



This is a repository copy of *Cyber attribution: technical and legal approaches and challenges*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/159651/>

Version: Accepted Version

Article:

Tsagourias, N. and Farrell, M. (2020) Cyber attribution: technical and legal approaches and challenges. *European Journal of International Law*, 31 (3). pp. 941-967. ISSN 0938-5428

<https://doi.org/10.1093/ejil/chaa057>

This is a pre-copyedited, author-produced version of an article accepted for publication in *European Journal of International Law* following peer review. The version of record Nicholas Tsagourias, Michael Farrell, *Cyber Attribution: Technical and Legal Approaches and Challenges*, *European Journal of International Law*, 2020, chaa057 is available online at: <https://doi.org/10.1093/ejil/chaa057>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

[forthcoming European Journal of International Law, 2020]

Cyber Attribution: Technical and Legal Approaches and Challenges

Nicholas Tsagourias* & Michael Farrell**

Abstract

Considering the role of attribution in the law of state responsibility, this article examines the technical and international law methodologies and determinants used when attributing malicious cyber activities falling below the use of force threshold to a state and identifies the challenges that arise which lead to responsibility gaps. The article goes on to discuss a number of proposals that aim to improve the effectiveness of the attribution process and in order to close the responsibility gaps. They include institutional proposals envisaging the creation of an international attribution agency; normative proposals advocating the revision of the legal determinants of attribution; and proposals concerning the standard of proof. The aim of the article is to reconstruct the theory and practice of cyber attribution in order to enhance the regulatory potential of international law in this area.

1. Introduction

Modern societies are increasingly dependent on digital technology and infrastructures and, for this reason, malicious cyber operations can cause serious harm to individuals, industry and states. Such harm can be physical, digital, economic, societal, political, psychological, reputational, it can affect security, or be a combination of all these factors. For example, the cost of cybercrime is reported to be between \$799 billion and \$22.5 trillion globally,¹ whereas the US government estimated the cost of malicious cyber activities on its economy to be between \$57 billion and \$109 billion in 2016.² These are dazzling figures, but even small-scale malicious cyber operations can cause serious and long-term harm. For example, tampering with even one voting machine can delegitimize the entire electoral process.

For this reason, individuals, industry and states have an interest in preventing and suppressing malicious cyber operations,³ in mitigating or redressing the harm they cause and, eventually, holding those responsible to account. Attribution is critical in this context because it refers to the process of assigning a particular malicious act to its author: the physical perpetrator but

* Nicholas Tsagourias is Professor of International Law at the University of Sheffield and Director of the Sheffield Centre for International and European Law (nicholas.tsagourias@sheffield.ac.uk).

** Dr Michael Farrell is Co-Executive Director of the Institute for Information Security & Privacy, Georgia Institute of Technology (michael.farrell@iisp.gatech.edu).

¹ P. Dreyer, *et al.*, *Estimating the Global Cost of Cyber Risk: Methodology and Examples*, RAND (2018), available at https://www.rand.org/pubs/research_reports/RR2299.html.

² The White House, The Council of Economic Advisers, *The Cost of Malicious Cyber Activity to the U.S. Economy* (2018), available at <https://www.whitehouse.gov/wp-content/uploads/2018/03/The-Cost-of-Malicious-Cyber-Activity-to-the-U.S.-Economy.pdf>.

³ The paper uses the phrases ‘cyber operation’, ‘cyber-attack’, and ‘cyber incident’ interchangeably to describe malicious acts below the use of force or armed attack thresholds.

even more importantly, the mastermind.⁴ Attribution therefore acts as a catalyst for taking appropriate and effective technical and legal action to prevent and suppress such activities and to establish responsibility. Conversely, non-attribution undermines the process of assigning responsibility and frustrates response action.

Against this backdrop, this article will study the legal methodologies and determinants involved in the process of attributing malicious cyber operations, which fall below the use of force and armed attack thresholds, to states as a prerequisite for engaging their responsibility. The discussion is confined to the law of state responsibility because attribution is one of its constitutive elements and because the function of the law of state responsibility is to maintain international legality by holding states responsible for their wrongful acts.⁵ Before doing this, the article will set out the technical methodologies and determinants of attribution because technical attribution supports and interacts with legal attribution, even if the respective methodologies, determinants and goals differ. To explain, whereas technical attribution is about the forensic investigation of a malicious cyber incident to identify the origins of an attack platform and underwrites technical decisions and actions to patch vulnerabilities and prevent further attacks, legal attribution is about the legal determination of ‘who did it’ on the basis of defined legal criteria in order to ascribe legal responsibility and initiate legal action. Furthermore, although the process of technical attribution can provide many clues about the author of a malicious cyber-attack, it is not sufficient in itself to hold a state legally responsible unless the legal determinants of attribution are also satisfied.⁶ To give an example, technical attribution may identify as the source of a cyber-attack a threat actor related to a state; however, the cyber-attack will not be attributed to that state as a matter of law if the threat actor acted in her private capacity in that instance. Moreover, although legal attribution relies on forensic evidence produced by technical attribution in order to make determinations and justify legal action in the form of, for example, indictments, sanctions, or countermeasures, forensic evidence needs to be interpreted and assessed according to legal criteria. In light of the above, the article will explore the dynamic interaction between the two processes, identify challenges and consider proposals to improve the international law methodology and determinants of

⁴ Tsagourias, ‘Cyber Attacks, Self-Defence and the Problem of Attribution’, 17 (2) *Journal of Conflict & Security Law* (2012) 229, at 233. See also: Lin, ‘Attribution of Malicious Cyber Incidents: From Soup to Nuts’, 70 (1) *Columbia Journal of International Affairs* (2016) 75; Clark and Landau, ‘Untangling Attribution’, *Harvard Law School National Security Journal* (2011), available at https://harvardnsj.org/wp-content/uploads/sites/13/2011/03/Vol.-2_Clark-Landau_Final-Version.pdf.

⁵ International Law Commission (ILC), Articles on Responsibility of States for Internationally Wrongful Acts (2001) Art. 2 (hereinafter referred to as ARSIWA). J. Crawford, *The International Law Commission’s Articles on State Responsibility: Introduction, Text and Commentaries* (2002), at 81-3; *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro)* (hereinafter referred to as Bosnia Genocide), Judgment, 26 February 2007, ICJ Reports (2007) 43, at paras 179, 379, 385. As the EU notes: ‘attribution to a State or a non-State actor ... should be established in accordance with international law of State responsibility’: Council of the European Union, General Secretariat of the Council, ‘Draft Council Conclusions on a Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities (“Cyber Diplomacy Toolbox”)', Doc. No. 9916/17, 7 June 2017, Annex, available at <https://data.consilium.europa.eu/doc/document/ST-9916-2017-INIT/en/pdf>, at para. 4.

⁶ As Novetta stated in its Sony report, it is unable to determine via technical malware analysis whether or not the attack was carried out by an identified nation-state: Novetta, *Operation Blockbuster: Unravelling the Long Thread of the Sony Attack* (2016), available at <https://www.operationblockbuster.com/wp-content/uploads/2016/02/Operation-Blockbuster-Report.pdf>, at 13.

attribution in order to make attribution more effective and close some of the responsibility gaps that currently exist.

This article will proceed as follows. Section two will discuss the current state of attribution of malicious cyber operations to states and identify their trends. Section three will discuss the methodologies and indicia of technical attribution. Section four will consider the legal methodology and determinants of attribution and apply them to certain cyber incidents in order to reveal the responsibility gaps to which they give rise. Section five will discuss the types of evidence – including technical evidence – and the standards of proof used in international law to establish attribution and will seek to reveal the legal uncertainty surrounding these issues. The article will then proceed in section six to discuss proposals for improving the legal methodology and, eventually, the effectiveness of attribution. More specifically, it will discuss institutional proposals envisioning the creation of an international attribution agency but will conclude that, at this point in time, it is neither desirable nor feasible to create such an agency. It will then consider a number of normative proposals with a view to revising the legal determinants of attribution. Specifically, it will discuss looser thresholds of control in the form of ‘overall control’ and ‘soft control’ and will introduce ‘implicit instructions’ as an attribution determinant. In the authors’ opinion, these normative proposals can capture more effectively the dynamics of cyberspace, the prominent role of non-state actors as vectors of malicious cyber operations,⁷ and the multifaceted interactions between non-state actors and states. Thus, they will assist in closing many of the responsibility gaps that currently exist. Finally, the article will propose the ‘preponderance of evidence’ as the most appropriate standard of proof in cases of cyber attribution because it ensures proper scrutiny of the available evidence without unreasonably obstructing attribution determinations. It is hoped that the article’s key findings and proposals will contribute to improving the methodology and practice of attribution in cyberspace and enhance the ability of international law to regulate this area and to hold states responsible.

2. The Current State of Attribution of Malicious Cyber Operations to States

Attribution in cyberspace has traditionally been presented as a challenge because of anonymization, the falsification of identities, the multi-stage nature of cyber operations, the dynamic landscape of cyber threats, the undifferentiated nature of cyber-tools, the human and technical resources involved in performing attribution, and the lengthy timescales involved. State attribution has been even more challenging for the same reasons but also because of the serious political and legal consequences that attribution or misattribution may trigger. However, the initial rarity of state attributions gradually gave way to ever more frequent attribution claims as a result of improvements in attribution capabilities coupled with the changing attitudes of states towards state attribution.⁸ Although we are not going to examine all of these attribution claims in this section, we will highlight certain important features they reveal about the attribution process which will provide some of the context to the discussion that follows.

⁷ In the case of groups, they are referred to as Advanced Persistent Threat (APT) actors.

⁸ See, for example, Council on Foreign Relations, *Cyber Operations Tracker*, available at <https://www.cfr.org/interactive/cyber-operations#Timeline>. According to research by Georgia Tech’s Internet Governance Project, between 2016 and the first quarter of 2018, 85% of the reported incidents were publicly attributed of which 15% were attributions made by governments. See: Mueller *et al.*, ‘Cyber Attribution: Can a New Institution Achieve Transnational Credibility’, 4(1) *The Cyber Defense Review* (2019) 107, at 111-112.

The first feature relates to the actors involved in attribution, which include governments, civil society,⁹ and the private sector, working separately or in collaboration.¹⁰ For example, one of the first and most influential attribution reports by a private company was by Mandiant (now FireEye) which named Unit 61398, a unit within the People's Liberation Army (PLA), as being the host of 'Advanced Persistent Threat 1' (APT1), which was linked to acts of cyber espionage. According to the report, '[t]he issue of attribution has always been a missing link in publicly understanding the landscape of APT cyber espionage' and then the report noted that 'it is time to acknowledge the threat is originating in China'.¹¹ Since then, there has been a substantial increase in attributions of cyber operations to states by private sector companies. Regarding state-to-state attributions, the first took place in 2014 when the US government publicly attributed the Sony attack to North Korea.¹² Since then, however, there have been many more state-to-state attributions including coordinated ones as when the US, UK, Australia, Canada, New Zealand, and Japan attributed WannaCry to North Korea;¹³ the UK, US, Denmark, Australia, Canada, and New Zealand attributed NotPetya to Russia;¹⁴ or when the UK and allies attributed the activities of APT10 involving theft of intellectual property and sensitive data in Europe, Asia and the US to the Chinese Ministry of State Security.¹⁵ There have also been coordinated state and private sector attributions as when US government agencies¹⁶ and the private security company CrowdStrike¹⁷ attributed the hacking of the

⁹ Citizen Lab, *Tracking GhostNet: Investigating a Cyber Espionage Network*, 28 March 2009, available at <https://issuu.com/citizenlab/docs/iwm-ghostnet>.

¹⁰ According to the National Cyber Security Strategy of the United States of America 'The United States will formalize and make routine how we work with like-minded partners to attribute and deter malicious cyber activities...': The White House, *National Cyber Strategy of the United States of America*, (2018), available at <https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf>, at 21.

¹¹ Mandiant, *APT1 Exposing One of China's Cyber Espionage Units*, (2013), available at <https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf>, at 6.

¹² FBI National Press Office, *Update on Sony Investigation*, 19 December 2014, available at <https://www.fbi.gov/news/pressrel/press-releases/update-on-sony-investigation>.

¹³ White House, *Press Briefing on the Attribution of the WannaCry Malware Attack to North Korea*, 19 December 2017, available at <https://www.whitehouse.gov/briefings-statements/press-briefing-on-the-attribution-of-the-wannacry-malware-attack-to-north-korea-121917/>.

¹⁴ The White House, *Statement from the Press Secretary*, 15 February 2018, available at <https://www.whitehouse.gov/briefings-statements/statement-press-secretary-25/>; Foreign & Commonwealth Office, *Foreign Office Minister Condemns Russia for NotPetya Attacks*, 15 February 2018, available at <https://www.gov.uk/government/news/foreign-office-minister-condemns-russia-for-notpetya-attacks>.

¹⁵ National Cyber Security Centre, *Press Release UK and Allies Reveal Global Scale of Chinese Cyber Campaign*, 20 December 2018, available at <https://www.gov.uk/government/news/uk-and-allies-reveal-global-scale-of-chinese-cyber-campaign>.

¹⁶ US Department of Homeland Security & Federal Bureau of Investigation, *Joint Analysis Report: GRIZZLY STEPPE—Russian Malicious Cyber Activity*, 29 December 2016, available at https://www.uscert.gov/sites/default/files/publications/JAR_16-20296A_GRIZZLY%20STEPPE-2016-1229.pdf; Intelligence Community Assessment, 'Assessing Russian Activities and Intentions in Recent US Elections' in Office of the Director of National Intelligence, *Background to 'Assessing Russian Activities and Intentions in Recent US Elections': The Analytic Process and Cyber Incident Attribution*, Doc. No. ICA 2017-01D, 6 January 2017, available at https://www.dni.gov/files/documents/ICA_2017_01.pdf, at 1.

¹⁷ Dmitri Alperovitch, *Bears in the Midst: Intrusion into the Democratic National Committee*, 15 June 2016, available at <https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/>.

Democratic National Committee (DNC) emails during the 2016 Presidential election to Russian state actors.

Secondly, existing attribution reports show that determinations of attribution are multi-sourced and can differ in the amount of information and evidence they contain. Moreover, their degree of analysis and their assessment methodology is often inconsistent or just unarticulated.¹⁸ This has raised questions about reliability and validity.¹⁹

Thirdly, current attribution claims demonstrate that attribution is a multifaceted and interactive process involving different processes each with their own particular determinants and techniques. They also reveal that the majority of existing claims concern technical or political attribution or a combination thereof. In the previous section we set out what technical attribution is, but what should be noted here is that not all technical reports attribute cyber-attacks to states or to threat actors related to states and, even where they do so, the remedies they contain are technical.

In contrast, political attribution is the determination of ‘who did it’ in the form of a state or an entity linked to a state on the basis of political analysis and assessment.²⁰ Political attribution is performed by political institutions as current state-to-state attributions show, it relies heavily on intelligence and is essentially subject to political considerations concerning the question of whether attribution will be done and when, whether it will be public or private, and what will be attributed and to whom; for example to a state or to a state actor linked to a state.²¹ Political attribution may lead to political action such as diplomatic demarches, public denunciations or restrictive measures, without necessarily attaching legal responsibility – this is the aim of legal attribution.

As far as legal attribution is concerned, it is currently performed within domestic law enforcement paradigms and concerns the gathering of evidence to prosecute individuals as, for example, in the case of Park Jin Hyok discussed in the next section²² or in order to impose sanctions on individuals or non-state actors for malicious cyber-attacks.²³ Although domestic

¹⁸ See, for example, ‘*Assessing Russian Activities and Intentions in Recent US Elections*’, *supra* note 16.

¹⁹ With regard to the Sony hack see, for example, Bruce Schneier, *We Still Don’t Know Who Hacked Sony*, 5 January 2015, available at <https://www.theatlantic.com/international/archive/2015/01/we-still-dont-know-who-hacked-sony-north-korea/384198/>. Also see: Jack Goldsmith, *The Sony Hack: Attribution Problems, and the Connection to Domestic Surveillance*, 19 December 2014, available at <https://www.lawfareblog.com/sony-hack-attribution-problems-and-connection-domestic-surveillance>.

²⁰ C. Guittou, *Inside the Enemy’s Computer: Identifying Cyber-Attackers*, (2017).

²¹ According to the UK Advocate-General ‘...the UK can and does attribute malicious cyber activity where we believe it is in our best interests to do so, and in furtherance of our commitment to clarity and stability in cyberspace. Sometimes we do this publicly, and sometimes we do so only to the country concerned. We consider each case on its merits’: United Kingdom Attorney General’s Office, *Cyber and International Law in the 21st Century*, 23 May 2018, available at <https://www.gov.uk/government/speeches/cyber-and-international-law-in-the-21st-century>.

²² See also: *United States v. Viktor Borisovich Netyksho et al.*, Case No. 1:18-cr-00215-ABJ, Indictment, 13 July 2018, available at <https://www.justice.gov/file/1080281/download>; *United States District Court, Southern District of New York, United States of America v. Zhu Hua and Zhang Shilong*, Indictment, Case No. 18 CRIM 891, 17 December 2018, available at <https://www.justice.gov/opa/press-release/file/1121706/download>.

²³ US Department of the Treasury, *Treasury Sanctions Russian Federal Security Service Enablers*, 11 June 2018, available at <https://home.treasury.gov/news/press-releases/sm0410>; US Department of the Treasury, *Treasury Sanctions North Korean State-Sponsored Malicious Cyber Groups*, 13 September 2019, available at <https://home.treasury.gov/index.php/news/press-releases/sm774>. See also: Council

legal processes of attribution may identify the links between said individuals and a state, they do not deal with the issue of state attribution and responsibility under international law for lack of competence. However, even concerned states have not so far invoked the issue of legal attribution and responsibility at the inter-state level. This may be due to political considerations to avoid for example aggravating the situation, but it may also be due to the uncertainty surrounding the scope of international law obligations in cyberspace²⁴ coupled with the difficulties surrounding the application of the legal determinants of attribution alluded to previously. That said, we believe the move to international law is a matter of time not only because the question of how international law applies to cyber-operations is gradually being settled but also because technical or political determinations of attribution, and any ensuing action, may be legally challenged and, more importantly, because attribution determinations need to comply with the legal standards of attribution if a state were to use available international law remedies such as countermeasures. It is for this reason that the article will pre-empt the discussion by examining the legal methodologies and determinants of attribution after discussing in the next section the technical ones.

3. Technical Cyber Attribution



Attributing a cyber-attack begins with a technical analysis of data that results from the attack. A series of actions is required to execute a successful cyber-attack. Analysts use this trail of actions and the related data, along with an established body of knowledge based on previous events that includes the methods and tooling of already known malicious actors, to attempt to trace these operations back to their sources. Although there is no standardized model of cyber-attack, from existing models²⁵ we can say that the cycle of a malicious cyber-attack includes a number of stages: the preparatory stage of target identification, reconnaissance and weaponization; the engagement and presence stage of delivery, exploitation, installation, and actions on objective; and the effects and consequences stage. Analysts collect as many data points as possible from each stage in order to associate them with online personas, individuals, and organizations. Data points that demonstrate potential relevance and/or uniqueness to a forensic investigation are considered indicators. Key indicator categories are tradecraft, infrastructure, malware, and intent as shown in Figure 1. The US Office of the Director of National Intelligence (ODNI) made this matrix publicly available to demonstrate the concepts behind their internal framework for attributing malicious cyber activity.²⁶

Decision (CFSP) 2019/797 of 17 May 2019 concerning restrictive measures against cyber attacks threatening the union or its member states, OJ L129I/13.

²⁴ See Efrony and Shany, 'A Rule Book on The Shelf? Tallinn Manual 2.0 on Cyber Operations and Subsequent State Practice', 112 (4) *American Journal of International Law* (2018) 583.

²⁵ Office of the Director of National Intelligence (ODNI), *Building Blocks of Cyber Intelligence "Cyber Threat Framework"* (2018), available at <https://www.dni.gov/index.php/cyber-threat-framework>.

²⁶ ODNI, *A Guide to Cyber Attribution*, 14 September 2018, available at https://www.dni.gov/files/CTIIC/documents/ODNI_A_Guide_to_Cyber_Attribution.pdf. See also: République Française, Ministère des Armées, *Droit International Appliqué aux Opérations dans le Cyberspace* (2019), at 11.

Data to associate with incident:  Sufficient  Limited

KEY INDICATORS FOR ATTRIBUTION





























CYBER INCIDENT		ADVERSARY	KEY INDICATORS FOR ATTRIBUTION				
			Tradecraft	Infrastructure	Malware	Intent	External Sources
2017	MARCH Major Compromises of Global IT Firms	RUSSIA					
		CHINA*					
		NORTH KOREA					
		IRAN					
		NON-STATE					
	MAY Wannacry Attacks	RUSSIA					
		CHINA					
		NORTH KOREA*					
		IRAN					
		NON-STATE					
	JUNE NotPetya Attacks	RUSSIA*					
		CHINA					
		NORTH KOREA					
		IRAN					
		NON-STATE					
	DECEMBER Saudi Petrochemical Facility Attack	RUSSIA					
		CHINA					
		NORTH KOREA					
		IRAN					
		NON-STATE					

Figure 1. Matrix of competing hypotheses of attribution for five cyber threat actor groups in relation to four major incidents. Graphic courtesy of ODNI.

The first indicator category, tradecraft, refers to the collective behaviour frequently used to conduct cyber-attacks which forms a pattern that can be seen across time and location. This is arguably the most important indicator category because human habits are more difficult to change than technical tools. Examples of tradecraft are payment and financial transactions, email and social media accounts, types of infection and delivery methods (e.g., infected USB drives or compromised websites), or actions on objective (activities inside the target/victim network). However, although an attacker's tools, techniques, and procedures (TTPs) can be unique tradecraft indicators, they can diminish in importance once they become public and other actors can mimic them.

The second indicator category, infrastructure, refers to the physical and/or virtual communication structures used to deliver a cyber capability or maintain command and control (C2) of capabilities. They include, for example, domain names, dynamic DNS services, IP addresses, proxy servers, and anonymity services. Attackers can buy, lease, share, and compromise servers and networks to build their infrastructure. They frequently establish infrastructure using legitimate online services, from free trials of commercial cloud services to social media accounts. Some cyber threat actors are reluctant to abandon infrastructure because of habit, cost, or time, while others will do so because they can rebuild it within hours. This indicator category is very powerful and is relied upon heavily along with tradecraft in many

attribution assessments. Although some attackers routinely change infrastructure between or even within operations to impede detection, skilled analysts actually use this to their advantage by drawing even more nodes and edges in a graph of malicious activity.

The third indicator category, malware, refers to malicious code (aka malware) designed to enable unauthorized functions on a compromised computer system. The functionality of malware is varied and depends on its purpose and target, but example functions include key logging, screen capture, audio recording, remote command and control, and establishing a 'backdoor' to ensure persistent access. One challenge with this indicator category is the increasing ease with which and the number of cyber threat actors that can modify another attacker's malware and repurpose it and the general availability of feature-rich malware on the dark web or cyber black market. Moreover, automation and machine learning systems are capable of making changes to malware and repurposing it quickly, with little human intervention.²⁷ Also, as with infrastructure, sophisticated threat actors routinely change malware between or within operations to impede detection and attribution.

The fourth indicator category, intent, refers to an attacker's commitment to carry out certain actions based on the surrounding context: geopolitical, social, economic, religious, and so on. For example, covert, deniable cyber-attacks are often launched against opponents before or during regional conflicts as in the case of the conflict between Georgia and Russia in 2008.²⁸ While some may consider an indicator in this category to be non-technical, it can provide confidence or weight to an indicator from another category. Intent can also be highly useful in suggesting to an analyst where s/he might look to find other relevant, technical indicators. However, a problem with intent is that multiple attackers might share the same intent, thus making it difficult to discriminate amongst and filter possible cyber threat actors and for this reason information from other sources is required.

In addition to the above, data or evidence from external sources such as the private sector or academia can be used.

Because of the complexity of actions involved in each stage of a cyber-attack, which also require anonymity-enhancing or identity-obfuscating techniques, and because of the time span of actions, mistakes are bound to occur. Forgetting, for example, to turn on a proxy or route through a particular virtual private network (VPN), selecting the wrong stored username or password from an autofill option on a web browser, typing a text string for name or payment information that is from another persona or real identity, or using a proper (real) name in correspondence are all examples of simple and small mistakes that can reveal a true identity and/or permit investigators to make an association with a false identity. Mistakes are therefore critical in attribution determinations because they can reveal patterns and relationships. Even one small error can allow a seasoned analyst with access to large data sets to begin to pull a thread through many disparate sources.

These data can be collected via different mechanisms. Companies that operate Internet infrastructure, run services, or third-party data resellers allow security researchers to purchase data or are compelled by state authorities to provide user, transaction and other data. Community repositories also exist where analysts can access large amounts of 'crowdsourced' data (e.g., VirusTotal²⁹ for malware). In some cases, data can be scraped from public web sites.

²⁷ See: Dustin Frazee, *Cyber Grand Challenge*, for more background on automation in cyber defense, available at <https://www.darpa.mil/program/cyber-grand-challenge>.

²⁸ Sarah White, *Understanding Cyberwarfare: Lessons from the Russia-Georgia War*, 20 March 2018, available at <https://mwi.usma.edu/understanding-cyberwarfare-lessons-russia-georgia-war/>.

²⁹ <https://www.virustotal.com>.

In the case of attribution done by a government, law enforcement agencies and intelligence agencies can also collect data.

In order to yield results, forensic investigations may require time to scour reams of data and assess them. What the ODNI matrix also shows is that analysts weigh the evidence in terms of both volume and veracity to determine a confidence level for their assessments. Certainty is rarely an option, let alone a realistic goal, in cyber-attack attribution as the ODNI matrix shows. Indeed, the ODNI matrix attributes levels of confidence to each indicator category from ‘sufficient’ to ‘limited confidence’. However, there is no published standard as to what constitutes a ‘sufficient’ amount of evidence to support an analytic judgement in making an attribution statement. Anecdotally, many analysts will refrain from making a public assessment when they have limited or no information from one or more of the indicator categories mentioned above and will wait until more evidence is gathered.³⁰ Forensic investigators may also rely on evidence from external sources, for example think tanks, non-profit groups, academics, media, and private industry, in order to confirm a finding or to strengthen the body of supporting evidence for the analytic judgement.

All these issues can be better illustrated by using an example namely, the criminal complaint against Park Jin Hyok³¹ who allegedly was behind a series of cyber-attacks, including the Sony and WannaCry attacks. According to the affidavit, attribution to the defendant was based, in part, on email and social media accounts that were connected to each other and which were used to send spear phishing messages; aliases, malware ‘collector accounts’, used to store stolen credentials; common malware code libraries; proxy services used to mask locations; and North Korean, Chinese, and other IP addresses used across multiple instances of malicious activities. The evidence was collected over many years. It relates to the tradecraft, malware and infrastructure indicator categories and is associated with the preparation and establishing presence and execution stages of the investigated cyber-attacks. The complaint was also based on information from reports produced by private security companies and information from various law enforcement agencies and investigatory agencies following search warrants and formal requests for evidence to foreign countries.³² Mistakes made by Park Jin Hyok were also critical in revealing his real credentials. The affidavit finally concluded that on the basis of the evidence there was sufficient probable cause for the requested complaint.³³

So far we have looked at the evidence and standards used in technical attribution but it should be noted that there is no single widely-accepted attribution model (although there have been a number of proposals in academic literature).³⁴ Cyber-attacks are dynamic in nature and models

³⁰ See, for example, the ODN matrix with regard to NotPetya.

³¹ United States District Court for the Central District of California, *United States of America vs Park Jin Hyok*, Unsealed Criminal Complaint, Case No. MJ-18-1479, 6 September 2018, available at <https://www.justice.gov/opa/press-release/file/1092091/download>. It should be acknowledged though that the case does not concern state-to-state attribution; it is referenced because of the detailed information it contains.

³² *Ibid.*, at paras 3, 4.

³³ *Ibid.*, at para 5. It should be noted, however, that this is not the standard used in a criminal trial.

³⁴ Wheeler and Larsen, ‘Techniques for Cyber Attack Attribution’, *Institute for Defense Analyses*, Paper P-3792, (2003), available at <https://apps.dtic.mil/dtic/tr/fulltext/u2/a468859.pdf>; Hunker, Hutchinson, and Marguiles, ‘Role and Challenges for Sufficient Cyber-Attack Attribution’, *Institute for Information Infrastructure Protection* (2008), available at <http://cobweb.dartmouth.edu/~thei3p/>; Lin, ‘Attribution of Malicious Cyber Incidents’, *Hoover Institution*, Aegis Series Paper No. 1607 (2016), available at https://www.hoover.org/sites/default/files/research/docs/lin_webready.pdf; Caltagirone, Pendergast, and Betz, ‘The Diamond Model of Intrusion Analysis’, *Center for Cyber Threat Intelligence and Threat Research*, Technical Report ADA586960 (2013), available at: <http://www.activeresponse.org/wp->

quickly breakdown or fail to account for new developments in the defensive and threat landscape. New protections are deployed by software and hardware vendors, while at the same time new weaknesses are discovered and targeted by malicious cyber actors. While some linear approaches have been used to illustrate the set of activities generally required for a successful cyber-attack, in reality the process is actually quite non-linear.

In concluding this section, it has been shown that technical attribution can be based on different categories of technical indicators and the extent to which multiple data points in each category are available. It has also been shown that methodological questions remain open and the level of confidence in the evidence varies; there is always a degree of granularity in technical attribution. Moreover, and as was said in the introduction, the aims and priorities of technical attribution differ from legal attribution. For this reason, technical attribution does not automatically translate into legal attribution and is not sufficient in itself to hold a state legally responsible unless the technically attributed malicious cyber-attack can also be attributed to a state as a matter of law and the technically produced evidence can also be validated in law. For this reason, the next section will examine the international law methodology and determinants of attribution.

4. Attribution in the Law of State Responsibility and its Determinants

Attribution in the law of state responsibility is a normative - not a factual or technical – process whose function is to assign a wrongful act to a state in order to engage its responsibility.³⁵ For this reason, its determinants are moulded by how the state is defined which in the law of state responsibility is reduced to the structures, entities and functions that make up its legal-political order. This makes the legal determinants of attribution quite narrow, requiring an identifiable, direct and close link between a state and an entity or between a state and the impugned conduct; a link that overrides the latter's independent existence. This occurs when an institutional, functional or agency link between a state and an entity or conduct is established.

The institutional link covers the relationship between a state and its *de jure* or *de facto* organs.³⁶ *De jure* organs are entities that are defined as such by the state's law. This would be the case, for example, with the GRU officers indicted in relation to the DNC hacking.³⁷ It will also cover entities, cyber defence groups or hacker groups incorporated into the state apparatus such as the Estonian Defence League;³⁸ Unit 61398 of the Third Department of the Chinese People's

[content/uploads/2013/07/diamond.pdf](https://www.odni.gov/uploads/2013/07/diamond.pdf); Santiago, Report Prepared for ODNI Public-Private Sector Analyst Engagement Program by the Office of Intelligence & Analysis, US Department of Homeland Security, *The Confidence Pyramid Conceptual Model* (2017), available at https://www.odni.gov/files/PE/Documents/PHASE-II_CYBER-ATTRIBUTION.pdf.

³⁵ Art. 2, ARSIWA. Attribution in the law of state responsibility is not about 'assigning responsibility for malicious cyber activity to a specific actor or sponsor' as the EU claims in a recent non-paper: Council of the European Union, European External Action Service, 'Implementation of the Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities – Attribution of Malicious Cyber Activities – Discussion of a Revised Text', Doc. No. 6852/1/19 REV 1, 18 March 2019, Annex, available at <http://www.statewatch.org/news/2019/mar/eu-council-cyber-6852-REV-1-19.pdf>, at 2.

³⁶ Art. 4 ARSIWA; Bosnia Genocide, *supra* note 5, at para 385.

³⁷ *United States v. Viktor Borisovich Netyksho et al.*, *supra* note 22. GRU is the Russian military intelligence agency and those indicted were members of Unit 26165 and Unit 74455.

³⁸ Estonian Defence League Act 2013, available at <https://www.riigiteataja.ee/en/eli/525112013006/consolide>.

Liberation Army;³⁹ Israel's Unit 8200 or Bureau 121, a hacking unit within the North Korean Reconnaissance General Bureau (RGB).⁴⁰

De facto organs are state instrumentalities. They include entities, groups or individuals who are completely dependent on a state and over whom the state exercises control 'in all fields'.⁴¹ From an examination of existing jurisprudence, it transpires that if an entity has been created by a state, operates on behalf of that state and has no real autonomy in decision-making, it is a *de facto* organ of that state. Contracted-out cyber groups or companies may also be included, but that would depend on the terms of the contract and how it is executed, in particular, whether the outsourced tasks are closely linked to the state such as security tasks, and whether their delivery is controlled by the state. That said, proving such a close relationship is particularly difficult and, as the ICJ said, it will be quite 'exceptional' to qualify entities as *de facto* organs.⁴² For example, the US government characterized three North Korean hacker groups including the 'Lazarus Group' as 'agencies, instrumentalities, or controlled entities of the Government of North Korea'⁴³ but it provided no evidence to substantiate the claim.

Moving on to the second modality of attribution, a functional link is established when an entity is empowered by a state to exercise governmental authority.⁴⁴ The delegation of authority can be specific or general depending on how it is stipulated in domestic law. For example, if the security of a governmental network is contracted-out to a private company, the contract between the government and the private company would amount to delegation if executive delegation is permitted by domestic law, otherwise, what would amount to delegation is the general authorizing legislation. In relation to this, it should be noted that judicial authorizations, for example, authorizations to conduct data searches, do not constitute delegation; they only certify the lawfulness of the search. With regard to the second prong of the test, what constitutes a governmental function varies, with the exception perhaps of certain intrinsically governmental functions such as defence. It all depends on the nature and purpose of the activity, the overall context within which such functions are exercised and the state's political identity. For instance, if a private cyber security company is authorized to defend the state against cyber-attacks, that would be a governmental function, but not when it is authorized to defend its own property from cyber intrusions. Would, however, ransomware attacks or cyber-crime constitute governmental functions if they support a state's economy or if they are part of a state's governance tools? Is espionage a governmental function?⁴⁵ With regard to espionage, one perhaps needs to distinguish political from industrial espionage, but even industrial espionage can be linked to core governmental functions such as the economy or national security. That said, and notwithstanding how such activities are characterized, it is highly unlikely that a state will explicitly authorize an entity to perform such activities; rendering Article 5, ARSIWA inapplicable. Equally, Article 5, ARSIWA, will not apply to attributions performed by private companies because there is no authorization, even if attribution is deemed to be a sovereign prerogative or touches upon a state's foreign relations. Article 5, ARSIWA may, however, play

³⁹ See, Mandiant, *supra* note 11.

⁴⁰ See Matthew Ha and David Maxwell, FDD Report, *Kim Jon Un's 'All Purpose Sword'*, 3 October 2018, available at <https://www.fdd.org/analysis/2018/10/03/kim-jong-uns-all-purpose-sword/>.

⁴¹ *Case Concerning Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)*, Judgment, 27 June 1986, ICJ Reports (1986) 14, at para 109 (hereinafter referred to as Nicaragua Case); Bosnia Genocide, *supra* note 5, at paras 390-394.

⁴² Bosnia Genocide, *supra* note 5, at para 393.

⁴³ US Department of the Treasury, *Treasury Sanctions North Korean State-Sponsored Malicious Cyber Groups*, *supra* note 23, <https://home.treasury.gov/index.php/news/press-releases/sm774>.

⁴⁴ Arts 5 and 6, ARSIWA.

⁴⁵ See R. Buchan, *Cyber Espionage and International Law* (2018), at 21-24.

a more prominent role in attributing cyber conduct if implicit authorization is accepted that is, authorization derived from informal relations and practices. However, as was said, Article 5, ARSIWA requires formal authorization, which reduces its operational functionality.

Although the narrow definition of what constitutes a state organ or a state empowered entity limits the scope of attribution *rationae personae*, attribution is expanded *rationae materiae* in these two instances because all their acts, including their *ultra vires* acts, are attributed to a state, provided that they were carried out under the cloak of state authority and they are not so far removed from official functions to be equated with private conduct.⁴⁶ There are many difficulties, however, in establishing when an entity acts with real or apparent authority in cyberspace, when s/he acts whilst on duty or off-duty, or in establishing which acts fall within official functions and which are private. For example, how can someone impersonating a private person give the impression that s/he operates with apparent authority? Also, is spear phishing or releasing documents stolen from private accounts or selling data acquired through espionage part of official functions? These are difficult questions to answer and, to the extent that these were some of the activities carried out by the indicted GRU officers during the 2016 Presidential election, Russia can plausibly claim that its organs acted in their private capacity. The difficulties in distinguishing private from apparent official conduct can, however, be circumvented if the unauthorized conduct is systematic and recurrent, ‘such that the state knew or ought to have known of it’.⁴⁷ In this case, said conduct will be attributed to the state concerned as if it were implicitly authorized by that state. This is a case of constructive attribution. Consequently, and to the extent that the activities of Russian state organs or agencies during the 2016 US Presidential election were as systematic and widespread as the Mueller indictment revealed,⁴⁸ they can be attributed to Russia, even if certain conduct was actually private in nature.

Finally, an agency link is established when a state instructs or directs a person to commit a wrongful act or when the state exercises control over the wrongful act.⁴⁹ Instructions imply orders which should be given in relation to each specific act that constitutes a violation of international law⁵⁰ and should be carried out as such by those instructed. Direction means guidance over the entity that commits the wrongful act in the sense of the state taking the lead. In both cases, the perpetrator implements or follows the state’s decision to commit the particular wrongful act and her will is subordinated to the will of the state. Regarding the criterion of control, it should be exercised over the specific conduct in question or over the operation in the course of which unlawful acts are committed and, as the ICJ has repeatedly said, it should be ‘effective’. Although the Court never defined ‘effective’, it is deemed to amount to domination over the act.⁵¹

Applying Article 8, ARSIWA to current attribution claims reveals the difficulties in establishing state attribution. Granted, the reports do not use legal descriptors to qualify the relationship between a state and a cyber group whereas the language they use varies and sometimes it is inconsistent, however it is important to see whether they can support legal attribution. The 2013 Mandiant Report for example says about APT1 that it is ‘government-

⁴⁶ Art 7, ARSIWA. *Ultra vires* acts are those acts that exceed an organ’s authority or contravene instructions.

⁴⁷ J. Crawford, *supra* note 5, at 108, para 8.

⁴⁸ See *supra* note 31, at para 1.

⁴⁹ Article 8, ARSIWA; J. Crawford, *supra* note 5, at 110-113.

⁵⁰ Bosnia Genocide, *supra* note 5, at para 400.

⁵¹ Nicaragua Case, *supra* note 41, at paras 116-117; Bosnia Genocide, *supra* note 5, at paras 398, 402-406, 413-414.

sponsored’ whereas in other parts it says that it ‘receives direct government support’, or that it acted with the ‘full knowledge and cooperation of the Chinese government’.⁵² Interpreting such language against the descriptors of Article 8, ARSIWA places ATP1 and its activities outside its scope. At this junction it should be noted that many reports use the term ‘government-sponsored’, ‘state-sponsored’ or similar language to describe the relationship between a state and a group or a cyber-attack but, although these terms are very common in political parlance, they are devoid of legal meaning and, in any case, they allude to a broader and indeed looser relationship between a state and an entity or an act than what Article 8 ARSIWA requires. To give other examples, it was said previously that the US government opined that the ‘Lazarus Group’ is controlled by North Korea, but this is not equivalent to ‘effective control’ whereas the ODNI report concluded that President Putin ‘ordered the campaign to influence the US elections’.⁵³ In relation to this, even if there were indeed instructions, they were too general and did not amount to a request to commit unlawful acts in order to lead to attribution according to Article 8, ARSIWA.

In addition to the above attribution determinants, Article 11, ARSIWA introduces another determinant by attributing to a state the acts of private actors when that state acknowledges and adopts them as its own. The adoption, however, needs to come from the highest levels of government and needs to be clear and explicit.⁵⁴ This would mean that the Stuxnet attack which undisclosed senior US officials acknowledged in a newspaper article⁵⁵ cannot be attributed to the US because the acknowledgement was not explicit and clear, there was no adoption and, more importantly, it did not concern an act committed by a third party which is the gist of Article 11 attribution.

The preceding discussion has thus demonstrated that existing attribution claims do not satisfy the attribution determinants found in the law of state responsibility but also that these determinants cannot be fulfilled easily in cyberspace (with the exception, perhaps, of *de jure* organs) which leads to responsibility gaps. These difficulties are compounded further by evidentiary difficulties. As the Russian presidential spokesman, Dmitry Peskov, said with regard to the accusation that Russia was responsible for the DNC hack, the United States ‘should either stop talking about [Russia being responsible for the DNC hack] or produce some proof at last’⁵⁶ whereas China reacted to the US accusations that it was responsible for the intrusion into the Office of Personnel Management by saying that they were neither ‘responsible nor scientific’ and stressed that it is ‘imperative to stop groundless accusations’.⁵⁷ For this reason, we shall discuss evidentiary issues associated with cyber attribution.

5. Evidence and the Standard of Proof in Cyber Attribution

⁵² Mandiant, *supra* note 11, at 2 and 59.

⁵³ ‘Assessing Russian Activities and Intentions in Recent US Elections’, *supra* note 16, at 1.

⁵⁴ *United States Diplomatic and Consular Staff in Tehran (United States of America v Iran)*, Judgment, 24 May 1980, ICJ Reports (1980) 3, at paras 63-74.

⁵⁵ David E Sanger, *Obama Order Sped Up Wave of Cyberattacks against Iran*, 1 June 2012, available at https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?_r=1&hp&pagewanted=all.

⁵⁶ Laura Smith-Spark, *Russia Challenges US to Prove Campaign Hacking Claims or Shut Up*, 16 December 2016, available at <http://edition.cnn.com/2016/12/16/europe/russia-us-hacking-claims-peskov/index.html>.

⁵⁷ Finklea *et al.*, *Cyber Intrusion into U.S. Office of Personnel Management: In Brief, Congressional Research Service*, 17 July 2015, available at https://digital.library.unt.edu/ark:/67531/metadc743551/m1/1/high_res_d/R44111_2015Jul17.pdf, at 3.

As was noted in section two, existing attribution reports are quite thin on the evidence they contain, and they are often quite obscure in their assessment methodology. Moreover, certain states such as the US, UK and France claim that they are under no obligation to disclose the evidence upon which attribution is made.⁵⁸ However, employing a methodologically sound process of analysing and assessing evidence, and a cogent standard of proof are critical for the credibility and validation of attribution because evidence can substantiate attribution determinations and justify subsequent actions. This is the same in legal as well as non-legal - technical or political - processes of attribution. For this reason, the 2015 UN GGE Report states that ‘accusations of organizing and implementing wrongful acts brought against States should be substantiated’.⁵⁹ It is, therefore, important to examine how cyber attribution can be substantiated as a matter of law because, as was said, states may be called upon to justify their determinations and the lawfulness of any action, for instance countermeasures, they may take. For this reason, in this section, we will discuss two issues: the types of evidence and the standard of proof required to establish cyber attribution according to international law. The discussion will be informed by the relevant international legal jurisprudence which, notwithstanding its relative under-development and its heavy judicial focus, provides the minimum context for understanding how evidence can be legally used and assessed. Moreover, it will provide insights that can be used beyond the legal context because, as the ODNI document presented in section three shows, there are parallels between legal and non-legal standards on evidence.⁶⁰

With regard to the first issue, the type of evidence that will be taken into account when making attribution determinations, it goes without saying that data and digital evidence obtained through forensic investigations constitute the primary category of evidence to be used in legal determinations of attribution. As indicated in section two, they may include, amongst other things, malware samples, compiler language, programming language, IP addresses, domain names, registration information for infrastructure, payment information for infrastructure, patterns/ordering of execution events, keyboard layout for malware creation, scripts and programs used on a victim network or host. What can also be used is documentary evidence such as cyber strategies, legislation, attribution reports, directives or intelligence reports (even in redacted form).

However, whether these items will be treated as legally significant depends on their relevance and probative value. The ICJ, for instance, takes into consideration a number of factors in this regard such as the source of the evidence and in particular its independence; the disinterested character of the investigation; whether the evidence is first-hand and contemporaneous or secondary and subsequent to the event; whether the assessment methodology was sound; or

⁵⁸ Egan, ‘International Law and Stability in Cyberspace’, 35 *Berkeley Journal of International Law* (2017) 169, at 177; *Cyber and International Law in the 21st Century supra* note 21; *Droit International Appliqué aux Opérations dans le Cyberspace, supra* note 26.

⁵⁹ UNGA, Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, UN Doc. No. A/70/174, 22 July 2015, at para 28(f).

⁶⁰ ‘Analysts evaluate three components when assigning probabilistic language and confidence levels: the timeliness and reliability of the evidence, the strength of the logic linking the evidence, and the type of evidence (direct, indirect, circumstantial, or contextual) It also explains that high confidence exists when the evidence is ‘beyond reasonable doubt with no reasonable alternative’; moderate confidence when the evidence is ‘clear and convincing’, and low confidence when ‘more than half of the body of evidence points to one thing, but there are significant information gaps’ ODNI, *A Guide to Cyber Attribution supra* note 26, at 4.

whether the evidence has been cross-examined or corroborated.⁶¹ This would mean that most of the existing attribution reports by state agencies will fail these tests for lack of independence, impartiality and external scrutiny, for the meagre amount of evidence they contain and for failing to explain their assessment methodology. Likewise, attribution reports by private security companies will most probably fail these tests. Their legal significance in proving state attribution can also be affected by whether they have been endorsed by the concerned government.⁶² As it transpires from existing government reports or statements, they often mention reports by the private sector but in a legally non-committal manner; whereas, on other occasions, they include information or evidence contained in such reports without identifying the provider or the relevant evidence.⁶³ In our opinion, none of this amounts to endorsement.

In relation to digital evidence, it may be easier to collect if it is open sourced but what can affect its legal probity is the fact that its probity depends on verification and authentication. This is because, for security or other reasons, states may not want to disclose their attribution technology, or the sources or personnel used to collect and analyse the evidence. Furthermore, the attribution technology used by a state may not be widely available in which case evidence cannot be verified.

In addition, there may be jurisdictional difficulties with the collection of hard evidence because it requires cross-border investigations. As indicated in section two, malware and scripts may be collected from within victim systems residing in different jurisdictions; whereas, other items, such as registration information, can be found in databases belonging to third parties that operate on the Internet, and which may reside in different jurisdictions. However, the state concerned may not consent to or cooperate with such investigations. This gives rise to three further questions: first, whether illegally obtained evidence can be used; secondly, whether adverse inferences can be made when the state refuses to cooperate; and, thirdly, what evidence, other than hard evidence, can be used in cases where obtaining hard evidence proves to be difficult. With regard to the first question, the use of illegally obtained evidence is not usually permitted in domestic legal proceedings but in international law things are perhaps different in light of the fact that the ICJ has not rejected such evidence but assesses its relevance and reliability in context. As to whether adverse inferences can be drawn if a state refuses to cooperate with the production of evidence, there is some support for this view,⁶⁴ but the ICJ seems to take a more cautious approach noting that it can draw its own conclusions from the non-production of evidence.⁶⁵ With regard to the third question, the use of circumstantial

⁶¹ *Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda)* (hereinafter *Armed Activities Case*), Judgment, 19 December 2005, ICJ Reports (2005) 168, at para 59; *Pulp Mills on the River Uruguay (Argentina v Uruguay)*, Judgment, 20 April 2010, ICJ Reports (2010) 14, at 72, para 168; *Bosnia Genocide*, *supra* note 5, at para 213. A. Riddell and B. Plant, *Evidence before the International Court of Justice* (2009) at 192.

⁶² *Case Concerning Application of the International Convention on the Elimination of all Forms of Racial Discrimination (Georgia v. Russian Federation)*, Preliminary Objections, 1 April 2011, ICJ Reports (2011) 70, at paras 73, 81.

⁶³ See *GRIZZLY STEPPE*, *supra* note 16, at 1; *Unsealed Criminal Complaint*, *supra* note 31; *Press Briefing on the Attribution of the WannaCry Malware Attack to North Korea*, *supra* note 13.

⁶⁴ R. A. Clarke and R. K. Knake, *Cyber War: The Next Threat to National Security and What to Do About It* (2010), at 178.

⁶⁵ *Bosnia Genocide*, *supra* note 5, at paras 44 and 204-6, also see: *Dissenting Opinion of Vice-President Al-Khasawhen*, *ibid*, at para 35 and *Dissenting Opinion of Judge ad hoc Mahiou*, *ibid*, at paras 56-63.

evidence becomes critical as the 2010 UN GGE Report⁶⁶ confirms. Circumstantial evidence is relational evidence; it includes surrounding factors and circumstances to prove a certain fact. They may include geopolitical factors, indicators of origin, motivation, the degree of sophistication of the attack, the scale and timing of the attack, linguistic indicators, common tooling, infrastructure and so on. In fact, similar evidence is also used to establish technical attribution as discussed in section two. Yet, any inferences made on the basis of circumstantial evidence should be reasonable in light of existing primary evidence.⁶⁷ It follows from this that without hard evidence, geopolitical factors, however persuasive, are not sufficient to attribute the Stuxnet attack to the US or Israel, and the same can be said about language indicators or IP addresses because they may be false flags.

The second issue to discuss is what standard of proof can clothe attribution determinations with the necessary degree of confidence. As was said in section two, existing attribution determinations do not reveal any consistently applied standard of proof or they fail to articulate any particular standard.⁶⁸ Likewise, the ODNI matrix suggests that technical attribution should be based on sufficient evidence, but it does not explain its threshold. That notwithstanding, existing reports indicate that attribution determinations should enjoy a degree of certainty in order to be credible and reliable. This is exactly what the standard of proof does. However, the law of state responsibility does not set its own evidentiary standards and international law in general is quite undecided, adapting the standard of proof to the facts of the case and the norms involved.⁶⁹ The ICJ, for example, opined that for charges of exceptional gravity, evidence needs to be fully conclusive and that this standard also applies to attribution,⁷⁰ but it was criticized for introducing criminal law standards to civil law cases.⁷¹ Elsewhere, the Court used formulations such as ‘sufficient’ or ‘conclusive’ evidence;⁷² or variations of the ‘on a balance of probabilities’, ‘in all probability’, ‘consistent with the probabilities’, ‘with a high degree of

⁶⁶ UNGA, Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (GGE Report), UN Doc. No. A/65/201, 30 July 2010, p.2.

⁶⁷ Nicaragua Case, *supra* note 41, at para 111; Bosnia Genocide *supra* note 5, Dissenting Opinion of Vice-President Al-Khasawhen, *ibid*, at para 51.

⁶⁸ For example, the UK government stated that ‘it is highly likely that North Korean actors known as the Lazarus Group were behind the WannaCry ransomware’: Foreign & Commonwealth Office, *Foreign Office Minister Condemns North Korean Actor for WannaCry Attacks*, 19 December 2017, available at <https://www.gov.uk/government/news/foreign-office-minister-condemns-north-korean-actor-for-wannacry-attacks>. The FBI attributed the Sony attack to North Korea on the basis of ‘enough information’, Update on Sony Investigation, *supra* note 12. Most reports, however, do not state any standard or if there is such a standard it remains unarticulated. See reports on DNC hacking, *supra* note 16.

⁶⁹ J. Crawford, *supra* note 5, at 124, para 4 but also see p.93, para 9. Regarding the ICJ, it applies various standards as subsequent footnotes will show.

⁷⁰ Bosnia Genocide, *supra* note 5, at para 209.

⁷¹ Meron, ‘Major Developments in International Law: A Conversation on the ICJ’s Opinion in Bosnia and Herzegovina v. Serbia and Montenegro’, 101 *American Society of International Law Proceedings* (2007) 215 at 216.

⁷² Nicaragua Case, *supra* note 41, at para 110; Armed Activities Case, *supra* note 61, at paras 91 and 172; Bosnia Genocide, *supra* note 5, at para 209.

probability'.⁷³ As far as the Tallinn Manual is concerned, it relegates this issue to judicial processes and admonishes states to act reasonably in the circumstances.⁷⁴

What the preceding discussion thus demonstrates is that although international law takes a more relaxed approach to evidence which may, in principle, assist cyber attribution, the absence of well-articulated standards of proof and their mutability create uncertainty which may affect the credibility, reliability and also the validity of legal determinations.

6. Proposals for Improving the Attribution Process

In the light of the challenges surrounding cyber attribution, we will now consider a number of proposals and whether they can make cyber attribution and its methodology more efficient and effective. The first proposal is of an institutional nature and envisages the establishment of an international attribution agency to centralize and streamline attribution determinations; the second proposal is normative and envisages the revision of the attribution determinants to capture the reality of cyber-threat actors and their connections to states; whereas the third proposal seeks to identify a standard of proof that can facilitate the making of reliable and credible attribution determinations.

A. International Attribution Agency

There are many blueprints for such an agency ranging from an agency with purely private-sector membership, an agency with private-public membership and a purely intergovernmental agency.⁷⁵ For its proponents, an international agency that centralizes and streamlines attribution determinations and processes will engender trust in attribution against the current state of decentralized, often inconsistent and methodologically obscure determinations. It will also lead to the standardization of attribution by establishing its own attribution methodology, rules of evidence, and decision-making process. Moreover, such an agency can assist governments or courts in making more informed and persuasive decisions. Other advantages are that it will contribute to the depoliticization of attribution but also to its democratization in view of the inequities in states' cyber capabilities. However, whether such an agency can attain these aims depends on its independence; the nature of its membership and its representative character; its financial transparency; its decision-making process and standards; its investigatory

⁷³ *Case Concerning the Land, Island and Maritime Frontier Dispute (El Salvador/Honduras: Nicaragua Intervening)*, Judgment, 21 September 1992, ICJ Reports (1992) 351, at paras 121, 155, 248; *Case Concerning Sovereignty over Pulau Ligitan and Pulau Sipadan (Indonesia/Malaysia)*, Judgment, 17 December 2002, ICJ Reports (2002) 625, at para 72; Nicaragua Case, *supra* note 41, at para 158.

⁷⁴ M. N. Schmitt, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, (2nd ed., 2017), at 81-2.

⁷⁵ J. S. Davis *et al*, RAND, *Stateless Attribution: Toward International Accountability in Cyberspace* (2017); Microsoft, *From Articulation to Implementation: Enabling Progress on Cyber Norms* (2016) available at <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/REVmc8>; Atlantic Council, *Confidence-Building Measures in Cyberspace: A Multistakeholder Approach for Stability and Security* (2014), available at <http://www.atlanticcouncil.org/images/publications/Confidence-Building-Measures-in-Cyberspace.pdf>; Microsoft, *An Attribution Organization to Strengthen Trust Online*, available at <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW67QI>; Droz and Stauffacher, ICT4Peace Foundation, *Trust and Attribution in Cyberspace: A Proposal for an Independent Network of Organisations Engaging in Attribution Peer-Review* (2018), available at <https://ict4peace.org/wp-content/uploads/2019/07/ICT4Peace-2019-Trust-and-Attribution-in-Cyberspace.pdf>. Elena Chernenko, Oleg Demidov and Fyodor Lukyanov, Council on Foreign Relations, *Increasing International Cooperation in Cybersecurity and Adapting Cyber Norms*, 23 February 2018, available at <https://www.cfr.org/report/increasing-international-cooperation-cybersecurity-and-adapting-cyber-norms>.

competence; its technical, political and legal expertise; its attribution methodology; the standard of proof it applies, and the vigorousness of its oversight mechanisms.⁷⁶

In our opinion, the above represent a rather demanding set of conditions and it is doubtful whether they can ever be satisfied. There are also other factors that cast doubt on the desirability as well as the feasibility of such an agency. The first and perhaps most important factor is that the functioning, utility, authority and effectiveness of such an agency will ultimately depend on the willingness of states to cooperate and accept its findings. Yet, many states view attribution as part of their sovereign prerogative, and they do not feel obligated to provide information when they make attributions. This is the US, UK and French position and also the position of the European Union and NATO.⁷⁷ Secondly, even if states exhibit some cooperative spirit in this regard, they will still want to protect their networks from prying eyes, something that will hinder the work of such an agency. Furthermore, victim or perpetrator states may demand access to the information the agency used in order to make its determinations which the agency may not be able to provide if such information is classified. Thirdly, there is the danger that such an agency may substitute its own attribution determination for that of courts or governments, even if its avowed function is to facilitate attribution determinations than to make conclusive determinations. To explain, the agency's determinations may have a direct bearing on legal determinations of attribution because of the close link between facts and the law in this area, particularly, if courts or other legal institutions lack the resources or expertise to scrutinise the agency's findings.⁷⁸ Likewise, states that lack the required technological or intelligence capacity may not be able to verify the agency's determinations. Even more worryingly, states may be compelled to take action following the agency's attribution determination against their more prudent judgement. If this were to happen, it would raise questions about the legitimacy of the agency, the legality of its determinations and its accountability. Fourthly, the timeline within which such an agency will produce its findings may affect the application of international law. In international law, there needs to be some form of temporal proximity between action and reaction as in the case of self-defence or countermeasures. Likewise, judicial proceedings should be completed within a reasonable time. Long timelines in making attribution determinations will either frustrate prospective legal action or delay it to such an extent that the temporal link would break down and the action would lose its legal justification. For these reasons, states or courts may ignore the agency and rely on their own attribution determinations unless, of course, the temporal requirements inscribed in law are also revised.

In light of the above, we believe that the creation of such an agency is quite premature and will just add another layer in the already fractured attribution process.

⁷⁶ See Davis *et al.*, *Stateless Attribution*, *ibid.*, at 25-34.

⁷⁷ 'Draft Council Conclusions on a Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities ("Cyber Diplomacy Toolbox") *supra* note 5, as well as 'Implementation of the Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities – Attribution of Malicious Cyber Activities – Discussion of a Revised Text', *supra* note 35 at 13: 'Every Member State is free to make its own determination with respect to attribution of a malicious cyber activity'; NATO, *Brussels Summit Declaration Issued by the Heads of State and Government Participating in the Meeting of the North Atlantic Council in Brussels 11-12 July 2018*, 11 July 2018, available at https://www.nato.int/cps/en/natohq/official_texts_156624.htm?selectedLocale=uk, at para 20. For the US, UK and French position see *supra*, note 56.

⁷⁸ See Pulp Mills, *supra* note 61, Dissenting Opinion of Judge ad hoc Vinuesa, at para 71 and Joint Dissenting Opinion of Judges Al-Khasawneh and Simma, at para 4.

B. Revision of the Attribution Determinants

The second proposal concerns the revision and readjustment of the attribution determinants focusing mainly on Article 8, ARSIWA. It should be said, in this respect, that revising the attribution determinants is neither unreasonable nor against the law and we will explain why. First, regarding the criterion of ‘control’, Article 8 ARSIWA does not specify its scope; whereas, the commentary notes that this standard should apply with a degree of flexibility. This means that there is room for readjustment. Secondly, the emergence of special responsibility and, for this reason, the emergence of special attribution regimes has been recognised by the ILC whereas, the ICJ admitted that the attribution determinants may vary depending on the area.⁷⁹ Thirdly, it should be recalled that the determinants of Article 8 ARSIWA were developed in an era where states created armed groups to fight proxy wars and these groups were dependent on states for direction and resources, such as weapons. The Article 8 ARSIWA descriptors thus reflect such unequal and vertical relations between states and non-state actors. However, the circumstances where Article 8 ARSIWA is called upon to be applied have changed and more so in cyberspace. Cyberspace offers a particularly facilitative environment for non-state actors to emerge and operate and many non-state actors in cyberspace are often self-sufficient, autonomous and pursue their goals either independently from states or in alignment with states but such alignment can be, variously, formal or informal, vertical or horizontal, continuous or *ad hoc*, attached or detached.⁸⁰ In this context, states may just offer political or ideological direction and support to non-state actors rather than material support. Fourthly, the current attribution determinants are conceptually informed by the distinction between the public and the private domain.⁸¹ However, the dividing line between private and public has nowadays been displaced and the relationship between non-state actors and states often resembles a networked relationship where non-state actors complement state policies and goals materially, functionally, or ideologically. This is particularly true in cyberspace. Fifthly, the facilitative environment cyberspace offers and the networked relationship between states and non-state actors can be exploited by states to achieve their policy goals by maintaining a certain level of plausible deniability and thus evade their responsibility. Consequently, if attaching responsibility to non-state actors as independent legal persons⁸² or attaching responsibility to states without attribution⁸³ is not currently on the international law agenda, and if international law is to maintain its relevance as a governance tool in cyberspace, it is necessary to start thinking about how to approach the attribution determinants in order to capture the role and place of non-state actors in cyberspace and the more subtle modes of interaction between states and non-state actors therein and in order to close the responsibility gaps that currently exist.

Amplifying the level of control a state needs to exercise over organized non-state actors in cyberspace, and moving away from ‘effective control’, can assist significantly in achieving these aims. The ‘overall control’ standard introduced by the ICTY is particularly apposite in this regard. ‘Overall control’ is established when a state equips, finances, coordinates or helps in the general planning of the activities of an organized group, short of issuing specific

⁷⁹ Art. 55, ARSIWA; Crawford, *supra* note 5, at 110-112; Bosnia Genocide, *supra* note 5, at para 405.

⁸⁰ T. Maurer, *Cyber Mercenaries* (2018).

⁸¹ Crawford, *supra* note 5, at 91.

⁸² See: Tsagourias, ‘Non-State Actors, Ungoverned Spaces and International Responsibility for Cyber Acts’, 21 (3) *Journal of Conflict and Security Law* (2016) 455.

⁸³ J. Healey, *Beyond Attribution: Seeking National Responsibility for Cyber Attacks*, Atlantic Council Issue Brief, 22 February 2012, available at <http://www.atlanticcouncil.org/publications/issue-briefs/beyond-attribution-seeking-national-responsibility-in-cyberspace>.

instructions.⁸⁴ From the ICTY's judgment, it transpires that the 'overall control' standard applies to all organized groups and is not restricted to armed groups only; whereas, more recent legal commentary views 'overall control' not only as a criterion for the classification of armed conflicts but also as an attribution criterion.⁸⁵

The advantages of an 'overall control' standard is that it takes an holistic, long-term, multifaceted and, indeed, a cumulative view of the relationship between states and organized non-state actors in contrast to the 'effective control' test which requires specific and overwhelming state input. Secondly, it attributes all the acts of the group to a state without the need to prove the state's input in each and every act as it is the case with 'effective control'.

Its merits can be demonstrated by considering a number of scenarios. The first concerns APT actors committing technically sophisticated attacks on a global scale for instance espionage, attacks on critical infrastructure systems or other high-end attacks. In order to design and execute such attacks, APTs need resources, time, technical knowledge, intelligence, platforms, organisational capabilities, all of which are traits of sophistication. Put differently, such operations require long term investment and close coordination in order to be executed irrespective of their success. Previously we said that the sophistication of the attack and consequently the sophistication of the actor can provide evidence of state involvement but the question to ask is how it can lead to attribution. This depends on discerning what type of relationship between a state and an APT can maintain such sophisticated operations in the absence of institutional, functional or agency links in view also of the targets and goals of the operation. It transpires that the 'overall control' indices of equipping, resourcing and planning provide the most probable description of the modality of relations between states and APTs in this instance by placing them on the spectrum of state control. Thus, and although it was not put in these terms, when the UK said that APT10 has an 'enduring relationship with the Chinese Ministry of State Security and operates to meet Chinese State requirements',⁸⁶ the 'overall control' criterion could form the basis of attributing its malicious cyber activities to China. Equally Stuxnet can be attributed to the US and/or Israel on that basis.

The second scenario concerns acts of cyber theft committed by privately-owned Chinese enterprises (POE). Any discussion of attribution in this case should take into consideration China's expansive notion of national security, which also includes its economy,⁸⁷ as well as the 2017 National Intelligence Law according to which 'any organization or citizen shall support, assist and cooperate with the state intelligence'.⁸⁸ It should also take into account China's economic model, which does not distinguish clearly between private and public

⁸⁴ Appeal Judgement, *Prosecutor v Tadić*, Case No, ICTY-94-1-A, Appeal Chamber, 15 July 1999, § 131; M. N. Schmitt, *Tallinn Manual 2.0*, *supra* note 74, at Rule 82, paras 3-8.

⁸⁵ ICRC, Commentary of 2017 on the Second Geneva Convention, available at https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Comment.xsp?action=openDocument&documentId=1A35EE65211A18AEC12581150044243A-107_B, at paras 287-295; Although the ICJ rejected the 'overall control' for purposes of responsibility, it acknowledged the existence of some flexibility as was said previously; *Bosnia Genocide*, *supra* note 5, at paras 405-6; Cassese, 'The Nicaragua and Tadić Tests Revisited in Light of the ICJ Judgment on Genocide in Bosnia', 18 (4) *European Journal of International Law* (2007) 649.

⁸⁶ 'UK and Allies Reveal Global Scale of Chinese Cyber Campaign', *supra* note 15.

⁸⁷ National Security Law of the People's Republic of China, (adopted 1 July 2015), Art. 2, available at http://eng.mod.gov.cn/publications/2017-03/03/content_4774229.htm.

⁸⁸ National Intelligence Law of the People's Republic of China (adopted 27 June 2017), Art. 7, available at http://cs.brown.edu/courses/csci1800/sources/2017_PRC_NationalIntelligenceLaw.pdf; See, however, Declaration of Jihong Chen and Jianwei Fang on behalf of Huawei, 27 May 2018, available at <https://thechinacollection.org/wp-content/uploads/2019/03/Huawei-Declaration.pdf>.

enterprises but establishes a network of relations between the state and these companies.⁸⁹ These factors blur the distinction between the state and the private sector and make the current attribution determinants inapplicable. To explain, POEs are not *de jure* organs whereas the network of relations between POEs and the government do not amount to complete dependence and control in order to make them *de facto* organs since the companies can make their own business decisions. There are no instructions, at least explicit and direct ones, and no direction since POEs make their own decisions. There is also no ‘effective control’ by the government over acts of cyber theft. Finally, there is no authorization to commit cyber theft and of course questions can be asked as to whether cyber theft is part of commercial or governmental functions.⁹⁰ If, however, one takes into account the material support POEs receive from the government in the form of aid or preferential loans, the government’s or the party’s participation in the planning of their activities through managerial or party appointments as well as all the surrounding laws, administrative guidelines and informal rules, it can be said that the government exercises ‘overall control’ over them and thus acts of cyber theft should be attributed to China.

Be that as it may, it should be noted that, even if the ‘overall control’ standard can extend the attribution net, it does not necessarily address all the difficulties surrounding cyber attribution. As was noted, the ‘overall control’ standard applies to organized groups; however, purely cyber groups may not exhibit the requisite level of organization. This means that the ‘overall control’ standard is more useful in relation to offline groups which also operate online. Furthermore, the required level of state input is still demanding but, as was said, a cyber actor may not need financial support, training or other assistance. Instead, non-state actors may lend their services to a state out of a sense of allegiance without the state participating in the planning of their activities.

The case of patriotic hackers is indicative in this regard. Patriotic hackers are individuals or groups who act in support of their country’s policies or goals.⁹¹ As the Honker Union of China, a now disbanded patriotic hacker group, declared, its campaign was to ‘safeguard national unity, protect China’s national sovereignty, resist foreign bullies, and deflate anti-China arrogance.’⁹² Patriotic hackers are not state organs; they are not empowered by the state; they are not explicitly and directly instructed by the state; they do not act under the state’s direction in the form of leadership and the state’s input is below the ‘overall control’ or ‘effective control’ thresholds⁹³ – even if the state encourages them or condones their actions. However, what defines their relationship with the state is their shared values and goals which, on the one hand, allows the state to influence their behaviour and to shape their actions whilst, on the other hand, it makes patriotic hackers receptive and responsive to such influence. This shows that even

⁸⁹ Milhaupt and Zheng, ‘Beyond Ownership: State Capitalism and the Chinese Firm’, 103 *Georgetown Law Journal* (2015) 665; Williams, ‘The ‘China, Inc.+’ Challenge to Cyberspace Norms’, *Hoover Institution, Aegis Series Paper No. 1803*, available at https://www.hoover.org/sites/default/files/research/docs/williams_webreadypdf1.pdf, at 69; see also: FireEye, *Red Line Drawn: China Recalculates Its Use of Cyber Espionage*, FireEye Special Report, (2016), available at <https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/rpt-china-espionage.pdf>.

⁹⁰ For public/commercial functions, see *Emilio Agustín Maffezini v the Kingdom of Spain*, Award on the Merits, ICSID Case No. ARB/97/7 (13 November 2000), at paras 52-83.

⁹¹ P. W. Singer and A. Friedman, *Cybersecurity and Cyberwar: What Everyone Needs to Know* (2014), at 111.

⁹² X. Wu, *Chinese Cyber Nationalism: Evolution, Characteristics, and Implications* (2007), at 56.

⁹³ Hang, ‘Freedom for Authoritarianism: Patriotic Hackers and Chinese Nationalism’ 5 *Yale Review of International Studies* (2014) 47.

softer versions of control which are value-oriented can be equally effective in aligning non-state actors and states and in engendering desired outcomes. For this reason, we propose a criterion of ‘soft control’⁹⁴ to capture cases where a state exerts influence over non-state actors due to shared values who then act in pursuance of a state’s policies or goals, thus complementing state action. On the basis of this criterion, attacks by Chinese patriotic hackers can be attributed to China and electoral attacks by Russian patriotic hackers following President Putin’s call that ‘[i]f they [hackers] are feeling patriotic, they will start contributing, as they believe, to the justified fight against those speaking ill of Russia’⁹⁵ will be attributed to Russia.

The case of patriotic hackers reveals another gap in the attribution determinants of Article 8, ARSIWA, namely the role of implicit instructions. Reverting to the example of Chinese patriotic hackers, their actions should be assessed against China’s military strategy, which views civilians and the army working together in times of peace or in times of war to support and defend the nation.⁹⁶ Within such a context where the dividing lines between state and private and between civilian and military are blurred, it can be said that Chinese patriotic hackers operate within a system of authority and submit to the will of the state willingly or sometimes unwillingly. Consequently, when the Chinese government identifies enemies of the state which are then attacked by patriotic hackers, the plausible inference that can be made is that they have acted under the implicit instructions of the government. The existence of instructions in this case can be proved through circumstantial evidence such as communications in governmental outlets, the type of attacks and their targets, the tempo of attacks, their timing, or the beginning and end of such attacks. The same reasoning can apply to Chinese POEs in the light of the formal and informal regulatory and managerial framework within which they operate as mentioned above, but there should be a case-by-case analysis to establish whether the POE implemented in the particular instance a governmental (implicit) instruction. In light of the above, we believe that implicit instructions should be included in Article 8 ARSWIA attribution determinants.

C. Standard of Proof

The third proposal tries to identify a suitable standard of proof for cyber attributions. It was said in section five that current attribution claims do not rely on any particular standard of proof and that international law is quite ambiguous in this regard. This offers a degree of flexibility, but in order to identify a standard of proof which is suitable for cyber attribution, it is important to understand the nature and function of the ‘standard of proof’. It should be said in this regard that the standard of proof is context-specific and hence it may vary depending on the domain and circumstances. Also, the standard of proof is gradated; it describes degrees of confidence but never absolute certainty. This is true in law as it is in political (intelligence) assessments, which use standards of low, medium or high confidence and it is also true in technical assessments as noted in section three. Regarding its function, the role of the standard of proof is to assist decision-makers to make reasonable determinations on the basis of the available evidence, which is especially important when the quantity of evidence is low, and its quality needs to be assessed as in the case of cyber evidence.

Applying the above considerations to cyberspace indicates that applying a very high standard of proof such as that of ‘beyond a reasonable doubt’ may increase the factual and legal

⁹⁴ It draws from the concept of soft power. See: J.S. Nye, ‘Soft Power’ 80 *Foreign Policy* (1990) 153.

⁹⁵ FT, *Putin Concedes ‘Patriotic’ Hackers Might Target Foreign Elections*, 1 June 2017, available at <https://www.ft.com/content/f607ac6c-46e6-11e7-8519-9f94ee97d996>.

⁹⁶ Peng Guang Qian, ‘The Twenty-First Century War: Chinese Perspectives’ in Y. Boyer and J. Lindley-French (eds), *The Oxford Handbook of War* (2012) 279.

reliability of attribution but it may frustrate determinations and consequential action; whereas, applying a low standard may affect the reliability of attribution and may lead to erroneous decisions with serious repercussions for the acting as well as the victim state.

In our opinion, the ‘preponderance of the evidence’ is the most appropriate standard for proving attribution in cases of less serious malicious cyber-operations. According to this standard, attribution will be established if more evidence supports a particular attribution finding than contradicts it. The ‘preponderance of the evidence’ is not a new standard but one that has already been used in international jurisprudence as was shown in section five and it is a standard which is also employed outside the legal context. For example, with regard to the attribution of the WannaCry attack, the NSA determined that the ‘preponderance of the evidence’ points to North Korea.⁹⁷

The main reason why this standard is promoted as the most appropriate standard for proving cyber attribution is because it balances the need for reliable attributions and the need for making such determinations and for holding those responsible to account. To explain, this standard ensures diligent scrutiny of the evidence which a lower standard cannot achieve but it does not unnecessarily impede decisions or actions which a higher standard would. Furthermore, it ensures that states will incur responsibility for their malicious cyber-operations which a higher standard cannot do because states would be able to operate below that higher standard but at the same time, it does not trivialise the institution of state responsibility and preserves its regulatory compass which a lower threshold would fail to do. Finally, from a practical perspective, it recognizes the difficulties in obtaining and in assessing cyber evidence and the need to make use of circumstantial evidence. In short, the ‘preponderance of the evidence’ standard reconciles legal, policy and factual considerations and ensures that reasonable determinations are made on the basis of available and sufficient evidence.

7. Conclusion

The preceding discussion has shown that attribution matters for the suppression of malicious cyber-operations but also that it is a multifaceted process. The article has presented the methodology and determinants of the technical and legal processes of attribution and explained how they interact. It has also considered what body of evidence can be used to prove attribution and what are the applicable standards of proof. The discussion identified flaws and gaps in the existing methodology, determinants and evidentiary standards which lead to responsibility gaps.

For this reason, the article discussed a number of institutional, normative and evidentiary proposals with the aim of improving the methodology and process of legal attribution in order to close the identified responsibility gaps. The article discussed first the possible contribution of an international attribution agency to the streamlining and standardization of attribution, but concluded that the creation, operation and effectiveness of such an agency would depend on state cooperation which is not forthcoming at this point in time. The article then put forward a number of normative proposals to re-adjust the legal determinants of attribution. They include amplified standards of control in the form of ‘overall control’ and ‘soft control’ and the inclusion of implicit instructions as an attribution determinant. In the authors’ opinion, these proposals can capture better the realities of cyberspace, the place and role of non-state actors therein and their multilevel relationships with states and, thus, address the responsibility deficit that the narrowness of the existing determinants generates. The article finally introduced the ‘preponderance of evidence’ as the most appropriate standard for proving attribution because

⁹⁷ ‘The NSA has Linked the WannaCry Computer Worm to North Korea’, *supra* note 13.

it guarantees adequate scrutiny of the evidence whilst also facilitating attribution determinations.

All in all, the article is informed by our belief that international law should not withdraw from cyberspace, but that it should face and shape cyber reality by establishing a regulatory framework within which states, individuals and other entities can operate and be held to account. Our normative and evidentiary proposals assist in building a legal framework for ensuring responsibility and for inculcating a culture of responsibility in cyberspace.