future attacks, etc.), meaning that the confidence gap may depend less on the certainty of attribution and more on the general incentives that states have for joining or not joining multilateral action,[84] and 2) the mere existence of a multilateral institution that commits non-victim states to responds seems to suggest that the source of that institutional connection may itself suffice to cause those states to join in imposing punishment without the extra assurance of a stricter attribution regime. For example, if states were bound to multilateral responses to a cyber-attack, then the fact of their being bound–as a matter of law, or as a matter of rational interest in securing future cooperation–might be enough to justify a state's decision to join the aggrieved state in issuing a multilateral response to an attributed source of cyber-attack. Consider, for example, the techniques that the

[84] This assumes, however, that states behave rationally. If states are risk-averse, and transactional and information costs makes states generally less inclined to punish cyber-aggressors compared to states that directly suffer an attack, then the law of attribution might account for this by adjusting rules of procedure to allow coalition parties (states that are bound to a multilateral response to cyber-aggression) to join a proceeding, which in turn may allow such states to receive access to evidence that might otherwise be under seal to other third-parties. *See* discussion *infra* 51-56.

United States employed to gather a coalition of states to participate in the Iraq War in 2003.[85] As a result, the unilateral or multilateral distinction likely will not alter the possibility of a trans-substantive set of rules for attribution.

The severity of possible countermeasures, however, may more seriously threaten the idea of a single trans-substantive law of attribution, largely on the intuitive principle that more serious countermeasures may demand more stringent procedural rules, causing such rules to depend upon the countermeasure that a state shall pursue.[86] While this may seem true in the abstract, it is worth exploring in the specific context of cyber-security and the possible state responses detailed above. Organizing the possible countermeasures by the seriousness of their magnitude, they can roughly be ordered as follows (from highest magnitude to lowest): military force, cyber countermeasures (or "hack-back" protocols), economic sanctions, and diplomatic punishments.

While military force covers a wide range of possible actions (from a full-scale military campaign to limited strikes and special operations), these actions nonetheless can be categorized as the most severe possible countermeasure in response to a cyber-attack. Given the general costs of military action and the danger of escalation,[87] military force is an increasingly rare option pursued by states.[88] Moreover, international law expressly places a general prohibition on the use

[85] *See, e.g.*, ANDREW JOSEPH LOOMIS, LEVERAGING LEGITIMACY IN SECURING U.S. LEADERSHIP: NORMATIVE DIMENSIONS OF HEGEMONIC AUTHORITY, 202 (2008); Barbara Slavin, *U.S. Builds War Coalition with Favors – and Money*, USA TODAY (Feb. 25, 2003).
[86] *See, e.g.*, *Mathews v. Eldridge*, 424 U.S. 319 (1976) (considering the private interest as one of the three key prongs in assessing the appropriate level of procedural due process); *Bridges v. Wixon*, 326 U.S. 135, 154 (1945) ("Though deportation is not technically a criminal proceeding, it visits a great hardship on the individual and deprives him of the right to stay and live and work in this land of freedom. That deportation is a penalty—at times a most serious one—cannot be doubted. Meticulous care must be exercised lest the procedure by which he is deprived of that liberty not meet the essential standards of fairness.")
[87] *See* Joseph S. Nye Jr., *Soft Power*, 80 FOREIGN POLICY 153, 157-58 (1990).
[88] *See, e.g.*, Therése Pettersson, Peter Wallensteen, *Armed Conflicts, 1946-2014*, 52 J. OF PEACE RESEARCH, 536 (2016); Joshua S. Goldstein & Steven Pinker, *The Decline of War and Violence*, BOSTON GLOBE (Apr. 15, 2016), https://www.bostonglobe.com/opinion/2016/04/15/the-decline-war-and-violence/lxhtEplvppt0Bz9kPphzkL/story.html.

of force.[89] Still, both politicians and military leaders have postured towards the possibility of military responses to being hacked,[90] leaving the option on the table when it comes to possible countermeasures against cyber-attacks, especially if the cyber-attack is serious enough to rise to the level of being classified as an act of force.[91] The specter of military action would no doubt trigger tremendous scrutiny and an exceedingly high bar of confidence to properly attribute the source of a cyber-attack. This is especially true given the infamy attached to the invasion of Iraq in 2003, which the United States initiated on the false assertion that Iraq possessed Weapons of Mass Destruction.[92]

Another category of countermeasure, the cyber "hack-back,"[93] might also rise to the level of seriousness linked to the use of military force. While cyber "hack-backs" may cover a potentially even broader array of activities than that of military force, several scholars have pointed out the possibility that cyber-attacks have for causing as much damage as a traditional, kinetic military attack, sometimes qualifying as force that falls under the international law of war.[94] To the extent that cyber hack-backs are considered the international equivalent of military

[89] *See* U.N. Charter art. 2, para. 4.

[90] *See* DEP'T OF DEFENSE, CYBERSPACE POLICY REPORT: A REPORT TO CONGRESS PURSUANT TO THE NATIONAL DEFENSE AUTHORIZATION ACT FOR FISCAL YEAR 2011, SECTION 934, at 2 (2011); Katie Bo Williams, *Clinton: Treat Cyber Attacks 'Like Any Other Attack'*, HILL (Aug. 31, 2016, 1:47 PM), http://thehill.com/policy/cybersecurity/293970-clinton-treat-cyberattacks-like-any-other-attack; Patrick Howell O'Neill, *U.S. Military and NATO Agree: Cyberattacks Could Trigger Real War*, Daily Dot (Jun 22, 2016, 10:22 AM), https://www.dailydot.com/layer8/dod-nato-cyber-attack-response/.

[91] *See* Hathaway, *supra* note 21; Daniel B. Silver, *Computer Network Attack as a Use of Force Under Article 2(4) of the United Nations Charter*, 76 COMPUTER NETWORK ATTACK AND INTERNATIONAL LAW 73, 80–82 (2002); 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, 909 (1999).

[92] *See* Martin Chulov & Helen Pidd, *Defector Admits to WMD Lies that Triggered Iraq War*, Guardian (Feb. 15, 2011, 7:58 P.M.), https://www.theguardian.com/world/2011/feb/15/defector-admits-wmd-lies-iraq-war.

[93] *See* Corey T. Holzer & James E. Lerums, *The Ethics of Hacking Back*, IEEE (2016); Vikas Jayaswal, William Yurcik & David Doss, *Internet Hack Back: Counter Attacks as Self-Defense or Vigilantism?*, IEEE 380 (2002).

[94] *See* Hathaway, *supra* note 821; David E. Graham, *Cyber Threats and the Law of War*, 4 J. OF NAT'L SEC. L. & POL'Y 87 (2010).

force, then, such countermeasures might also demand a particular set of procedural rules to justify an attribution in those contexts, given their high stakes.

Does this doom the project of creating a trans-substantive law of attribution? Not at all– laws can account for punishments of differing degrees of severity by simply modifying relevant procedural rules or requirements to trigger particular punishments. Consider, for example, U.S. Copyright law, which contains provisions that can impose civil damages, enhanced civil damages, or criminal liability.[95] All three punishments for infringement attach to the same general system of copyright law; the particular punishment turns on the juncture of the defendant's mens rea. "Willful" infringement can earn enhanced statutory damages, while "purposeful" infringement can earn criminal penalties.[96] Thus, higher levels of penalty can still attach to the same framework of law, even if the higher penalty deserves consideration of some higher standard of proof. The relevant question, then, is whether or not that difference in penalty can have its corresponding effect on procedural rules confined to a single category of rule (e.g. adjusting copyright infringement standards merely by solely notching up the mens rea standard, thereby preserving a single overarching framework of copyright law).

Here, in the context of attribution, the same adjustment of law can account for differences in punishment subsequent to the attribution of an attack to a particular state. It is true that state-to-state adjudication may care less about the particular mens rea involved, since mens rea focuses on individual mindsets, and states are composed of a multitude of individuals, making a state's "mens rea" a legal fiction. Nonetheless, a law of attribution can adjust its standards of scrutiny based on the burden of proof it requires. The standards for burden of proof, like mens rea, are a core element of procedure that can be notched higher or lower based on the severity of

[95] *See* 17 U.S.C. § 504(b); 17 U.S.C. § 504(c); 17 U.S.C. § 506(a).
[96] *Id.*

the chosen remedy. If anything, the mens rea requirement is merely one particular means of fine-tuning the burden of proof, and the evidentiary standard of proof presents another holistic way to incorporate the seriousness of a penalty into the generalized requirements of a procedural framework.

Given the possibility of creating a trans-substantive law of attribution, then, the next step is to begin outlining the main features and characteristics of such a system, beginning with the foundational elements that will shape the structure of the overall law.

**Adversarial or Civil System**

One main design choice to make is whether a law of attribution would operate under an adversarial framework, as typified by the American and British legal systems, or under an inquisitorial framework,[97] as typified by most of the European, Asian, and South American countries' legal institutions.[98] The choice between an adversarial or inquisitorial framework is largely reflective of a philosophy of legal process that then shapes the rules and overall design of the system. An adversarial legal framework is one primarily characterized as a system where impartial decision makers (judge or jury) issue judgments on disputes based on evidence and arguments presented by the parties (and their legal representatives).[99] This system is one where

[97] *See, e.g.*, Franklin Strier, *What Can the American Adversary System Learn from an Inquisitorial System of Justice?*, 76 JUDICATURE 109 (1992-1993).
[98] *See Alphabetical Index of the 192 United Nations Member States and Corresponding Legal Systems*, JURIGLOBE, http://www.juriglobe.ca/eng/syst-onu/index-alpha.php (last visited Apr. 1, 2017). The inquisitorial system is also sometimes referred to as the "continental system." *See generally*, John H. Langbein, *The German Advantage in Civil Procedure*, 52 U. CHI. L. REV. 823 (1985); Hein Kötz, *Civil Justice Systems in Europe and the United States*, 13 DUKE J. OF COMPARATIVE & INT'L L. 61 (2003). These systems have also been referred to as "nonadversarial systems," see, e.g., Edward A. Tomlinson, *Nonadversarial Justice: The French Experience*, 42 MARYLAND L. REV. 131 (1983). Though the label "inquisitorial" is subject to some controversy, *see* Kötz, *Civil Justice Systems in Europe and the United States*, 66 (describing the labels "inquisitorial" as "misleading because it conjures up the Spanish Inquisition, Kafka's castle, and bureaucratic omnipotence,"), the suggested connotations of the term "inquisitorial" do not seem to reflect the contemporary understanding of inquisitorial legal systems.
[99] *See* Bruce L. Hay & Kathryn E. Spier, *Burdens of Proof in Civil Litigation: An Economic Perspective*, 26 J. OF LEGAL STUDIES 413, 413 (1997).

the production of evidence and arguments are primarily driven by the adversarial parties themselves. A inquisitorial framework, meanwhile, is one where the judge plays the primary role in investigating facts, and the parties and their attorneys' role in gathering evidence is much more limited.[100]

While many inquisitorial systems still retain a number of "adversarial" features,[101] this shift in emphasis from the parties to the judge has a key ripple effect on the overall legal system.[102] As John Langbein notes, the German courts' inquisitorial design significantly shapes the rest of its civil procedure. For example, he points out that the inquisitorial system produces a much more flexible sequence for the various stages of litigation. While an adversarial model has set sequences for plaintiff and defendant presentation or participation in various parts of the litigation, "in German procedure the court ranges over the entire case, constantly looking for the jugular–for the issue of law or fact that might dispose of the case."[103] Consequently, the inquisitorial system, at least in Germany, has an "episodic character," where the flexibility of inquisitorial processes allow a continuous trial process that allows rehearing of issues through multiple points in time.[104] Additionally, Langbein notes that the inquisitorial structure has a significant impact on the use of witnesses and the role they play in producing facts or evidence before the court. In the adversarial system, the parties are largely responsible for supplying the witnesses, preparing the witnesses, and direct and cross-examining the witnesses.[105] The inquisitorial system, meanwhile, has the judge manage the task of summoning witnesses and

[100] *See* Langbein, *supra* note 94 at 824.
[101] *Id.*; *see also* Kötz, *supra* note 94 at 66-67 (describing similarities between the two systems).
[102] *See generally*, Langbein, *supra* note 94 (describing the differences that the German inquisitorial system has on the substantiation of a complaint, judicial case management, discovery, solicitation and examination of witnesses, and expert testimony).
[103] *Id.* at 830.
[104] *Id.* at 831.
[105] *See* Martin Marcus, *Above the Fray or Into the Breach: The Judge's Role in New York's Adversarial System of Criminal Justice*, 57 Brooklyn L. Rev. 1193, 1194 1992).

directing their examination in court.[106] These are but two examples of the larger effects that an adversarial or inquisitorial system may have in influencing the overall character of a legal institution's civil procedure. Consequently, when constructing a law of attribution, this feature of legal design should be one determined at the outset.

Arguments can be mustered in favor of either system. Advocates of the adversarial system extol the virtues of adversarial cross-examination as the most robust tool for exposing falsehoods,[107] point to potential efficiency in a system whereby parties specialize in presenting and securing evidence, and judges specialize in inferring from given evidence,[108] and point to the possibility that a inquisitorial judge may prejudges the outcome of a case, omitting crucial evidence or arguments that might shed further light on the dispute.[109] Advocates of the inquisitorial system point to the excessive partisanship and showmanship that shades into an adversarial process which may end up distorting the facts and evidence[110] and tilt the system into one that favors those with more resources and better counsel.[111] Amongst all this back and forth, scholars have employed a number of theoretical and empirical models to test the efficacy of such systems. Some mathematical models suggest that there would be little difference in either system's capacity to produce accurate or ideal outcomes,[112] while other models or studies say

---

[106] *Id.* at 828, 837.
[107] *See* John H. Wigmore, 5 *Wigmore on Evidence* 1367, at 32 (Chadbourn rev, Little Brown, 1974).
[108] Jeffrey S. Parker, *Daubert's Debut: The Supreme Court, the Economics of Scientific Evidence, and the Adversarial System*, 4 Supreme Court Economic Rev 1, 12-13 (1995).
[109] *See* Kötz, *supra* note 94 at 65. *But see* Carrie Menkel-Meadow, *The Trouble with the Adversary System in a Postmodern, Multicultural World*, 38 WILLIAM AND MARY L. REV. 5 (1996) (suggesting that even a binary oppositional system does not present a sufficiently high number of viewpoints to capture the nuances of truth).
[110] *See* JEROME FRANK, COURTS ON TRIAL: MYTHS AND REALITY IN AMERICAN JUSTICE, 86 (1949); Kötz *supra* note 94 at 65; Langbein *supra* note 94 at 833.
[111] *See* Russell G. Pearce, *Redressing Inequality in the Market for Justice: Why Access to Lawyers Will Never Solve the Problem and Why Rethinking the Role of Judges Will Help*, 73 FORDHAM L. REV. 969 (2004); Gillian K. Hadfield, *The Price of Law: How the Market for Lawyers Distorts the Justice System*, 98 MICHIGAN L. REV. 953 (2000).
[112] *See* Luke M. Froeb & Bruce H. Kobayashi, *Evidence Production in Adversarial vs. Inquisitorial Regimes*, 70 Economic Letters 267 (2001).

that the outcome depends on the particular data that an individual is measuring.[113] While the debate between models of legal design has long raged on, and will likely see no resolution in the near future, it is no controversial claim to suggest that perhaps each model may operate better in different contexts. Consider Langbein, who, despite favoring the general efficacy of inquisitorial systems, acknowledges that the adversarial system may receive unique justifications in the context of criminal law.[114]

Here, I suggest that the adversarial model is more uniquely suited to the context of attribution. This is largely because the advantages of inquisitorial legal systems are nullified by the international setting. First, inquisitorial systems depend upon a preexisting, centralized judicial authority that can be trusted to objectively seek the truth, and such an institution is lacking in the international realm. Second, because attribution frequently relies on technical evidence, and evidence that is acquired through espionage or other covert intelligence gathering, the parties themselves will almost always be in the best position to acquire and present such evidence in attribution disputes.

First, the inquisitorial system's dependence upon the judiciary to drive its procedure is largely a weakness in the international context. While a number of international courts do exist, these courts have incomplete jurisdiction, or are dedicated to specialized subject-matter that fails

[113] *See* Francesco Parisi, *Rent-Seeking Through Litigation: Adversarial and Inquisitorial Systems Compared*, 22 Int'l Rev. of L. and Econ. 193 (2002) (concluding that the adversarial system's costs are more apparent when evaluating each system through the lens of the Nash Equilibrium and considering litigation expenditure); Blair H. Sheppard & Neil Vidmar, *Adversary Pretrial Procedures and Testimonial Evidence: Effect of Lawyer's Role and Machiavellianism*, 39 J. of Personality and Social Psychology 320 (1980) (measuring the adversarial system's likelihood of reducing bias in the judge hearing a case); E. ALLAN LIND & TOM R. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE, 27-30 (1988) (reviewing studies that favored the adversarial system based on subjects' perceived "ratings of procedural fairness and satisfaction").

[114] *See* Langbein *supra* note 94 at 842.

to cover the attribution question presented here.[115] The International Court of Justice ("ICJ") is the best possible option in the current international framework, given its generally broad consideration of subject-matter.[116] However, even the ICJ has limited reach–the ICJ can settle disputes between states only to the extent that states seek to make use of it.[117] Following the court's ruling against the United States in *Nicaragua v. United States*,[118] for example, the United States withdrew from the compulsory jurisdiction of the ICJ.[119] Moreover, the enforcement powers of the ICJ are limited by the fact that enforcement is carried out through the Security Council, which allows members of the Security Council to thwart enforcement of its rulings, as the United States did in *Nicaragua*.[120] Since the inquisitorial system's emphasis on the managerial judge presumes a heightened degree of trust in the legitimacy of the institutional judiciary that directs much of its proceedings, the political nature of these international claims may make states less likely to participate in a process driven more by courts than the parties themselves.

Second, the inquisitorial system presumes that its judges have enough expertise to seek out the relevant information that will resolve a case, such as knowing which (expert) witnesses to seek and how to conduct their examination. But in the context of attribution, this may very well be a case of judges not knowing what they do not know. Given the technical nature of cyber-attacks and attribution, parties may justifiably view a generalized court as less reliable in taking

---

[115] Because this paper is concerned with state-to-state disputes, courts like the International Criminal Court, for instance, provide no answer because their jurisdiction is solely limited to prosecuting *individuals* for their conduct under international law.
[116] *See* HUGH THIRLWAY, THE INTERNATIONAL COURT OF JUSTICE, 27 (2016).
[117] *Id.*
[118] Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States), 1986 I.C.J. 14 [hereinafter *Nicaragua*].
[119] *See* Paul W. Kahn, *From Nuremberg to the Hague: The United States Position in* Nicaragua v. United States *and the Development of International Law*, 12 YALE J. OF INT'L L. 1, 2 (1987).
[120] *Subjects of UN Security Council Vetoes*, https://www.globalpolicy.org/component/content/article/102/40069.html.

the lead on the production of facts and evidence. Even if this concern could be addressed by conducting its proceedings under a panel of judges with technical expertise,[121] such judges would still fall short when it comes to their relative position in ascertaining the precise facts at issue in a particular dispute. A judge might not have as much familiarity with each states' cyber capabilities and their operations, nor the underlying evidence that might support one state's allegations that another was responsible for a cyber-attack. Since much of the evidence surrounding cyber-attacks and cyber-security might also be derived from covert intelligence operations,[122] the adversarial system might be more appropriate since the parties themselves would be in the best position to present or decide when to present certain sensitive evidence.[123]

The choice of an adversarial system for the attribution framework sets up a general picture of what our law of attribution might look like. Such a system will have an impartial adjudicator (how precisely those adjudicators are selected will be a subject for discussion later), and will largely be driven by the parties in terms of both legal argumentation and the production of facts and evidence. Consequently, such a system would contain a procedural sequencing similar to that of the American legal system, from initiation to discovery to the presentation of arguments, where arguments are structured around the parties' respective phases of argumentation.

---

[121] The Federal Circuit Court of Appeals, for instance, is known for specializing in technical matters, given the fact that it has nearly exclusive jurisdiction over patent appeals in the United States. *See Court Jurisdiction*, United States Court of Appeals for the Federal Circuit. http://www.cafc.uscourts.gov/the-court/court-jurisdiction.

[122] *See, e.g.*, Sanger & Fackler, *supra* note 69; Shane Harris, *U.S. Spies Say They Tracked 'Sony Hackers' for Years*, DAILY BEAST (Jan. 2, 2015, 6:55 PM), http://www.thedailybeast.com/articles/2015/01/02/u-s-spies-say-they-tracked-sony-hackers-for-years.html; Sam Biddle, *Top-Secret Snowden Document Reveals What the NSA Knew About Previous Russian Hacking*, INTERCEPT (Dec. 29, 2016, 10:26 AM), https://theintercept.com/2016/12/29/top-secret-snowden-document-reveals-what-the-nsa-knew-about-previous-russian-hacking/; Kate Connolly, *German Spy Chief Says Russian Hackers Could Disrupt Elections*, GUARDIAN (Nov. 29, 2016, 10:34 PM), https://www.theguardian.com/world/2016/nov/29/german-spy-chief-russian-hackers-could-disrupt-elections-bruno-kahl-cyber-attacks.

[123] For an economic analysis of how the burden of production might be optimized in an adversarial system, see generally, Hay & Spier, *supra* note 96.

**Standard of Proof**

With an adversarial framework in place, the next part of the picture to fill in is in establishing how the adversarial parties will succeed in proving their claim of attribution: the burden of proof they have in successfully proving their claim. The term "burden of proof" generally refers to two distinct concepts: the burden of persuasion and the burden of production (of evidence).[124] Since much of the section above addresses the burden of production (being placed on the parties in an adversarial setting), the term "burden of proof," as I use here, will refer to the burden of persuasion. Broadly speaking, the burden of persuasion refers to the confidence a trier of fact should have in coming to a legal conclusion after receiving all of the relevant facts and arguments presented by a case.[125]

The burden of proof is perhaps the most significant procedural rule that has bearing on the substantive outcome of a case. Robert Belton describes the burden of proof as "one of the most important procedural notions in our legal system" since "it helps implement the substantive laws by instructing the factfinder on the degree of confidence he should have in the correctness of factual conclusions for a particular type of case."[126] After all, the same set of facts may lead to entirely different outcomes based on the burden the parties have to prove their case.[127]

---

[124] *See* James Fleming, Jr., *Burdens of Proof*, 47 VIRGINIA L. REV. 51, 51 (1961); J. THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW 355-59 (1898).
[125] *Id.* at 52.
[126] Robert Belton, *Burdens of Pleading and Proof in Discrimination Cases: Toward a Theory of Procedural Justice*, 34 Vand. L. Rev. 1205, 1207 (1981).
[127] Consider the raised pleading standard established in *Iqbal v. Ashcroft*, 556 U.S. 662 (2009) and *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007), which Arthur Miller criticized as collapsing the distinction between summary judgment and the motion to dismiss phase (heightening the latter to the level of the former, which in effect forced the former standard to heighten in order to distinguish itself from the latter). *See* Arthur R. Miller, *From Conley to Twombly to Iqbal: A Double Play on the Federal Rules of Civil Procedure*, 60 DUKE L.J. 1, 15, 18 (2010). Although the pleading standard occupies a different context from the merits-phase of meeting a burden of proof, pleading standards entail their own burdens of proof for a case to proceed, which is the precise issue attracting controversy from the rulings in *Twombly* and *Iqbal*. Empirical studies to date have determined that the heightened pleading standard established by *Iqbal* and *Twombly* have had a statistically significant effect on diminishing

Some scholars have criticized the gradations between burdens of proof as having no clear or meaningful distinctions in the minds of a judge or jury.[128] However, these criticisms have been raised on a theoretical level; often, the empirical evidence mustered in support of these arguments have been based on surveys asking individuals to define or assign a probability value to various burdens of proof in the abstract.[129] But answers to surveys on the meaning of these burdens of proof may not be conclusive because the meaning of such terms are always understood in practice in relation to specific sets of facts.[130] Thus, a lack of consensus on the particular meaning of "clear and convincing" may not reflect factfinders' actual agreement as it pertains to a particular case, where a given set of factfinders may all agree that a party's evidence has established "clear and convincing" evidence. Furthermore, these theoretical arguments dismissing the role of the standards of proof seem unpersuasive when considering the empirical effect that the burdens of proof have had on the outcomes of cases in practice.[131] Given the

---

plaintiffs' access to the courts. *See* Theodore Eisenberg & Kevin M. Clermont, Essay, *Plaintiphobia in the Supreme Court*, 100 CORNELL L. REV. 193, 209 fn. 53 (2014) (analyzing over 18,000 cases to find a 14% increase in defendant's chance of winning pre-trial adjudication post-*Twombly*, and a 36% increase in the case of pro se plaintiffs); Patricia W. Hatamyar, *The Tao of Pleading: Do Twombly and Iqbal Matter Empirically?*, 59 AM. U. L. REV. 553, 556 (2010) (finding that after *Twombly*, the number of 12(b) motions to dismiss granted increased from 46 percent to 48 percent, and that after Iqbal, granted 12(b) motions rose to 56 percent); Joseph A. Seiner, *Pleading Disability*, 51 B.C. L. REV. 95, 118 (2010) (noting that dismissals increased from 54.2 percent to 64.6 percent in disability cases after *Twombly*).

[128] *See* C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293 (1982). In fact, some studies suggest that burdens of persuasion may have the opposite effect–that the standard way of explaining the burden of proof beyond reasonable doubt, in fact, may lead juries to be more likely to convict in criminal cases than in civil ones. *See* Lawrence Solan, *Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt*, 78 TEX. L. REV. 105 (1999-2000).

[129] *See* McCauliff, *supra* note 125.

[130] *See* Louis Kaplow, *Burden of Proof*, 121 YALE L.J. 738, 809 (2012) ("Answers to surveys on the meaning of 'more likely than not' may convey little, for the suggestion here is that its meaning in practice can depend very much on the circumstances."); Erik Lillquist. *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 146-83 (2002-2003) (suggesting that the variability of jury understanding of "reasonable doubt" may be an appropriate response to the particular types of cases observed by the jury).

[131] *See* Dennis J. Devine, Laura D. Clayton, Benjamin B. Dunford, Rasmy Seying & Jennifer Pryce, *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOLOGY, PUBLIC POLICY, AND LAW 622 (2001) (observing in a literature review that five studies demonstrated that "the wording used to convey the standard of proof has a substantial impact on jury verdicts"); Ashley Provencher, Josh Gupta-Kagan & Mary Eschelbach-Hansen, *The Standard of Proof at Adjudication of Abuse or Neglect: Its Influence on Case Outcomes at Key Junctures*, 17 SOCIAL WORK & SOCIAL SCIENCES REVIEW 5 (2014); citations *supra* note 121.

significant weight that the burden of proof has on a legal system's proceedings, it is important for us to decide how high we wish to set the burden of proof for the law of attribution.

There are three classic standards used for the burden of proof: proving a case by the preponderance of the evidence, proving a case by clear and convincing evidence, and proving a case beyond a reasonable doubt.[132] A preponderance of the evidence standard straightforwardly requires that a factfinder believes the existence of the fact (or legal outcome) to be more likely than its nonexistence,[133] roughly allocating the burdens of proof equally across both parties.[134] A clear and convincing standard is described by the Supreme Court as an "intermediate standard," that imposes somewhat higher requirements for persuasion than that of clear and convincing, though still a level of persuasion short of that reserved for those beyond a reasonable doubt.[135] Finally, the standard of "beyond a reasonable doubt" represents the highest burden of proof, meant to ensure the highest possible protection for the defendant against the possibility of an erroneous judgment.[136]

Though this spectrum for burdens of proof is well established, the normative underpinnings for when each standard ought to apply is much less clear. James Fleming wrote that "[t]here is no satisfactory test for allocating the burden of proof in either sense on any given issue."[137] Robert Belton echoed that statement, noting that "the courts have not yet developed any universal rule or set of policy considerations for courts to rely on in determining how the three burdens should be allocated between the parties."[138] It is true that the preponderance standard has long been the standard for civil proceedings in the United States, and reasonable

[132] *See* J.P. McBaine, *Burden of Proof: Degrees of Belief*, 32 CALI. L. REV. 242, 245 (1944).
[133] *See* Belton, *supra* note 123 at 1220.
[134] *See* Addington v. Texas, 441 U.S. 418, 423 (1979).
[135] *Id.* at 424.
[136] *Id.*
[137] Fleming, *supra* note 121 at 58.
[138] Belton, *supra* note 123 at 1217.

doubt has likewise been the principle rule for American criminal justice proceedings.[139] However, these standards have become associated with their respective proceedings mostly as a matter of tradition, lacking particularized justification, particularly for the standard used in civil proceedings.[140] This is especially clear when contrasting the United States' legal system to those of other countries. A number of countries with inquisitorial traditions, such as Germany, apply the reasonable doubt standard to all legal questions that their courts confront, no matter the subject matter.[141] So different burdens of proof can most certainly be employed for any one particular legal context. In the case of attribution, how do we go about choosing which burden of proof to apply?

While there may be no singular test for choosing a standard for the burden of proof, there are general principles that do shape this selection. As Belton notes, "Many different burden allocation tests have emerged from the cases and literature, but there is little consensus on a favored approach. All the tests, however, are grounded in considerations such as policy rationales, fairness, and the probability that the event in question actually occurred."[142] Fleming also concludes that similar overarching principles of fairness, convenience, and policy drive the decisions setting a standard for burdens of proof.[143] Besides these more general principles, Fleming acknowledges the relevance of other considerations, such as a party's relative access to evidence, the extent to which a party's contention departs from ordinary human experience, and substantive considerations that might employed the burdens of proof as handicaps against disfavored contentions.[144]

---

[139] *Id.*
[140] *See* Kaplow, *supra* note 127 at 742.
[141] *See* Kevin M. Clermont & Emily Sherwin, *A Comparative View of Standards of Proof*, 50 AM. J. OF COMPARATIVE LAW 243, 245 (2002).
[142] Belton, *supra* note 123 at 1217-18.
[143] *See* Fleming, *supra* note 121 at 60.
[144] *Id.* at 58-61.

While Belton and Fleming's descriptions seem conventionally true, they also don't seem to be very helpful. Fairness, convenience, and policy, as broad justifications, could apply to almost any legal construction, and in any direction. The more specific considerations that they proffer are a step in the right direction, but even then, the confluence of multiple considerations risks making the endeavor into a multi-factor marionette, one that can be pulled in any particular manner based on the puppeteer and the string that they wish to pull.

Instead, Louis Kaplow places these considerations along a more concrete frame of reference, approaching the burdens of proof with an economic analysis of how each burden of proof might best accomplish the legal system's goals.[145] The burden of proof is specifically seen as a tool for adjusting two main probabilistic outcomes: the probability of imposing liability on someone who conducted harmful behavior, and the probability of creating erroneous liability on someone behaving benignly or productively.[146] For Kaplow, the burden of proof must walk the tightrope balance between deterring harmful acts and avoiding the chilling of productive ones. In this line of thought, it is essential to consider asymmetric error costs, [147] since these error calculations often dictate how our procedural rules tilt the playing field, including the way we set our burdens of proof.

The classic example is that of criminal punishment–because it is "better to let ten guilty persons go free than to convict one innocent person," we justify "many defendant-favoring rules of criminal procedure," including a high burden of proof.[148] For attribution, the error costs seem

[145] *See generally*, Kaplow, *supra* note 127.
[146] *Id.* at 745-46.
[147] *See* Jacob Gersen & Adrian Vermeule, *Thin Rationality Review*, 114 MICH. L. REV. 1355, 1395-96 (2016); *see also* David H. Kaye, *Statistical Significance and the Burden of Persuasion*, 46 Statistical Inference in Litigation 13, 16 (1983) (describing the Supreme Court's reasoning behind burden of proof cases as involving the error costs at play).
[148] *Id.*

less clearly skewed towards one side or the other. Is it better to let a cyber-attacking state go free than to punish one innocent state? Assuming that the cyber-attack is serious enough to rise to the level of armed force,[149] and assuming the range of countermeasures short of a military strike,[150] it is not necessarily clear whether the harm of the former is less serious than the latter, especially if the latter is supposed to be constrained by rules of proportionality.[151]

For a law of attribution, then, the preponderance of the evidence is most suitable to achieve the overall aims for a system of attribution; in cases where military action is the only (or threatened) response to a cyber-attack, the burden of proof should ratchet up to the reasonable doubt standard. As a baseline burden of proof, demonstrating attribution by a preponderance of the evidence seems most appropriate for two main reasons. First, a lower burden of proof produces a lower evidence threshold that increases the chance of producing legal judgment, increasing the risk of liability, and promoting the deterrence of malicious behavior. Second, it allocates the burden of persuasion roughly equally among parties, challenging both parties to optimally produce information and evidence regarding the origins of a cyber-attack.

While it is possible to conceive of an even lower burden of proof (strict liability, for example), the preponderance standard is the most preferable point of balance because it mandates that a certain degree of information to be presented to establish a prima facie case, and then renders a judgment based on a comparative analysis of the information provided by both

---

[149] In other words, a cyber-attack that produces net effects equivalent to a kinetic armed strike. *See* Hathaway, *supra* note 22. Examples might include a cyber-attack that disrupts or destroys critical civilian infrastructure, such as a program disabling a power grid.

[150] Recall that the law of attribution might justifiably treat attribution for the purposes of military action as a unique category deserving of a higher burden of proof. *See* discussion *supra* 28-30. In this case, the asymmetric error costs of war might be quite similar to the classic asymmetric error costs of a criminal conviction, in which case the reasonable doubt standard offers the appropriate burden of proof to offset the disproportionate harm of erroneous military conflagration.

[151] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol 1), art. 51(5)(b), June 8, 1977, 1125 U.N.T.S. 3 [hereinafter Protocol Additional I]; *see also id.* art. 85(3)(b).

parties. This requirement then encourages a competitive information production from both the accusing party (the basis of their claim) as well as the accused party (who then has an interest in sharing countervailing intelligence to identify the true actor). The preponderance standard thereby results in an optimal level of information production, and greater information produced about international cyber-attacks more broadly helps tackle the uncertainty and transaction costs in state-to-state interactions that afflict the field of cybersecurity[152] and international relations more generally.[153]

A critic might object that the preponderance standard is an unfair one to the country defending itself from claims of attribution. After all, the preponderance places the burden equally across both parties, but the critic might argue that states in the defensive role are actually in a *weaker* position than that of the state bringing claims. Not only is there an asymmetry in information, since the state bringing an attribution claim may have (or claim to have) covert intelligence supporting its position, but the state in a defensive role also is essentially forced to rebut the allegations by proving a counterfactual–that they did not in fact launch the cyber-attack. Given the potentially complex technical skills needed to conduct an attribution, and the fact that a number of countries may have a dearth of individuals possessing such skills, some states may simply not have the resources to carry out countervailing attribution efforts to given the particular challenges raised by attribution. And unlike, say, the individual in a criminal or civil case, who can give an account of her alibi, the complex, many-membered state generally cannot give a full accounting of the entirety of its functions to display its honesty.

[152] *See generally*, Jason Li, Xinming Ou & Raj Rajagopalan, *Uncertainty and Risk Management in Cyber Situational Awareness*, CYBER SITUATIONAL AWARENESS – ISSUES AND RESEARCH (2010).
[153] *See* Brian C. Rathburn, *Uncertain about Uncertainty: Understanding the Multiple Meanings of a Crucial Concept in International Relations Theory*, 51 INT'L STUDIES QUARTERLY 533 (2007).

The counter, of course, is that corporations give an accounting of themselves all the time when acting as the defendants in civil suits. And while it is true that proving a counterfactual is difficult, especially in the case of hacking, this objection assumes that the prima facie case for attributing an attack to a state has already taken place. As discussed earlier, such a task is still a challenge, even using the preponderance standard. The preponderance standard is traditionally represented as the idea that a party needs to prove their claim with anything above a 0.5 probability.[154] But it is not enough to assume that an agnostic fact-finder begins exactly on the 0.5 line, and can be nudged over by the accuser. While it is practically true that adversarial frameworks force a factfinder to perform a comparative analysis of the two parties' claims,[155] the 0.5 probability assumes that the defendant is merely negating the plaintiff's claims, when in reality the defendant frequently proposes one or more counter narratives.

Rather than a strict tug-of-war of probabilistic truth over the plaintiff's narrative, then, a case turns on the ratio of the probabilistic truth of the plaintiff in relation to the probabilistic truth of the defendant's possible counter narrative(s).[156] Considering the context of cyber-attacks, the objection that the preponderance standard is plaintiff-skewed therefore makes a Bayesian probability error; rather than presuming the absolute truth of the plaintiff's accusations of attribution, these claims must be compared against the underlying probability that any one of a vast number of possible global actors was responsible for the attack. A defendant state can then reference any number of the technological or circumstantial bases for doubting an attribution.[157]

---

[154] *See, e.g.*, MCCORMICK'S HANDBOOK OF THE LAW OF EVIDENCE § 339, at 794 n.56 (2d ed. 1972); Edward K. Cheng, Reconceptualizing the Burden of Proof, 122 YALE L.J. 1254, 1256 (2013); Kaplow *supra* note 127 at 779; Ronald J. Allen, *Burdens of Proof*, 13 J. OF LAW, PROBABILITY & RISK 195, 203 (2014); Vern R. Walker, *Preponderance, Probability and Warranted Factfinding*, 62 BROOK. L. REV. 1075, 1097 (1996).
[155] Cheng, *supra* note 151 at 1259-60.
[156] *Id.* at 1259-62.
[157] *See* discussion *supra* 9-16.

Moreover, the information asymmetry that supposedly favors the accusing state is likely to be less favorable in practice because factfinders tend to express a greater degree of skepticism towards parties that withhold information. This has specifically been examined in the context of international, state-to-state adjudications before the ICJ, where the ICJ has responded to the withholding of evidence, usually on grounds of security, by liberally construing circumstantial evidence in favor of the party that lacks any access to the evidence that is withheld.[158] The principles behind the ICJ's actions logically extend to other forms or forums of international adjudication. If anything, the ICJ's response is a rather mild response to the withholding of evidence, given many domestic courts' tendency to make an actively adverse inference from the fact that a party withholds evidence.[159] Accordingly, a preponderance of the evidence standard would not result in an unfair plaintiff advantage when applied to the law of attribution.

Generally, then, a preponderance of the evidence standard fits the goals of attribution, since it provides the optimal balance of deterrence and information production since a lower burden lowers the barriers to attribution (and hence, increases the potential for countermeasures) while still requiring a requisite level of persuasion that would incentivize the production of relevant intelligence and information regarding the cyber-attack. In cases where a military strike is proposed or threatened as a countermeasure, the law of attribution should ratchet its burden of proof to the reasonable doubt standard, much for the same reasons that the standard is employed in American criminal law. The reasonable doubt standard recognizes the tremendously disproportionate error rates that accompany so serious of a penalty, and just as the risk of

[158] *See* Michael P. Scharf & Margaux Day, *The International Court of Justice's Treatment of Circumstantial Evidence and Adverse Inferences*, 13 Chi. J. of Int'l L. 123, 149-50 (2012).
[159] Dale A. Nance, *Adverse Inferences About Adverse Inferences: Restructuring juridical Roles for Responding to Evidence Tampering by Parties to Litigation*, 90 B.U. L. Rev. 1089, 1094 (2010).

erroneous criminal punishment presents a disproportionately intolerable harm, so too would an erroneous military conflict, perhaps on an exponentially higher scope and scale.

### Attributing Cyber-Attacks by Non-state Actors to States: State Responsibility Doctrine

The law of attribution is starting to look clearer. We have an adversarial model, following stages of procedure akin to the American and British legal systems, including rules for initiating an action, the back and forth sequencing of complaint and answer, and the adversarial discovery framework for producing evidence. We also have a general standard of proof to determine when a party has successfully proven that another state is responsible for launching a cyber-attack. But say that a state defends itself from attribution by placing the blame on "non-state actors" who happen to have operated within its borders. How do we attribute the malicious activity of non-state hackers to that of a state?

This is a particular problem for the law of attribution and cybersecurity, given the fact that the relatively low cost of conducting a cyber-attack opens up the option up to myriad non-state actors,[160] who may act for a variety of motivations. And all the typical problems associated with simply attributing an attack create further attenuation between the individual conducting the

[160] *See* Joseph S. Nye, Jr., *Cyber Power*, Belfer Center for Science and Int'l Affairs, 4-6, 9-11 (May 2010), http://www.belfercenter.org/sites/default/files/legacy/files/cyber-power.pdf. While digital technology has lowered the cost of entry to distribute cyber-attack capabilities more diffusely across a number of actors, as a note of caution, it is important to remember that certain high-magnitude cyber-attacks are still out of the reach of many, and that individuals do not have the same exact capabilities of government. *See Id.* at 11. Certain types of cyber-attacks may be as accessible by individuals as they are by governments–DDOS and botnet attacks, for example. But other sophisticated tools, such as ones that require decryption or zero-day exploits, are much less accessible to your ordinary hacker. Contrary to certain claims by individuals that their ten-year-old son "can do anything with a computer," Catherine Rampell, *How Trump's 10-Year-Old Son Could Guide U.S. Cybersecurity*, CHI. TRIBUNE (Jan. 3, 2017, 1:55PM), http://www.chicagotribune.com/news/opinion/commentary/ct-cybersecurity-computers-internet-trump-perspec-0104-20170103-story.html, young hackers can't quite do everything, at least to the same extent as governments. As Nye puts it, "A teenage hacker and a large government can both do considerable damage over the internet, but that does not make them equally powerful in the cyber domain. Power diffusion is not the same as power equalization." Nye, *Cyber Power* at 11.

hack and any chain of command or control infrastructure that might tie that actor to a state. After all, hackers don't wear uniforms in cyberspace. Thus, a law of attribution must address the inevitable result where it follows the trail to an individual hacker, and then faces the problem of how to connect that person to a state for the purposes of legal responsibility.

Fortunately, the state responsibility doctrine is a legal problem that exists beyond the realm of cyber-attacks, and has consequently been addressed before in other contexts.[161] Generally speaking, international law already possesses a state responsibility doctrine for attributing the malicious behavior of non-state actors to a state. The International Law Commission's 2001 Draft Articles on State Responsibility set out the ways in which international courts have found states to be responsible for non-state actors.[162] Articles 4 and 8 of the Draft Articles on State Responsibility have subsequently been recognized as customary international law by the ICJ,[163] and courts, commentators, and other sources have come to widely recognize these articles as setting the standard view of state responsibility doctrine under customary international law.[164] For example, both the first edition of the Tallinn Manual and the recently released second edition both draw heavily on the ILC's Draft Articles to formulate their conception of state responsibility doctrine in the setting of cyber-attacks.[165]

---

[161] *See, e.g.*, Oona A. Hathaway, Emily Chertoff, Lara Domínguez, Zachary Manfredi & Peter Tzeng, *Ensuring Responsibility: Common Article 1 and State Responsibility for Non-State Actors*, 95 Tex. L. Rev. (Forthcoming in 2017).

[162] International Law Commission, *Draft Articles on Responsibility of States for Internationally Wrongful Acts*, U.N. Doc. A/56/10 (2001) [hereinafter Draft Articles].

[163] 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, ¶¶ 385 (Article 4), 398 (Article 8) (Feb. 26) [hereinafter *Bosnian Genocide*].

[164] *See* Hathaway, Chertoff, Domínguez, Manfredi & Tzeng, *supra* note 158 at 9 (quoting JAMES CRAWFORD, STATE RESPONSIBILITY: THE GENERAL PART, 43 (2013) as saying that the ILC's Draft Articles "are considered by courts and commentators to be in whole or in large part an accurate codification of the customary international law of state responsibility").

[165] TALLINN MANUAL 1.0 at 29; TALLINN MANUAL 2.0 at 79.

The Draft Articles on State Responsibility find that a non-state actors' wrongful behavior is attributable to a state if a non-state actor is acting as an organ of the state or is acting under the instructions, directions, or control of the state. Article 4 of the Draft Articles holds actions attributable to a state when they are conducted by individuals who may be recognized as organs of the state. As Article 4 states:

> 1. The conduct of any State organ shall be considered an act of that State under international law, whether the organ exercises legislative, executive, judicial or any other functions, whatever position it holds in the organization of the State, and whatever its character as an organ of the central Government or of a territorial unit of the State.
>
> 2. An organ includes any person or entity which has that status in accordance with the internal law of the State[166]

As made clear in the commentary on Article 4, Article 4 also extends to individuals who may be considered *de facto* organs of the state.[167] Meanwhile, Article 8 of the Draft Articles also finds the actions of non-state actors attributable to a state if they are "acting on the instructions of, under the direction, or under the control of" a state.[168] The conditions for state responsibility described in Articles 4 and 8 generally have been understood as tests for the control a state has, either over the individual actor, or the action the individual actor has taken.[169] These control tests, in turns, echo the control tests that have been employed in rulings by courts like the ICJ.[170]

However, there are a number of limitations to the existing international law on state responsibility. Oona Hathaway, for instance, criticizes the current framework as creating perverse incentives, whereby states can still escape responsibility by handing illegal tasks to non-

[166] Draft Articles, *supra* note 158, art. 4.
[167] *See* International Law Commission, *Draft Articles on Responsibility of States for Internationally Wrongful Acts, with Commentaries*, at 47-48, *in* Report of the International Law Commission on the Work of Its Fifty-third Session, U.N. GAOR, 56th Sess., Supp. No. 10, U.N. Doc. A/56/10 (2001) [hereinafter Draft Articles Commentary], art. 4, cmt. 11.
[168] Draft Articles, *supra* note 158, art. 8.
[169] *See* Hathaway, Chertoff, Domínguez, Manfredi & Tzeng, *supra* note 160 at 11.
[170] *See, e.g.*, *Bosnian Genocide*, *supra* note 159; *Nicaragua*, *supra* note 115.

state actors so long as they maintain minimal oversight.[171] She also argues that the control test in fact disincentives efforts to control rogue or malicious behavior, since the attempts to impose control might create a sufficient degree of control to hold the state responsible for wrongdoing that the non-state actor commits, in spite of state efforts to police it.[172] Peter Margulies, significantly, criticizes the scope of state responsibility doctrine as applied to the task of attributing cyber-attacks, noting that the Draft Articles' control tests require a high bar of specific, comprehensive control, and that such standard would exclude very significant examples of states directing non-state actors in conducting a cyber-attack.[173]

Fortunately, these comments aren't just critical, but constructive, too, and Hathaway and Margulies propose adjustments to remedy these shortcomings in state responsibility rules. Margulies suggests the "virtual control test," where "the burden shifts to a state to demonstrate it was not responsible for a cyber attack when the state funds and equips a private entity or individual who subsequently engages in a cyber attack."[174] Under this test, Margulies appears to require some prima facie indication linking the accused state to the non-state entity.[175] However,

---

[171] *See* Hathaway, Chertoff, Domínguez, Manfredi & Tzeng, *supra* note 160 at 25-27.
[172] *Id.* at 27-28.
[173] *See* Peter Margulies, *Sovereignty and Cyber Attacks: Technology's Challenge to the Law of State Responsibility*, 14 MELB. J. INT'L L. 496, 506-07, 510-11 (2013).
[174] *Id.* at 5. Later in his article, Margulies expands the virtual control standard to also include burden-shifting to cases where "knowingly provides sanctuary to a private entity that subsequently engages in a cyber attack against another state." *Id.* at 19.
[175] Margulies is not very clear on the precise legal conditions for when the burden-shift happens. For example, he does not explain what the accusing state's burden of production or proof is, or what level of mens rea is required in order to trigger the burden-shift. Would the mere allegation of funding and equipping suffice to trigger the burden-shifting? Would the provision of computers for an entirely different purpose count as "funding and equipping" an entity for the virtual control test (if, for example, a rogue librarian with access to a government-provided computer decided to hack someone)? Margulies instead explains his virtual control test with a hypothetical example. He writes, "Suppose that Utopia was the victim of a cyber attack . . . After a sophisticated digital forensics investigation, Utopian officials concluded that the attack originated from an IP address assigned to the Ruritanian Resistance Group ('RRG'). . . . Initial intelligence reports suggested that the RRG received funding and software from Ruritania. Ruritania's assistance to the RRG therefore met the 'virtual control' standard outlined here." Presumably, Utopia has made some sort of public demonstration of the results of its "digital forensics investigation" and "[i]nitial intelligence reports" in order to then trigger legal burden-shifting upon Ruritania (else the existence of those facts would not be legally relevant), indicating some sort of initial, prima facie burden on Utopia, though the precise requirements of that initial burden are still not clarified by his example.

this suggested approach to state responsibility runs some of the risk described by Hathaway in the current regime, where the potential attachment of liability to any existing relationship between the government and a non-state actor might instead incentivize governments to relinquish any control over the non-state actors within its reach. Margulies might counter that the "funding and equipping" requirement means that the virtual control test only requires Governments to exercise such oversight in cases where it materially supports such entities, that Governments naturally have an incentivize to fund non-state entities in all manner of contexts, and that in cases where they do so, there should be a presumed expectation of oversight. The problem with this argument is Margulies' undefined definition of what it means to fund and/or equip a non-state entity, and the potentially broad scope of these terms essentially erases this limitation.[176]

Of course, these concerns are easily remedied by defining these terms with greater specificity; alternatively, Hathaway's proposal of an affirmative defense to claims of state responsibility can complementarily tackle the problem of perverse incentives. Hathaway proposes a similarly broad obligation on behalf of states to "ensure respect" under Common Article 1 of the Geneva Conventions by ensuring that non-state actors within its reach do not engage in cyber-attacks.[177] While this raises a parallel fear about incentivizing states to distance themselves instead of regulating, Hathaway addresses this concern with the idea that states

[176] *Cf.* Holder v. Humanitarian Law Project, 561 U.S. 1 (2010), which held that "[m]aterial support meant to 'promot[e] peacable, lawful conduct' can further terrorism" merely by freeing up resources. *Id.* at 25. Even when the material support statute at issue had a mens rea requirement, the Court interpreted the mens rea requirement merely to require knowledge that the entity at issue was a designated foreign terrorist organization, not knowledge that the support at issue may be used to support terrorist activity. Thus, there is a dual problem of not knowing the mens rea sufficient to trigger the burden-shifting, and not knowing which elements to which the mens rea requirement might apply.

[177] Hathaway, Chertoff, Domínguez, Manfredi & Tzeng, *supra* note 160 at 1, 40.

should have an affirmative defense if states can prove that they took "reasonable steps" to prevent violations of international law.[178]

By incorporating these proposals into its procedural rules, the law of attribution can not only advance the doctrines of state responsibility, it can do so to successfully address the novel challenges of cyber-attack attribution with the similarly novel solutions that Hathaway and Margulies present. A more charitable association between non-state actors and the state they are tied to–through the virtual control test–combined with an affirmative defense of "reasonable care" should allow a law of attribution to attribute individuals' cyber-attacks to states, while allowing states the proper means of protecting themselves from liability when they take good faith measures to prevent wrongdoing.

**Sensitive Intelligence & Evidentiary Rules**

Suppose a state has suffered a cyber-attack and wishes to bring a legal claim attributing that attack to another state. With everything laid out so far, the state knows the procedure for initiating an action, and the back-and-forth sequencing of complaint and answer, summary judgment arguments back and forth, and the production of the evidence. Here, in this last step, the state runs into a problem: what happens if significant portions of the evidence it relies on is derived from covert intelligence?[179] Moreover, states may have plausible factual bases for attributing an attack, but may not want to disclose such evidence on legitimate grounds, since cyber-attackers could learn from those points of attribution and avoid leaving the same fingerprints in the future. Here, the law of attribution faces the challenge of reconciling the need

[178] *Id.* at 42-46.
[179] As noted previously, many of the recent major cyber-attacks have been attributed to actors on the basis of covert intelligence. *See supra* notes 71-72, 120.

to present such evidence with states' desires to preserve the secrecy of their confidential intelligence and their sources.

The adversarial system addresses this dilemma to some extent: since the parties have control over pushing forward a claim, one answer is to simply dismiss this problem out of hand and say, tough luck, the onus falls on the state to decide what to do in such a situation. Under a cost-benefit calculation, this position would say that such disclosure is to price to pay for seeking recourse against a cyber-aggressor, and that it would entirely be up to the state to weigh the benefits of seeking recourse versus the costs of disclosing information about their covert intelligence capacities. The problem with this approach is that it assumes that the costs of cyber-attacks are purely internal to the states subject to the precise attack at issue. If, however, we understand cyber-attacks to be a general, global, and iterative phenomenon,[180] and that a state unchecked in its cyber-aggression will proceed to conduct future cyber-attacks against others, then the act of attribution (and the fact that it enables countermeasures to deter future attacks) produces positive externalities that are not accounted for in the "tough luck" mindset.

Consequently, a law of attribution should strive to accommodate the state's secrecy and attribution interests simultaneously by finding a way to allow states to present sensitive intelligence as evidence while preserving the secrecy of such evidence from the broader public. This is not the first time, though, that courts have grappled with the role of sensitive intelligence in court. Courts have long balanced the sensitive security concerns of states with the public role of courts, and have developed a number of managerial tools to protect the information produced

[180] Which is particularly true of cyber-attacks, given how easily the tools of cyber-attack can be disseminated to other actors. For example, almost immediately after the Mirai botnet attacks, the code used for the attack was dumped online for anybody to copy and use themselves. *See* Robert Hackett, *Why a Hacker Dumped Code Behind Colossal Website-Trampling Botnet*, Fortune (Oct. 3, 2016), http://fortune.com/2016/10/03/botnet-code-ddos-hacker/.

or used in a hearing. There are two primary procedures that a law of attribution can incorporate to accommodate states' desires to protect classified information. First, courts can have procedures for hearing evidence *ex parte* and *in camera*, and second, courts can seal their dockets and records when such records concern classified information.

A number of national courts employ such procedures to secure classified information when it is necessary to prove a claim in court. In the United States, the Foreign Surveillance Intelligence Act of 1978 (FISA) created the Foreign Intelligence Surveillance Court,[181] which reviews federal law enforcement and intelligence officers' request for surveillance warrants.[182] The Foreign Intelligence Surveillance Court conducts its proceedings *ex parte* and *in camera*, with few of its rulings ever reaching the public.[183] These procedural moves are not limited to specialized courts. The Classified Information Procedure Act allows U.S. courts in criminal cases to review classified information *ex parte* and *in camera* to determine if the evidence is essential for a fair trial or criminal due process requirements.[184] And, as a general matter, in civil claims brought before a federal court, Federal Rule 26 allows sealing of court records on good cause.[185]

Other countries possess similar procedures for shielding proceedings or evidence used at trial. The United Kingdom passed the Justice and Security Act in 2013, creating closed material procedures (CMPs), which are secret court hearings where only the judge and specialized

[181] Foreign Intelligence Surveillance Act (FISA), 50 U.S.C. §§ 1801-1885c.
[182] 50 U.S.C. §1804.
[183] *See* Conor Clarke, Essay, *Is the Foreign Intelligence Surveillance Court Really a Rubber Stamp? Ex Parte Proceedings and the FISC Win Rate*, 66 STAN. L. REV. 125 (2014); Eric Lichtblau, *In Secret, Court Vastly Broadens Powers of N.S.A.*, N.Y. TIMES (Jul. 6, 2013), http://www.nytimes.com/2013/07/07/us/in-secret-court-vastly-broadens-powers-of-nsa.html.
[184] Classified Information Procedures Act (CIPA), 18 U.S.C. app. III §§ 1–16; *see also* Fred F. Manget, *Intelligence and the Criminal Law System*, 17 STAN. L. & POL'Y REV. 415, 424 (2006).
[185] Fed. R. Civ. P. 26(c); *see also* Hon. T. S. Ellis, III, *Sealing, Judicial Transparency and Judicial Independence*, 53 Vill. L. Rev. 939, 945 (2008); David A. Schulz, *Rethinking Confidentiality and Access in Civil Litigation*, 23 COMM. LAW. 24, 25 (2005-2006).

security-cleared advocates are given access to any sensitive intelligence at issue in the case.[186] Similarly, the Netherlands possesses the Act on Shielded Witnesses, which provides for a special procedure where a special magistrate can hear representatives of the Netherlands' two main intelligence agencies on whether certain information should stay secret, or whether certain witnesses should have their identity cloaked in anonymity.[187] Such evidence is used in Dutch administrative, civil, and criminal cases, and this procedure, like that of the U.S. FISA Courts, is largely conducted *ex parte* and *in camera*, though it is possible for the parties to the case to be present when the special magistrate evaluates the sensitive intelligence.[188] Germany and Spain, meanwhile, prohibit the use of secret evidence at trial, though testimony or anonymous information based on secret evidence may sometimes be permitted for use.[189]

*Ex parte* and *in camera* procedures benefit the law of attribution in a number of ways. Adding these types of proceedings creates flexibility for the system, allowing the factfinders to analyze the issues that sensitive intelligence raises on a case-by-case basis. *Ex parte* proceedings, in particular may allow the factfinder to negotiate with a party on issues of disclosure, since parties may tend to overestimate the cost of disclosing their own information, as a form of loss-aversion.[190] *In camera* proceedings allow sensitive evidence to have their full evidentiary value, while mitigating the cost of disclosure more generally.[191]

---

[186] Justice and Security Act (2013), http://www.legislation.gov.uk/ukpga/2013/18/contents; *see also* Directorate-General for Internal Policies, Policy Dep't, *National Intelligence and Secret Evidence in Legislation and Before the Courts: Exploring the Challenges*, Study for the LIBE Committee, 21-25 (2014) [hereinafter *National Security and Secret Evidence*], http://www.europarl.epa.eu/RegData/etudes/STUD/2014/509991/IPOL_STU(2014)509991_EN.pdf.

[187] Staatsblad van het Koninkrijk der Nederlanden (2006), www.eerstekamer.nl/behandeling/20061024/publicatie_wet_14/document3/f=/w29743st.pdf; *see also National Security and Secret Evidence*, *supra* note 184 at 25-26.

[188] *See National Security and Secret Evidence*, *supra* note 184 at 25-26.

[189] *Id.* at 27-28.

[190] *See* Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, 5 J. OF ECON. PERSPECTIVES, 193, 199-203 (1991).

[191] Of course, procedures need to be put in place to impose sanctions on a state for breaking the terms of the *in camera* proceedings, which a state may do in reckless rage were a court to make an adverse finding against it. Even if both parties comply with the nondisclosure requirements of the proceeding, however, *in camera* proceedings may

There are, of course, costs to having secrecy rules in a legal proceeding. Transparency in a legal proceeding tends to bestow upon it a greater air of legitimacy,[192] while secrecy instead might serve to undermine it. Furthermore, if one of the overriding goals of the law of attribution is to justify a later countermeasure in the eyes of the international community, a secret hearing to justify an attribution might leave many in the international community skeptical of the counter-acting state's claim to legitimacy. Can a law of attribution legitimize countermeasures behind closed doors?[193]

This is a difficult question whose answer revolves around the question of where courts or legal judgments derive their authority. While it is true that the open display of a judicial proceeding may command some legitimacy to the process by virtue of its transparency, it does not follow that openness is the only sufficient or necessary element of what makes a judicial judgment binding. After all, the countries discussed previously have successfully incorporated measures of secrecy into their legal systems without undermining the legitimacy of their legal rulings.[194] Of course, those institutions did not begin with closed proceedings, nor do most of them necessarily shield the majority of their cases behind closed proceedings. It may be that

---

still have shortcomings since the information will inevitably be disclosed to the opposing party. This is especially concerning in the realm of attribution, given the fact that sensitive intelligence that tends to attribute an attack is most likely sensitive intelligence that the attributing state collected from the attributed state, and the disclosure is most undesirable when it results in the spying state revealing its intelligence to the very state who is being spied on.

[192] *See* Richmond Newspapers, Inc. v. Virginia, 448 U.S. 555, 569 (1980) (describing "the importance of openness to the proper functioning of a trial; it gave assurance that the proceedings were conducted fairly to all concerned, and it discouraged perjury, the misconduct of participants, and decisions based on secret bias or partiality"); Press-Enter. Co. v. Superior Court of California for Riverside Cty., 478 U.S. 1, 9 (1986) (holding that "openness in criminal trials, including the selection of jurors, 'enhances both the basic fairness of the criminal trial and the appearance of fairness so essential to public confidence in the system'").

[193] A judgment of attribution need not necessarily be tied to a subsequent countermeasure or sanction against the state determined to be responsible for a cyber-attack. In this case, attribution might serve as a symbolic shaming, "outing" the guilty party to the world. It seems doubtful, though, that states would expend the time and resources to acquire a legal judgment of attribution purely for just its symbolic effect.

[194] Though the more secretive proceedings do tend to attract some criticism and controversy. *See, e.g.*, Alan Butler, *Standing Up to Clapper: How to Increase Transparency and Oversight of FISA Surveillance*, 48 NEW ENG. L. REV. 55 (2013); Ellen Yaroshefsky, *Secret Evidence is Slowly Eroding the Adversary System: CIPA and FISA in the Courts*, 34 HOFSTRA L. REV. 1063 (2005-2006).

closure of certain records is accepted by the society bound by such law because those judicial institutions have already built up legitimacy through a general openness of proceedings through time.

While this may seem difficult for a new international legal system to do, there is some suggestion from surveys of public opinion that international courts derive their legitimacy in the public eye not from the specific legitimacy that individual courts have accrued, but from the general trust that the public has in international institutions and their own systems of law.[195] If members of the public trust in international institutions generally and in their own domestic courts, that trust bleeds over into support for international courts. This comports with broader jurisprudential accounts of authority, which suggest that it is the office or institution of courts that claim authority, and not merely the pure power to persuade.[196] Thus, it is not necessarily the public presentation of evidence or the persuasiveness of a particular adjudicator's reasoning that compels adherence to the ruling of an adjudicator[197]–rather, it is the fact of process that produces this credibility. After all, in the United States, the large majority of cases brought before federal appellate courts are terminated via unpublished "no-opinion" orders, indicating that the

[195] *See* Eric Voeten, *Public Opinion and the Legitimacy of International Courts*, 14 THEORETICAL INQUIRIES IN LAW 411 (2013). While it's true that the public opinion of citizens may not map perfectly onto the views of states, and we are concerned with legitimacy in the eyes of state in this context, states themselves are bound by their entanglement and commitment to many of these international institutions, meaning that they, too, are probably subject to buy-in in terms of these legal institutions' legitimacy.

[196] *See, e.g.*, Joseph Raz, *Authority, Law and Morality*, 68 MONIST 295, 299 (1985). Raz offers his preemption thesis, a component of authority, as "[t]he fact that an authority requires performance of an action is a reason for its performance which is not to be added to all other relevant reasons when assessing what to do, but should replace some of them." *Id.* By describing the judgment of authority as not merely one "to be added to all other relevant reasons when assessing what to do," *Id.* Raz is acknowledging that authority is not merely an exercise in persuasion among all the other factors that might persuade an individual, but instead ascribes authority to the general aspect of the institution that itself provides a heuristic authority superseding or supplanting the general process of pure reasoning that might otherwise produce further controversy.

[197] After all, courts' opinions fall subject to criticism, academic and in popular opinion, all of the time. *See, e.g.*, David L. Shapiro, *In Defense of Judicial Candor*, 100 HARV. L. REV. 731, 731 (1987).

resolution of legal controversies does not demand a purely transparent window into the legal process.[198]

In fact, other international courts have maintained their legitimacy, despite including the use of secret proceedings. The European Court of Human Rights, for instance, encountered this precise issue in *A v. United Kingdom*, where the ECHR reviewed the United Kingdom's procedure for permitting detention of an individual on evidence that included "secret material," concluding that there was no excessive or unjustified secrecy employed.[199] The accumulation of this empirical experience, from both national and international courts, demonstrates that the law of attribution can very easily employ the methods of *in camera* and *ex parte* proceedings. Of course, these procedures cannot be applied haphazardly, but must judiciously be used with the appropriate procedural rigor, but their existence allows states to present sensitive intelligence in claims of attribution while preserving the secrecy of that intelligence.

**The legal framework for a law of attribution**

In sum, the law of attribution possesses the following characteristics. First, it operates as an adversarial institution, where both claims and the record are largely developed by the parties. Second, being an adversarial framework, the rules of procedure also temporally sequence the stages of a case in the back-and-forth manner that characterizes a typical adversarial legal proceeding. Third, upon reaching the merits, an accusing state must prove their claim of attribution by the preponderance of the evidence, except in instances where the accusing state

[198] *See* Patricia Wald, *The Rhetoric of Results and the Results of Rhetoric: Judicial Writings*, 62 U. CHI. L. REV. 1371, 1373 n. 3 (1995).
[199] A v. United Kingdom, 49 EHRR 29 (2009); *see also* Daniel Alati, Ronnie Dennis, Ryan Goss, Alecia Johns, Esther Kuforiji, Paul Troop & Keiran Hardy, *The Use of Secret Evidence in Judicial Proceedings: A Comparative Survey*, Research Paper Prepared for the Joint Committee on Human Rights, 17 (2011).

wishes to employ a military countermeasure. In cases where a state has not disclosed its planned countermeasure, or where such an option is still uncertain, the case may proceed on the preponderance standard, but will not be sufficient to justify later military action. Fourth, to meet this burden of proof, states will have the option of employing procedures like *in camera* review, *ex parte* hearing, and the sealing of records in order to use sensitive evidence to prove their claims. Fifth and finally, the state proving their attribution claim needs to specifically prove that the attack can be linked to individuals operating on the behest of a state or under the control of a state, where the control test will be interpreted charitably under the "virtual control test" espoused by Margulies. Simultaneously, though, states will have the affirmative defense of demonstrating due diligence in policing the relevant non-state actors.

## Part III: Models for Implementing the Law of Attribution

With the legal framework for attribution drawn out, how can this theory be fully fleshed out and brought to life? The next part of this paper addresses the more policy-oriented side of attribution, which mainly raises questions of setting: where the judgment will take place, and by whom. These questions of venue and forum are invariably tied to the crucial, practical requirement of answering of why states will have the incentive to even participate in such a legal system. The issue of state compliance with international institutions or laws is, of course, a vast subject of discussion all in itself.[200] Structural explanations of international law and institutions

[200] *See, e.g.*, GLOBAL GOVERNANCE (ed. Lisa Martin, 2008); JACK L. GOLDSMITH & ERIC A. POSNER, THE LIMITS OF INTERNATIONAL LAW (2005); Oona A. Hathaway, *Between Power and Principle: An Integrated Theory of International Law*, 72 U. CHI. L. REV. 469 (2005); Mattias Kumm, *The Legitimacy of International Law: A Constitutionalist Framework of Analysis*, 15 EUR. J. INT'L. L. 907 (2004); George W. Downs & Michael A. Jones, *Reputation, Compliance, and International Law*, 31 J. OF LEGAL STUDIES S95 (2002); Andrew T. Guzman, *A Compliance-Based Theory of International Law*, 90 CAL. L. REV. 1823 (2002); Beth A. Simmons, *Capacity, Commitment, and Compliance: International Institutions and Territorial Disputes*, 46 J. OF CONFLICT RESOLUTION 829 (2002); Beth A. Simmons, *International Law and State Behavior: Commitment and Compliance in International Monetary Affairs*, 94 AM. POLITICAL SCI. REV. 819 (2000); Anne-Marie Slaughter, Andrew S.

run the gamut, from Kantian philosophy[201] to rational choice theory.[202] And discussions of state compliance in specific subject-matters have popped up in nearly every context, including criminal law,[203] environmental law,[204] and human rights law.[205]

While this paper can proffer general, structural analysis regarding state incentives to participate, the problem of state cooperation or compliance is as much a political question as it is a legal one. In order to produce a fully predictive claim for how states might involve themselves in such a legal framework, a proposal would have to call upon international relations, both broadly in theory and specific to this historical moment, behavioral economics to analyze incentives, costs, and the probabilities of behavior given the various actors in play, and specific historical and psychological analysis of many of the players who might be important in bringing about such a legal regime.

A full answer to the questions raised by the challenge of international compliance is far too broad to answer within the bounds of this paper. This paper instead takes the more modest approach of discussing the general incentives for state buy-in by surveying various other forms of international adjudication. Thus, I examine three examples of international adjudication: the International Court of Justice, the World Trade Organization's dispute settlement process, and ad hoc systems like the US-Iran Tribunal. Each institution reflects a different approach to international adjudication, providing models for how international institutions have succeeded in

---

Tulumello & Stepan Wood, *International Law and International Relations Theory: A New Generation of Interdisciplinary Scholarship*, 92 AM. J. OF INT'L L. 367 (1998); Harold Hongju Koh, Book Review, *Why Do Nations Obey International Law?*, 106 YALE L.J. 2599 (1997).

[201] *See* Fernando R. Tesón, *The Kantian Theory of International Law*, 92 COLUM. L. REV. 53 (1992).

[202] *See* ANDREW T. GUZMAN, HOW INTERNATIONAL LAW WORKS: A RATIONAL CHOICE THEORY (2008).

[203] *See* Beth A. Simmons & Allison Danner, *Credible Commitments and the International Criminal Court*, 64 INT'L ORG. 225 (2010).

[204] *See, e.g.*, Daniel Brodansky, *The Legitimacy of International Governance: A Coming Challenge for International Environmental Law?*, 93 Am. J. of Int'l L. 596 (1999).

[205] *See, e.g.*, BETH A. SIMMONS, MOBILIZING FOR HUMAN RIGHTS: INTERNATIONAL LAW IN DOMESTIC POLITICS (2009).

getting states to participate in their systems. The ICJ presents the option of incorporating the law of attribution within an existing forum that has broad subject matter jurisdiction. The WTO's dispute resolution process reflects an adjudicatory system with specialized subject matter, and the US-Iran Tribunal models an ad hoc, state-to-state approach, that may more flexibly resolve conflicts between two particular states, but lacks the power create more lasting legal power.

**International Court of Justice**

The International Court of Justice is the paradigm example of an international legal institution. Established by the United Nations Charter in 1946,[206] the ICJ was the only international court in existence for much of the twentieth century.[207] Consequently, the ICJ not only serves as a model for creating a new international legal system–it provides an existing forum where the law of attribution might be incorporated. As a general matter, the ICJ has broad subject-matter jurisdiction to hear any international law claim brought before it, so long as it is brought with the consent of both parties.[208] Incorporating the law of attribution into the ICJ would have the advantage of attaching the law of attribution to a preexisting institution that has established credibility, institutional history, and fully developed rules and resources.

Prior to the creation of the ICJ, several attempts had been made at creating international institutions for state-to-state dispute resolution. The Permanent Court of Arbitration, for example, was created following the Hague Peace Conference of 1899.[209] Despite its name, the Permanent Court of Arbitration was not a permanent standing court, but instead provided an

[206] *See* THIRLWAY, *supra* note 115 at 3.
[207] *See* Pierre-Marie Dupuy, *The Danger of Fragmentation or Unification of the International Legal System and the International Court of Justice*, 31 N.Y.U. J. of Int'l L. and Politics 791, 791 (1999).
[208] *See* THIRLWAY, *supra* note 115 at 35.
[209] *See* ROBERT KOLB, THE ELGAR COMPANION TO THE INTERNATIONAL COURT OF JUSTICE, 6 (2014).

administrative organization where states could select arbitrators from a pool of candidates and create their own tribunal to resolve disputes.[210] And although the PCA provided a set of procedural rules, these rules were mere defaults that would be overridden by whatever choice of rules the state parties elected to institute themselves.[211] After the creation of the PCA in 1899, a follow-up conference took place in 1907, where several states, including the United States, proposed the creation of an actual, permanent court.[212]

Though the proposals in 1907 failed to gain traction at the time, the devastation wrought by World War One spurred movement towards the creation of an international court, finally culminating in the precursor to the ICJ: the Permanent Court of International Justice ("PCIJ).[213] The PCIJ was created in 1921 under the League of Nations.[214] In its twenty-five year tenure,[215] the PCIJ produced 32 judgments, all of which were implemented.[216] The PCIJ also issued 27 advisory opinions in this period, with states adhering to or acting upon most of these advisory rulings.[217] All in all, the PCIJ laid a successful groundwork for the later ICJ.[218]

The ICJ was created with the establishment of the United Nations Charter in 1946, and was modeled closely after the PCIJ.[219] The ICJ is composed of fifteen judges elected by the Security Council and General Assembly.[220] These members are elected in separate elections, with nine year terms, with elections focusing on the judges as individuals, and not as

---

[210] *Id.*
[211] *Id.*
[212] *Id.* at 7.
[213] *Id.* at 12.
[214] *Id.*
[215] The PCIJ existed from 1921 until 1946, when the present ICJ was established. *See* THIRLWAY, *supra* note 115 at 3.
[216] *See* KOLB, *supra* note 208 at 12.
[217] *Id.*
[218] The dissolution of the PCIJ was mainly due to its close attachment to the League of Nations, which itself was dissolved in the aftermath of the WW2. *See id.* at 22-24.
[219] *See* THIRLWAY, *supra* note 115 at 3.
[220] See THIRLWAY, supra note 115 at 9.

representatives of their countries.[221] The ICJ also incorporates a number of rules to ensure the independence of its judiciary. These include rules requiring members of the court to make a solemn declaration of impartiality in the performance of their duties, and the ICJ strives to eliminate conflicts of interest[222] by prohibiting its members from "exercis[ing] any political or administrative function, or engag[ing] in any other occupation of a professional nature" in their time as judges on the court.[223] Furthermore, members of the ICJ cannot be removed unless the rest of the members unanimously find that the judge has failed to fulfill their duties.[224]

Articles 34 through 38 of the Statute of the International Court of Justice lay out the ICJ's jurisdiction, giving it grounds to consider all legal disputes[225] concerning:

> a. the interpretation of a treaty;
> b. any question of international law;
> c. the existence of any fact which, if established, would constitute a breach of an international obligation;
> d. the nature or extent of the reparation to be made for the breach of an international obligation.[226]

---

[221] *Id.* The specific length of the nine year term is a holdover from the PCIJ, and attempts to strike the balance of providing judges with a secure tenure so as to not have their decision making corrupted by the politics of re-election; simultaneously, appointment for life was undesirable since the court strove to have the judicial membership represent the diverse bodies of nations that were party to the court. *See id.*

[222] Cases involving a judge's state of national origin do not create cause for recusal; reasons for recusal are determined in Articles 17 and 24 of the Statute, which require the judge not to participate only if the judge has previously participated in the case for one of the parties, Statute of the International Court of Justice, art. 17 ¶ 2, or in cases involving a "special reason" for recusal, Statute of the International Court of Justice, art. 24. ¶¶ 1-2.

[223] *See id.* at 11-12.

[224] *See* Statute of the International Court of Justice, art. 18 ¶ 1.

[225] Here, someone might object that the requirement of a legal "dispute" precludes the ICJ from hearing a claim of attribution because the limitation of jurisdiction to "disputes" sounds similar to the standing requirement in U.S. law. The party making this claim might argue that the attribution is an incomplete claim since the declaratory ruling of attribution is insufficient to redress the real harm at issue (the cyber-attack). This argument, however, is no obstacle given the ICJ's broad interpretation of what counts as a dispute. ICJ rulings demonstrate that the elements of showing a dispute simply entail "the claim of one party is positively opposed by the other," and that "the matter is one of substance, not of form." *Id.* at 54 (citing *South West Africa (Preliminary Objections)*, 1962 ICJ Rep 328 and *Application of the International Convention on the Elimination of All Forms of Racial Discrimination, Georgia v. Russian Federation*, 2011 ICJ Rep 84 ¶ 30).

[226] Statute of the International Court of Justice, art. 36 ¶ 2.

Cyber-attacks, and the law of attribution, certainly touch upon legal questions falling within the ICJ's purview. Cyber-attacks potentially rise to a level of armed force in violation of Article 2(4),[227] while also posing potential violations of the doctrines of state sovereignty and neutrality.[228] Attribution, as a necessarily ancillary question to that of cyber-attack, implicates such questions of international law. While the ICJ has not yet heard any disputes concerning the use of cyber-attacks, the jurisdictional scope outlined above appears to place such disputes well within its bounds.[229]

With this general overview, we can now ask: what factors led to the ICJ's formation, and how lessons might those teach for implementing the law of attribution? It is difficult to dissociate the creation of the ICJ (and its predecessor, the PCIJ) from the historical moments that gave birth to these two institutions. The First and Second World Wars no doubt played a significant role not only in the creation of these courts, but the international organizations that these courts are tied to. As historical lessons, they appear to teach the story of international law arising in response to international tragedy. As a narrative, this is both encouraging and troubling. It is encouraging because it suggests the possibility of states embracing the creation of new international laws and institutions to deal with contemporary challenges like that of cyber-attack and global cybersecurity. It is troubling because it may be that states are compelled to create such institutions only when such challenges have grown to the degree where they result in an

[227] *See* Hathaway, *supra* note 23; Waxman, *supra* note 43.
[228] *See* TALLINN MANUAL 2.0 at 11-29, 553-562.
[229] *See List of Cases Referred to the Court Since 1946 by Date of Introduction*, International Court of Justice, http://www.icj-cij.org/docket/index.php?p1=3&p2=2 (last visited Apr. 17, 2017). The closest case appears to be a ruling issued in *Timor-Leste v. Australia*, in which concerned Australia's seizure of documents and data from legal advisors to Timor-Leste. Questions relating to the Seizure and Detention of Certain Documents and Data (Timor-Leste v. Australia), Judgment, 2013 I.C.J. Rep. 156 (Mar. 3). The third prong of the ICJ order, for instance, commands "that Australia shall not interfere in any way in communications between Timor-Leste and its legal advisers in connection with" a pending maritime arbitration. *Id.* In this case, however, the seizure of electronic data simply accompanied the physical seizure of documents from an office, meaning that the ruling did not examine the issue of cyber-attack, cyber-espionage, or any other related digital breach of sovereignty.

international catastrophe or event that causes widespread harm. Such broad generalizations, of course, are not the end all be all for the practical implementation of the law of attribution. After all, more localized events like the Estonia cyber-attack have spurred groups such as the one that came together to create the Tallinn Manual and its sequel, hinting at the possibility of preemptive, rather than reactive implementation of international law.

**WTO Dispute Settlement System**

A second model for implementing the law of attribution would be through an institution such as the World Trade Organization's dispute settlement process. Unlike the ICJ model, which provides a standing court with broad subject-matter jurisdiction, The WTO's dispute settlement system is a model that attaches an adjudicatory process to an international body with a specific subject-matter focus. Employing this kind of model would have the advantage of implementing the law of attribution through a specialized body of factfinders who might be best equipped to address the technical complexity of the evidence and techniques by which states and their experts trace malicious digital activity back to its creators.

The WTO was created under the Marrakesh Agreement, one of the several agreements made in the 1994 Uruguay Round.[230] The WTO was generally formed to promote and oversee global trade, and the WTO's dispute settlement system is one of the express functions laid out in Article III of the Marrakesh Agreement that are meant to help the institution achieve such a goal.[231] Meanwhile, the structure and procedure of the WTO's dispute settlement process is laid out more precisely in the Understanding on Rules and Procedures Governing the Use of Disputes

[230] Marrakesh Agreement Establishing the World Trade Organization, Apr. 15, 1994, 1867 U.N.T.S. 154 [hereinafter Marrakesh Agreement].
[231] *Id.* art. III.

("DSU").[232] Under Article 1 of the DSU, the dispute settlement process can be applied to disputes covered under a number of specified agreements, including the 1994 General Agreement on Trade and Tariffs[233] and the Agreement on Trade-Related Aspects of Intellectual Property Rights.[234]

The dispute settlement process is administered by the Dispute Settlement Board ("DSB"), which oversees the operation of WTO's settlement panels and the implementation of their rulings. The actual function of the panels is determined by the rules set out by the DSU. These rules include provisions for establishing adjudicatory panels, the composition of such panels, panel procedures, and various other ground rules for how the panel is to perform its decision making process. For instance, the DSU prescribes the conditions for initiating a dispute settlement panel, stating that the DSB shall create a settlement panel when a complaining party requests one "in writing," and that such request "shall indicate whether consultations were held, identify the specific measures at issue and provide a brief summary of the legal basis of the complaint sufficient to present the problem clearly."[235] Additionally, the DSU regulates the composition of its panels, imposing requirements such as the fact that none of the panelists may be from a country party to a dispute (unless stipulated to by both parties).[236] In terms of the decision making process, the DSU's provisions also require its panels to create specific timelines

[232] Understanding on Rules and Procedures Governing the Settlement of Disputes art. 1, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 2, 1869 U.N.T.S. 401 [hereinafter DSU].
[233] Multilateral Agreements on Trade in Goods, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1869 U.N.T.S. 401.
[234] Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, THE LEGAL TEXTS: THE RESULTS OF THE URUGUAY ROUND OF MULTILATERAL TRADE NEGOTIATIONS 320 (1999), 1869 U.N.T.S. 299, 33 I.L.M. 1197 (1994) [hereinafter TRIPS Agreement], available at https://www.wto.org/english/docs_e/legal_e/27-trips.pdf.
[235] DSU art. 6 ¶¶ 1-2.
[236] *Id.* art. 8 ¶ 3.

for its decisions,[237] specific stages of review and the procedures for those specific stages,[238] and the types of information that the panel may review or consult.[239] Accordingly, the DSU lays out a comprehensive regime for adjudication.

Naturally, such an institution has attracted scholarly attention to how effective it has been in inducing state participation and compliance. On the issue of state participation, a more specialized forum may raise the concern that more powerful states with a vested interest in this subject area may use such an institution merely as a means to throw their weight around. Chad P. Bown, for example, produced an empirical study suggesting that a country's retaliatory capacities, legal capacities, and its role in international political-economic relationships were significant in measuring that state's likelihood of participating in the dispute resolution system.[240] If this is the case, one concern might be that a specialized institution simply becomes a tool for powerful states to institutionalize their dominant power in certain domains such as trade or cybersecurity. Of course, this problem may simply be a feature of asymmetric international power, or a problem of wealth inequality affecting law more generally.[241]

---

[237] *Id.* art. 12 ¶¶ 3-12.
[238] *Id.* art. 15.
[239] *Id.* art. 13 (giving panels the right to "seek information and technical advice from any individual or body which it deems appropriate" so long as notice is provided to the parties); *id.* art. 18 ¶ 1 (forbidding *ex parte* contacts concerning the case under consideration).
[240] Chad P. Bown, Participation in WTO Dispute Settlement: Complainants, Interested Parties, and Free Riders, 19 WORLD BANK ECON. REV. 287, 308 (2005) (concluding that "Even after controlling for the economic importance of disputed sector market access, variables that serve as proxies for the institutional bias generated by the current rules of the system also affect the nonparticipation choice. . . . despite market access interests in a dispute, an exporting country is less likely to participate in WTO litigation if it has inadequate power for trade retaliation, if it is poor and does not have the capacity to absorb substantial legal costs, if it is particularly reliant on the respondent country for bilateral assistance, or if it is engaged with the respondent in a preferential trade agreement").
[241] *See, e.g.*, Edward Glaeser, Jose Scheinkman & Andrei Shleifer, *The Injustice of Inequality*, 50 J. MONETARY ECON. 199 (2003); Beverly Moran & Stephanie M. Wildman, *Race and Wealth Disparity: The Role of Law and the Legal System*, 34 FORDHAM URB. L.J. 1219, 1235-36 (2007) ("Access to lawyers and the legal system is another form of wealth. . . . legal rules have tremendous impact on the protection of property rights, the creation of bargaining power, and the determination of wealth distribution. Just as legal rules act to concentrate other types of wealth, such as education, housing, and tax benefits, legal resources are yet another type of wealth that remains unevenly distributed.").

In the end, even if there is a participation bias towards certain states, if we understand systems of law to be valuable not merely in adjudicating claims for one party or another, but for the positive externalities that the institution of law brings in creating greater predictability and cooperation among states, then the skew in participation may be a tolerable price to pay. Other empirical studies suggest that this may be the case. Michael Bechtel and Thomas Sattler, for instance, find that there is minimal difference in the economic benefits given to complainant parties and passive third parties that sign onto the claims brought by complainants before the WTO.[242] Such results indicate that "weaker" states have the option of freeriding on the efforts of more powerful, gaining the benefits of increased trade, and that the adjudicatory process produces spillover benefits that may benefit state more broadly. And to the extent that the WTO dispute settlement process has been effective in engendering compliance from parties that do come before it,[243] the compliance produced by this process, and the positive externalities that follow, may very well provide the tale of a successful international adjudicatory regime.

Not only does the WTO dispute resolution system offer a model of international adjudication–the story of how the TRIPS agreement came to be incorporated into the WTO offers a practical lesson for how certain legal regimes might be folded into international institutions with larger buy-in. In Private Power, Public Law, Susan Sells traces the history of how the TRIPS agreement came to be woven into the fabric of the WTO.[244] In this historical narrative, Sells draws attention to the "central player in this drama," the "US-based twelve

[242] Michael M. Bechtel & Thomas Sattler, What is Litigation in the World Trade Organization Worth?, 69 INT'L ORG. 375, 395-96 (2015).
[243] *See* Robert Howse, *The World Trade Organization 20 Years On: Global Governance by Judiciary*, 27 EUR. J. INT'L L. 9 (2016); Bruce Wilson, *Compliance by WTO Members with Adverse WTO Dispute Settlement Rulings: The Record to Date*, 10 J. INT'L ECON. L. 397 (2007).
[244] SUSAN K. SELLS, PRIVATE POWER, PUBLIC LAW: THE GLOBALIZATION OF INTELLECTUAL PROPERTY RIGHTS (2003). The TRIPS agreement was an agreement that institutionalized a stringent and enforceable global intellectual property regime. *See* TRIPS Agreement, *supra* note 233; *id.* at 1.

member Intellectual Property Committee" that consisted of twelve chief executive officers representing various industries.[245] The role of concentrated lobbying, then, can play a prominent role in implementing certain regulatory regimes into international law, and in getting states to act as strong advocates of such systems. Given the increasingly high risk that cyber-attacks pose to private commercial entities–the Sony attack, for example, or the Yahoo cyberattack[246]–there is a definite opportunity for commercial companies to play a prominent role in lobbying to successfully institutionalize international regimes like the law of attribution.

**Mass Claims Commissions (US-Iran tribunal)**

A third model for implementing a law of attribution would be through ad hoc tribunals, such as the Iran-United States Claims Tribunal created in 1981.[247] The United States-Iran tribunal is an example of a purely bilateral mass claims commission that came into existence through a treaty made between two states.[248] Unlike the prior two models, then, this sort of system arises in response to a specific set of claims between two parties. This approach has the advantage of flexibility, allowing implementation tailored to specific circumstances and parties involved, though it also comes at the cost of having its effect be limited in scope, both in terms

[245] SELLS, 1.
[246] *See* Mike Levine & Emily Shapiro, *How Russian Agents Allegedly Directed Massive Yahoo Cyberattack*, ABC NEWS (Mar. 15, 2017, 4:34 PM), http://abcnews.go.com/US/russian-agents-facing-charges-yahoo-hacking-attacks/story?id=46142396.
[247] Declaration of the Government of the Democratic and Popular Republic of Algeria Concerning the Settlement of Claims by the Government of the United States of America and the Government of the Islamic Republic of Iran, 19 January 1981, art. II(2) [hereinafter "Claims Settlement Declaration"]. Though it was created to adjudicate a specific set of claims between Iran and the United States, the Iran-United States Claims Tribunal, like the ICJ, was also physically seated at the Hague. *See* KOLB, *supra* note 208 at 53.
[248] While there are examples of mass claims commissions that operated through the United Nations (such as the UN Compensation Commission), as opposed to directly between two states, this section's emphasis is on the bilateral nature of such ad hoc arrangements, not their particular function specific to mass claims.

of the parties subject to such an ad hoc tribunal, and to the historical events that are justiciable under the tribunal.

The Tribunal was created as part of agreement to resolve the Iranian Hostage Crisis.[249] In the Revolution of 1979, Iranians stormed the U.S. embassy in Tehran, taking 69 people captive.[250] While a number of the hostages were released, 52 remained captive for 444 days.[251] The Algiers Accords helped broker an agreement between the United States and Iran, where Iran would release the American hostages in exchange for the United States removing trade sanctions and the freeze it had imposed on a number of Iranian assets.[252] Significantly, the Algiers Accord also sought to address a multitude of private claims that U.S. citizens raised against Iran, and that Iranian citizens raised against the U.S.[253] The Algiers Accord addressed these by shifting them from litigation to arbitration–and hence, the formation of the Iran-United States Claims Tribunal.

The Claims Settlement Declaration formally established the Tribunal, including the terms of its jurisdiction, composition, and arbitral rules.[254] Jurisdictionally, the Tribunal was limited to hearing to two categories of claims[255]: 1) claims "of nationals of the United States against Iran and claims of nationals of Iran against the United States, and any counterclaim which arises out of the same contract, transaction or occurrence that constitutes the subject matter of that national's claim,"[256] and 2) official claims "of the United States and Iran against each other

[249] *See* Richard M. Mosk, *Lessons from the Hague – An Update on the Iran-United States Claims Tribunal*, 14 PEPP. L. REV. 819, 819-20 (1987).
[250] Muhammad Sahimi, *The Hostage Crisis, 30 Years On*, PBS FRONTLINE (Nov. 3, 2009, 1:30 PM), http://www.pbs.org/wgbh/pages/frontline/tehranbureau/2009/11/30-years-after-the-hostage-crisis.html.
[251] *Id.*
[252] *See* Hosk, *supra* note 247 at 820.
[253] *Id.*
[254] Claims Settlement Declaration, art. II(1).
[255] Besides limiting claims based on their substance, the Tribunal also limited claims procedurally by requiring them to be filed with the Tribunal by Jan. 19, 1982. *See Id.* Thus, the Tribunal's procedural rules also served to limit and funnel the historical scope of the claims that the Tribunal would reach.
[256] *Id.*

arising out of contractual arrangements between them for the purchase and sale of goods and services."[257] In establishing its adjudicators, the Claim Settlement Declaration determined that the Tribunal was to be composed of nine members, three appointed by the United States, three appointed by Iran, with those six members then appointing the last three members of the Tribunal.[258]

For its procedures, the Tribunal adopted the arbitral rules of the United Nations Commission on International Trade Law (UNCITRAL).[259] These rules, in turn, created a comprehensive set of procedures that governed the stages of hearing, including the method of conducting examination, the production of evidence, and the use of expert testimony.[260] These rules also provided a significant degree of flexibility and discretion to the arbitration Tribunal in its use of various procedural mechanisms, such as when or how it would incorporate expert evidence.[261] The incorporation of the UNCITRAL rules, then, provides an example of how a preexisting set of rules can be incorporated or woven into specific ad hoc adjudicatory institutions. This in turn suggests a similar possibility for how ad hoc institutions might do the same with the law of attribution.

As a general matter, the Iran-United States Claims Tribunal was successful in processing a large number of claims on both sides. All of the claims brought by the United States were

---

[257] *Id.* art. II(2).
[258] *Id.* art. III(1).
[259] *Id.* art. III(2).
[260] http://www.uncitral.org/pdf/english/texts/arbitration/arb-rules-2013/UNCITRAL-Arbitration-Rules-2013-e.pdf
[261] *See* Karl-Heinz Bockstiegel, *Applying the UNCITRAL Rules: The Experience of the Iran-United States Claims Tribunal*, 4 BERKELEY J. INT'L L. 266, 267 (1986) ("It is clear that the broad base and inherent elasticity of the UNCITRAL Rules are features which have proved invaluable in laying a firm foundation for the development of these rules. Changes have been introduced, however, to accommodate the special needs of this unique arbitral body as its work has proceeded."); Michael Straus, *The Practice of the Iran-U.S. Claims Tribunal in Receiving Evidence from Parties and from Experts*, 3 J. INT'L ARB. 57, 63 n. 7 (1986) (noting, for example, the discretion granted to the Tribunal under Article 25(4) to allow persons identified as a party or party representative to remain in the room during a hearing, as part of the discretion "to determine the manner in which witnesses are examined," as well as the general exercise of discretion in evaluating conditions for summoning and presenting expert testimony).

decided,[262] and those decided in favor of US claimants were all paid in full.[263] On the Iranian side, the United States recently agreed in 2016 to pay a settlement of $1.3 billion dollars to settle one of its longstanding claims.[264] For some, then, the Tribunal presented much for cause for celebration.[265] These supporters point to the Tribunal's track record, and the fact that it has processed over 3900 cases since its inception, which generally covers all but a few large and complex claims between the two states.[266] Beyond the number of cases it has addressed, others, like Richard M. Mosk, have lauded the Tribunal for its ability to practically and successfully implement a full suite of procedural rules for adjudicating its cases, rules that helped to effectively navigate complicated cases such that its procedures "may serve as guides for future tribunals."[267] In fact, the Tribunal is also serving as a guide in other ways–one study by Christopher Gibson and Christopher Drahozal demonstrates that Iran-United States Claims Tribunal decisions have been cited as precedent by the ICSID Tribunal,[268] suggesting that an ad hoc Tribunal may still allow its implementation of law to have broader effect beyond the immediate historical controversies that it adjudicates.

There are limitations, however, on raising attribution claims with an ad hoc approach. Despite the fact that the Iran-United States Claims Tribunals' decisions have been cited in other

---

[262] IUSCT, Communiqué, 16 May 2016, available at: http://www.iusct.net/General%20Documents/Communique%2016.1%20(9%20May%202016).pdf.
[263] *See* Charles N. Brower, *Lessons to be Drawn from the Iran-U.S. Claims Tribunal*, 9 J. INT'L ARB. 51, 51 (1992).
[264] *See* Elise Labott, Nicole Gaouette & Kevin Liptak, *US Sent Plane with $400 Million in Cash to Iran*, CNN (Aug. 4, 2016), http://www.cnn.com/2016/08/03/politics/us-sends-plane-iran-400-million-cash/.
[265] Others have levied a number of criticisms towards the way the Tribunal functioned. Charles N. Brower, for example, noted that the judges "could never seem to agree on anything very much and adopt a uniform Tribunal jurisprudence, even on fairly simple issues." Brower, *supra* note 261 at 54. Brower also took serious issue with the Tribunal's ability to as well as the fact that some 2500 of these claims were resolved with lump-sum payments, precluding a truly individualized assessment of claims that, in his eyes, produces an inadequate remedy. Charles N. Brower, *Lessons to be Drawn from the Iran-U.S. Claims Tribunal*, 9 J. INT'L ARB. 51 (1992).
[266] Stephen M. Schwebel & Ruth Teitelbaum, *The Latest Award from the Iran-United States Claims Tribunal: The Line Between Approximation of Damages and Ruling ex Aequo et Bono*, 109 AM. J. INT'L L. 369, 369 (2015).
[267] *See* Mosk, *supra* note 248 at 822-23.
[268] Christopher S. Gibson & Christopher R. Drahozal, *Iran-United States Claims Tribunal Precedent in Investor-State Arbitration*, 23 J. INT'L ARB. 521, 540, 543-44 (2006).

tribunals, more general surveys of arbitration citations demonstrate that arbitration courts' citation of cases tends to vary significantly according to context; while the Convention on Contracts for the International Sale of Goods and the ICC had relatively few citations to prior awards, the Court of Arbitration for Sports and domain name arbitration systems had nearly ubiquitous citation of precedent in their rulings.[269] In the case of attribution, it is easy to see these rulings going to the way of the former. Given the wide range of factual variation in cyber-attack attribution cases–ranging from the type of cyber-attack[270] to the level of secrecy attached to a state's evidence supporting attribution–tribunals would likely be reluctant to rely too heavily on prior cases given their potential for factual dissimilarity.

But ad hoc tribunals also face a particularly unique challenge in establishing the incentives for participation. Because they frequently arise out of bilateral agreements, they depend on states having (or treating each other as having) relatively equal standing. Moreover, they depend upon particular historical contexts where both states have sufficient grievances against the other to provide the incentive to form such tribunal in the first place. While such a circumstance is certainly possible in the cyber-attack context–states may have scourged each other with mutual cyber-aggression–it is difficult to imagine states owning up to this fact and approaching the other with the desire to call it quits, and it is especially difficult to imagine states having sufficiently equal leverage in this context to force both to the bargaining table. And even where there is sufficient incentive for states to form these ad hoc tribunals, a crucial limitation is that ad hoc tribunals are reactive to such harm, and therefore seem very after-the-fact and

[269] Christopher R. Drahozal, *Empirical Findings on International Arbitration: An Overview*, OXFORD HANDBOOK ON INTERNATIONAL LAW, 38-39 (Forthcoming) (citing a study by Gabrielle Kaufmann-Kohler).
[270] *See, e.g.*, Bonnie Zhu, Anthony Joseph &Shankar Sastry, *A Taxonomy of Cyber Attacks on SCADA Systems*, 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, 380, 383-87 (listing types of attacks as including things such hardware attacks, buffer overflows, SQL injections, diagnostic server attacks, Address Resolution Protocol Spoofing, chain/loop attacks, SYN floods, and DNS forgery).

retrospective, rather than forward looking.[271] While it is true that the previous two models can only adjudicate claims over attacks that have already happened, the sheer fact of a standing judicial institutions represents a temporal longevity that allows its decisions to cast a greater shadow on the future. Thus, the ad hoc model, while perhaps most effective in particular factual circumstances that might call for it, is more generally a model that seems less effective for implementing the law of attribution.

## Conclusion

When describing the origins of the International Court of Justice, Robert Kolb breaks down its path into three parts:

> The organization of a comprehensive scheme of arbitral justice;
> The attempt to create a permanent and compulsory 'arbitral court';
> The creation of an institutional court, linked to the League of Nations – the Permanent Court of International Justice (PCIJ).[272]

Crucially, the first step to the creation of this regime was the creation of the legal scheme– something has to first be imagined before it can be created. And with each step, the vision of law becomes incrementally more specific, until that vision has taken the form of an actual institution of law. The law of attribution proposed here seeks to begin drawing that vision for how states can redress the threat of cyber-attacks through law. The law of attribution, of course, is a far more modest project than the initial concept of an international court of justice. But it is

[271] *See* Ralph Zacklin, *The Failings of Ad Hoc International Tribunals*, 2 J. Int'l Crim. Just. 541, 542 (2004). While Zackling appears equally critical of standing international courts' (i.e. the International Criminal Court) ability to do better, more recent systematic assessments demonstrate that standing courts like the ICC do have some deterrent effect. *See* Hyeran Jo & Beth A. Simmons, *Can the International Criminal Court Deter Atrocity?*, 70 Int'l Org. 443 (2016).
[272] Kolb, *supra* note 208 at 5.

nonetheless an important one, and one made all the more possible by the foundations laid by prior institutions of international law.

It imagines a legal framework for attributing a cyber-attack to the state responsibility, and proposes the procedural rules that would allow a state to legitimately make such a claim. By adopting an adversarial model, the law of attribution can situate both parties to balance the burden of producing adequate information in such a subject of uncertainty. Through the default burden of proof–proving attribution by a preponderance of evidence–the law of attribution can account for the technological difficulties of proving attribution by allowing the law to recognize when circumstantial evidence can suffice to link an attack to its source. Furthermore, by using the test of virtual control, the law of attribution can more expansively hold states accountable for the non-state actors linked to them, with an affirmative defense of due diligence to create a safe harbor for states that exercise the appropriate level of oversight over such actors. Finally, procedural rules allowing for *ex parte* and *in camera* review of evidence would allow states to accommodate both their concerns about the secrecy of their sensitive intelligence, while also having the capacity to use such relevant evidence in bringing a claim of attribution.

Through such rules, the law of attribution aims to make transparent the source behind cyber-attacks. Cyber-attacks have long been able to go unchecked underneath the veils of secrecy,[273] and states have long been able to elude responsibility for conducting such attacks. While state actors like the United States may have once believed themselves to have a disproportionate advantage in the realm of cyber-warfare,[274] the increasing proliferation of cyber-attacks may have sprawled beyond any single state's control, threatening not only the

[273] *See generally*, FRED KAPLAN, DARK TERRITORY: THE SECRET HISTORY OF CYBER WAR (2016).
[274] *See* Danny Vinik, *America's Secret Arsenal*, Politico (Dec. 9, 2015, 4:57 AM), http://www.politico.com/agenda/story/2015/12/defense-department-cyber-offense-strategy-000331.

security of states, but the stability of their private and civic institutions as well. With the increasing costs of insecurity and uncertainty associated with a world of unfettered cyber-attacks, states may soon come to recognize the need for legal institutions to begin reining them in by holding each other accountable.

Nonetheless, recent years seem to show some tears in the international fabric. With the occurrence of events like Brexit, and the increasing rise of individuals like Trump and Marine Le Pen who endorse protectionist policies,[275] there appears to be a retreat from the international institutions that characterized much of the growth of international law in the past few decades. All of this is compounded by the increasing threat posed by the rise of cyber-attacks, especially its more pernicious uses in potentially interfering with electoral politics and the legitimacy of domestic institutions. All of these threats together would appear to undermine faith in the ability and stability of state sovereignty and international law.

It is easy to get caught in the political winds of the present moment and lose sight of the longer path forward. But the increasing uncertainty today is all the more reminder of the need for further development in international law, not further retreat from it. Imagining the new legal frameworks that we might implement is one step. But the theory of law is only one part of the fight. Theory alone cannot rest on its laurels–the practical concerns and affairs of the world run amok unless such theory can be bent to meet to practical concerns of parties in the world, state and otherwise. The procedural rules set forth by the law of attribution dictate not just the technical features that must be met for a claim to succeed, but the practical costs that accompany them. In doing so, it concretizes the costs of legal institutions to weigh against the costs of

[275] *See The Politics of Anger; Liberalism After Brexit*, ECONOMIST (Jul. 2, 2016), http://www.economist.com/news/leaders/21701478-triumph-brexit-campaign-warning-liberal-international-order-politics.

uncertainty in the ungoverned status quo. It may be that states and their constituents can tolerate a world without law to check the threat of cyber-security. But with a surer sense of what costs the law of attribution may entail, states may soon come to realize that the havoc of unbounded cyber-attacks may in fact be too costly to ignore.