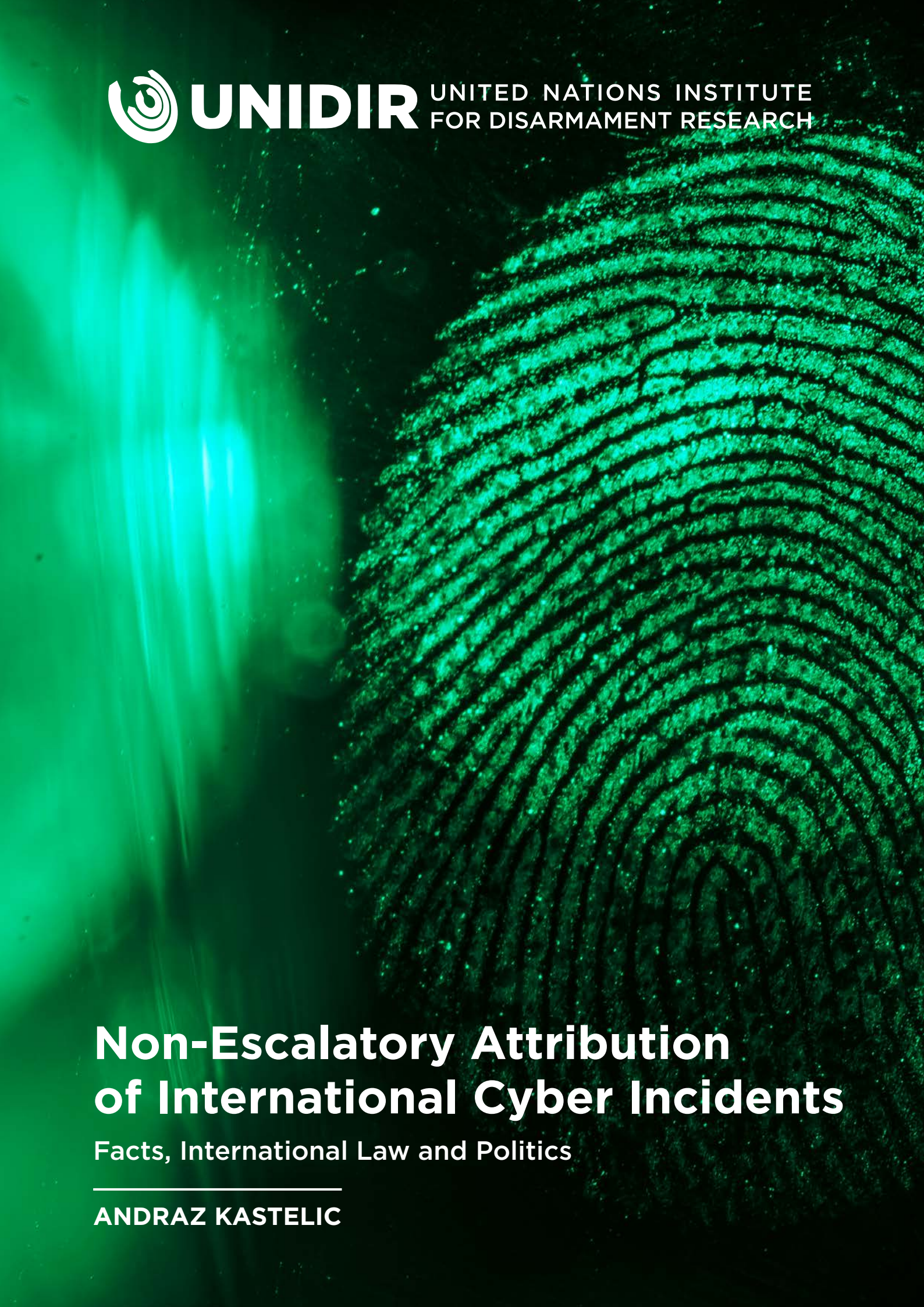

UNIDIR UNITED NATIONS INSTITUTE FOR DISARMAMENT RESEARCH

Non-Escalatory Attribution of International Cyber Incidents

Facts, International Law and Politics

ANDRAZ KASTELIC

ACKNOWLEDGEMENTS

Support from UNIDIR's core funders provides the foundation for all of the Institute's activities. This study was produced by the Security and Technology Programme, which is funded by the Governments of Germany, the Netherlands, Norway and Switzerland and by Microsoft. The author wishes to thank the following individuals for their invaluable advice and assistance on this paper: Cecile Aptel (UNIDIR), Gerardo Isaac Morales Tenorio (Mexico), Giacomo Persi Paoli (UNIDIR), Samuele Dominioni (UNIDIR), Serge Droz (Forum of Incident Response and Security Teams); and the participants in the UNIDIR multi-stakeholder dialogue "Political, Technical and Legal Aspects of Attribution: Multi-Stakeholder Dialogue on the Norms of Responsible State Behaviour in Cyberspace", held on 6 July 2021: Kristen Eichensehr (University of Virginia School of Law), Max Smeets (ETH Zurich), Rolliansyah Soemirat (Indonesia), and Stéphane Duguin (CyberPeace Institute).

ABOUT UNIDIR

The United Nations Institute for Disarmament Research (UNIDIR) is a voluntarily funded, autonomous institute within the United Nations. One of the few policy institutes worldwide focusing on disarmament, UNIDIR generates knowledge and promotes dialogue and action on disarmament and security. Based in Geneva, UNIDIR assists the international community to develop the practical, innovative ideas needed to find solutions to critical security problems.

NOTE

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The views expressed in this publication are the sole responsibility of the author. They do not necessary reflect the views or opinions of the United Nations, UNIDIR, its other staff members or its sponsors.

ABOUT THE AUTHOR

Andraz Kastelic is the lead Cyber Stability Researcher of the Security and Technology Programme at UNIDIR. Prior to joining UNIDIR, he held various research positions at international organizations and research institutions around the world.

Table of contents

Executive summary

Attribution – the process of allocating responsibility for a malicious cyber operation – is comprised of three distinct and intertwined aspects: factual or technical, legal, and political. This paper analyses these three aspects through the prism of the normative expectations of responsible State behaviour in cyberspace. As a result, the paper makes a number of suggestions of how to consider the challenges of attribution and how to operationalize norm B of the 2015 Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.

To this end, and with a view to facilitating non-escalatory reaction to an information and communications technology (ICT) incident, the prevention of conflicts and the peaceful settlement of disputes, this paper suggest that States consider the following measures:

Section 3

a) Developing technical capacity to enable investigation of ICT incidents. This includes capacities related to the collection and custody of evidence.

b) Establishing and institutionalizing mechanisms for cooperation between relevant domestic stakeholders. The attribution process would benefit from engaging the policy and political community, law enforcement, the intelligence apparatus, the private sector, academia, the diplomatic community and legal experts.

c) Establishing cooperative mechanisms with international partners. This will facilitate cooperation when an incident occurs, facilitate prevention of potential escalatory mis-attribution and facilitate trust among States.

d) Indicating the degree of confidence in the conclusions of technical or factual investigation and the analysis of the cyber incident. This will enable the decision makers to make an informed decision regarding the communication of the attribution claim and its format. It will also assist them as they consider the response.

e) Approaching the attribution in law in accordance with the provisions of the international law of State responsibility if the cyber operation in question is deemed to be internationally wrongful.

f) Developing interpretation of the international law of State responsibility in the context of ICT operations.

g) Sharing those interpretations with the wider international community. This would contribute to the progressive development of the international law applicable to State conduct in cyberspace.

h) Engaging in dialogue with all the States involved, and only resorting to any international reactions by way of, for example, retorsion or countermeasures when the dialogue fails.

Section 4

i) Taking into account the respective benefits and potential pitfalls of public and private attribution, when deciding on the format of attribution based on its desired effects, each State should critically in cases the evidence before joining a declaration of attribution, just as they would in cases where attribution rests on claims reported by private enterprises.

j) Aiming to substantiate the attribution claims, regardless of whether the cyber operation constitutes a breach of international law.

k) Striving to rest the attribution claims on multiple sources of reliable and objective evidence, prioritising outcomes of technical forensic analysis, to be supplemented by circumstantial evidence.

l) Continuing to engage in international discussions related to the standard of proof expected in the context of attribution.

m) Resorting to various confidence-building measures and sharing the standards of proof with the international community. This would contribute to defining the limits of acceptable attribution practice and to the development of the customary international law.

n) Using various international forums to further discuss the proposal for an international attribution mechanism, including the envisioned benefits and pitfalls.

1. Introduction and content

Information and communications technologies (ICTs) have endowed our societies with immense opportunities.[1] The challenges of the digital domain to international peace and stability however "risk overshadowing [the] benefits".[2] Malicious cyber operations are more and more frequent[3] and so are the accompanying attribution claims.[4] Attribution of malicious cyber operations, however, is a challenging endeavour and erroneous claims of responsibility "could result in significant consequences, including in unintended armed responses and escalation".[5]

Building on the work of past United Nations processes on international ICT security[6] and acknowledging the challenges of attribution, the 2021 Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security (GGE) suggested:

> "In case of ICT incidents, States should consider all relevant information, including the larger context of the event, the challenges of attribution in the ICT environment, and the nature and extent of the consequences."[7]

This norm sets an expectation that a State targeted by a malicious cyber operation will consider the context of the incident when deliberating its reaction, as accurate attribution is key to a non-escalatory response to a malicious inter-State cyber operation. This is particularly the case in the context of an ICT incident that amounts to a breach of international law, especially when the initial malicious cyber operation is considered to have amounted to an armed attack, triggering the right to self-defence by cyber as well as other means.

In order to facilitate implementation of the norm, this paper elaborates on:

- The concept of attribution, including its purpose and the related challenges
- The three aspects of attribution: technical or factual, legal, and political, which, despite the doctrinal demarcation between them, are very much related and often intertwined[8]
- Aspect-specific suggestions for operationalization of the norm
- The proposals for an international attribution mechanism

1 UNGA (2021a, Sec. II).
2 UNSG (2021, 62).
3 UNGA (2021a, Sec. II).
4 Egloff and Wenger (2019, 1); Tsagourias and Farrell (2020).
5 UNODA (2021).
6 UNGA (2015a); UNGA (2013).
7 UNGA (2021a, Norm 13 (b)).
8 Lin (2016, 13).

2. The concept of attribution

Attribution denotes the allocation of direct or indirect responsibility for a malicious cyber operation. It involves the determination of the origin or authorship of the cyber operation. Generally, thus, a cyber operation can be attributed to a natural person or a group of natural persons, an ICT system of origin, or a State. Given the context of advancing responsible State behaviour in cyberspace, the primary focus of the norm and therefore of this paper is the attribution of a malicious cyber operation to a particular State or group of States.

Attribution involves a process of several distinct and potentially overlapping elements:

– Evidence collection and analysis
– Legal analysis
– Decision-making and communication

Evidence collection and analysis are part of the fact-finding stage of attribution, providing the foundation for the legitimacy of the following political act of attribution of a malicious cyber operation. In the event that a cyber operation amounts to a breach of the international rights of a State, technical or factual investigation and analysis are key elements in the legal test of attribution, in invocation of State responsibility and in employment of countermeasures as authorized by customary international law.

Countermeasures *are non-punitive, compliance-inducing measures intended to secure cessation of and reparation for internationally wrongful conduct. Otherwise unlawful, wrongfulness of a countermeasure taken in response to an internationally wrongful act is precluded according to the customary international law of State responsibility.*[9]

Countermeasures must be

- *non-forcible in nature*[10]
- *reversible as far as possible*[11]
- *quantitively as well as qualitatively proportionate to the initial wrongdoing*[12]

The methods of countermeasures are not limited to the methods of the initial wrongdoing or, in other words, if an internationally wrongful act is perpetrated by way of a cyber operation, countermeasures need not be a cyber operation.

2.1. The purpose of attribution

The ultimate purpose of attribution is to induce compliance either by deterrence or by imposing costs on the wrongdoing party. This promotes accountability and the rule of international law.

In pursuing deterrence, attribution could be a way of signalling the limits of expected State behaviour in cyberspace.[13] The limits may be the universally agreed voluntary norms

9 UNGA (2002, Annex, Art. 22).
10 UNGA (2002, Annex, Art. 50, para. 1(a)).
11 ICJ (1997); ILC (2001, 312).
12 Case concerning the Air Service Agreement of 27 March 1946 between the United States of America and France (1978, Paras 83–90); ICJ (1997).
13 For example, "Belgium's cybersecurity policy also provides for a new attribution mechanism which is conceived as a deterrent tool. If we want to effectively prevent and deter malicious cyber activities in an environment where cyber attacks are growing in number and in complexity, formal attribution of malicious cyber activity targeting a vital organization in Belgium is an important instrument." Kingdom of Belgium (2021, 2). Consider also the positions of the United States Department of Defense (2015, 11): "Attribution is a fundamental part of an effective cyber deterrence strategy"; and the United Kingdom (Foreign and Commonwealth Office (2019)): "[attribution] can be a powerful deterrence tool when deployed effectively".

adopted by the United Nations General Assembly in 2015[14] or other arbitrary limits of acceptable behaviour.[15] Additionally, attribution may signal the attribution capabilities of the targeted State, thus dispelling the popular notion that attribution is nearly impossible and the myth of the anonymity of malicious cyber deeds.

Malicious cyber operations can also be attributed with a view to imposing costs on the wrongdoing party[16] and in doing so alter its rational choice calculus.[17] Attribution coupled with a detailed description of the malicious cyber operation can expose the exploits used, giving the opportunity to the ICT community to patch the vulnerabilities, and so rendering the malicious tools less effective in the future. Moreover, public attribution claims may impose reputational costs on the responsible party. Scholarship has emphasized the importance of reputation in international relations.[18] However, the potential for compliance to be induced by reputational costs incurred due to public ascription of responsibility for non-conformist behaviour remains seriously disputed.[19]

Beyond reputational costs, attribution is key to legitimate imposition of reparations, which is another form of cost for wrongdoers. Being the necessary element of State responsibility, attribution legitimizes the injured State to take countermeasures in order to secure cessation of the internationally wrongful cyber operation and reparations for its consequences.[20]

2.2. Attribution challenges

Several States have pledged investments in their attribution capabilities in the past decade.[21] Nevertheless, the challenges associate with attribution[22] of cyber operations persist and stem from the technological structure of ICT systems. The origin of a malicious cyber operation can be obfuscated by such underlying technology mechanisms as:

- Use of botnets[23]
- The privacy of domain registration[24]
- Various spoofing and false flag techniques[25]
- Use of proxies, virtual private networks (VPNs) and onion routing[26]
- Dynamic allocation of Internet Protocol (IP) addresses
- Covert communication, including encryption[27]
- Using compromised third-party infrastructure

14 UNGA (2015b).
15 Fischerkeller (2021).
16 According to the Kingdom of the Netherlands (2021). "There must be consequences for bad behavior in cyberspace."
17 The Council of the European Union (2017a) has stated that "clearly signaling the likely consequences of a joint EU diplomatic response to such malicious cyber activities influences the behavior of potential aggressors in cyberspace thus reinforcing the security of the EU and its Member States."
18 "The extent to which a violation is known by the relevant players affects the reputational consequences of the violation. Obviously, if a violation takes place, but no other state has knowledge of it, there is no reputational loss. The reputational consequences will also be less if only a small number of countries know of the violation."
19 Guzman (2002, 1863).
20 Huth (1997, 72–99).
21 See section 4.
22 Federal Council (2018, 10); United States Department of Defense (2015, 11–12).
23 See, for example, Mission of Brazil to the UN New York (2021, 1); China (2021).
24 Boebert (2011, 45).
25 Kaspersky (2021c); Kaspersky (2021b); Skopik and Pahi (2020).
26 "Onion routing is a technique by which a series of routers participation in an encryption network. Any client who seeks to conduct online activity with anonymity then sends their internet communications through the onion routing network. The client secures their online communication with several layers of encryption, and selects a set of onion routers that will each individually have the key to decrypt one layer of encryption on the communication, until the communication ultimately reaches its destination fully decrypted. Because each router only has a single layer of decryption, no single router knows the overall path that the communication takes." Tran (2018, 389 fn51).
27 Boebert (2011, 43–47); Khraisat et al. (2019). For an overview of defence evasion techniques of cyber operations, see, for example, MITRE (2021). Note that not all techniques on this list are aimed specifically at obfuscation of the origin.

Further analysis of the range of obfuscation techniques is beyond the scope of this paper and a wealth of dedicated literature exists on this particular topic;[28] suffice to say that the techniques are plentiful, diverse and, much like the threat landscape, ever evolving.[29] This is also one of the reasons why the development of attribution capabilities may require a significant and continuous investment, which may not be available to all States.

Even with the growing number of cases of public attribution of malicious cyber operations, "the lack of accountability in cyberspace"[30] appears to be a challenge. The consequent absence of deterrence may encourage the proliferation of offensive cyber capabilities.[31]

2.3. Misattribution and the potential for escalation

The challenges of attribution elaborated above mean that the likelihood of misattribution should not be dismissed. Although there are perhaps endless theoretical scenarios with the potential to lead to misattribution, it can simply result from an inadequate quality or quantity of evidence, which fails to provide sufficient proof of origin or authorship. Misattribution can also result from hasty or biased political decisions before completion of the factual investigation, which can be a lengthy procedure.[32]

In the initial stage of attribution, the investigators and analysts can be deceived by various spoofing and false flag techniques employed by the author of the malicious cyber operation.[33] Moreover, the political decision-making phase may lead to erroneous attribution conclusions based on a disproportionate reliance on circumstantial evidence and on conclusions by inference. Legal operation of attribution can lead to incorrect results if the nexus between the responsible natural person and a State is not established in line with the relevant provisions of the customary law of State responsibility. This includes attribution claims based on the inaccurate assumption that the geographic origin of a cyber operation automatically entails responsibility of the State with jurisdiction over the particular territory[34] or that the responsible group of individuals acted on behalf of that particular State.

In certain circumstances, "misattribution is a real possibility and can carry serious consequences for international relations, peace and security".[35] For instance, if a State injured by a cyber operation reacts with measures of retorsion taken against the innocent State, this may result in deterioration of the bilateral relationship.[36] This potential for deterioration is further envisioned in the case where the innocent State decides to respond to misattribution with more than rejection of responsibility.

---

28 See, for example, Wheeler et al. (2003); Nicholson et al. (2012, 188–197).
29 Egloff (2020).
30 UNSG (2021, 62).
31 "The political and technical difficulty of attributing and assigning responsibility for cyber-attacks encourages actors to adopt an offensive posture." UNSG (2018, 8–9).
32 "Identifying a targeted attack, profiling the attackers and creating attribution factors for the different threat actors is a long and in depth task; it can take years. Attribution that works is always based on many years worth of previously accumulated data and involves a highly-skilled team of researchers with experience in forensics and Investigation." Kaspersky (2021a).
33 Note, for example, the position of Brazil arguing "cyber operations can be designed to mask or spoof the perpetrator, which in turns increase the risks of miscalculated responses against innocent actors." UNGA (2021b, 22).
34 "[T]he Group recalls that the indication that an ICT activity was launched or otherwise originates from the territory or the ICT infrastructure of a State may be insufficient in itself to attribute the activity to that State." UNGA (2021a, 71(g)).
35 UNSG (2018).
36 Note, for example, the position of Estonia, arguing "States have the right to apply these measures as long as they do not violate obligations under international law. These measures could, for example include the expulsion of diplomats or applying restrictive measures to officials of a third country such as asset freezes or travel bans." UNGA (2021b, 29).

Measures of retorsion *are unfriendly but lawful reaction measures, intended to compel another State to cease behaviour that is considered unfriendly or undesired but is still consistent with the international obligations of the State. "Acts of retorsion involve non-violent, lawful acts by both parties, without resort to armed force."*[37] *Retorsion can manifest in, for instance, various forms of sanctions, embargoes or severing of diplomatic ties among others.*[38]

A threat to peace and security can materialize when the initial cyber operation deprives the targeted State of its international rights. If, for instance, the State injured by the malicious cyber operation launches countermeasures based on erroneous attribution, those countermeasures in themselves constitute a violation of the international legal obligations.

What is more, a self-defence response to a cyber operation amounting to an armed attack based on misattribution can trigger a forcible response and thus considerably worsen international peace and security;[39] in the event of a cyber operation of sufficient scale and magnitude, such a response could potentially include the use of nuclear weapons.[40]

37 Joyner (2006).
38 UNGA (2002, Annex, ch II cmt 3).
39 Kastelic (2020); Wan et al. (2021).
40 United States Department of Defense (2018, 38).

3. Technical and factual aspects of attribution

Following an ICT incident, the attribution process begins with the technical or factual investigation of the incident. **Investigation predominantly focuses on uncovering the origin of the malicious cyber operation and the natural person in control of the originating ICT system.** As such it seeks to establish responsibility in fact.

Technical aspects of attribution frequently consider such elements of the incident as malware (signatures, complexity, code clues) and the systems used by the malicious cyber actors, paying particular attention to elements such as command and control servers, domains used, IP addresses, and traffic analysis among other things.

Scholarship has developed a taxonomy of technical investigation techniques that lists the following clusters of technical attribution methods:

- Manual attribution
- Traceback mechanisms
- Stepping-stone attribution
- Payload attribution
- Honeypots
- Internet redesign[41]

The technical investigation part of attribution is not the exclusive domain of the competent national authorities. Non-State actors, such as private enterprises as well as academia, frequently publish technical reports detailing the operation or operators of malware.[42] This may be used in the investigation efforts of the competent national authorities.[43] The above-mentioned analysis may indeed aid the official investigation of the incident although it does also carry associated risks.[44]

In an attempt to attribute a malicious cyber operation, technical analysis of the incident is supplemented with other sources of information that are not strictly focused on the technology and malware at the centre of the cyber operation.[45]

3.1. Socio-political methods of investigation

The investigation of tactics, techniques and procedures focuses on the methods of the malicious actors.[46] Additional socio-political and behavioural methods of investigation complement the technical analysis; these can seek to establish attribution based on indicators such as who benefitted from the malicious operation;[47] language (mis)used by the malware authors;[48] keyboard layout; domain registration information;[49] behavioural biometric data (keystrokes, mouse and touch input);[50] and the operating hours of the alleged orchestrators of a cyber operation.[51]

41 Nicholson et al. (2012, 191–196).
42 See, for example, Hackdig (2021); Mandiant (2013); Burt (2021); Bencsáth et al. (2011); Marczak et al. (2021).
43 Seitz (2019).
44 Romanosky and Boudreaux (2021).
45 "Attribution could be established, based on an analysis of technical data and all-source intelligence, including on the possible interests of the aggressor". Council of the European Union (2017a, 13). See also, for example, House of Representatives (2010).
46 National Institute of Standards and Technology (n.d.).
47 Baezner and Robin (2017, 8).
48 Kaspersky Lab (2017).
49 FireEye (2014).
50 Tsimperidis et al. (2021, 835); Keromytis (2016).
51 Fire Eye (2018, 27).

3.2. Suggestions for operationalization

In order to strengthen the technical or factual investigation of the incidents and to facilitate accurate and non-escalatory attribution, States could consider the following measures:

a) Developing technical capacity to enable investigation of ICT incidents. This includes capacities related to the collection and custody of evidence. Given the interconnected nature of ICT systems and the fact that the investigation would be likely to benefit from international cooperation mechanisms, States should consider assisting other countries with the development of relevant investigation capacities.[52]

b) Consider establishing mechanisms for cooperation between relevant domestic stakeholders. While the technical and intelligence authorities should probably be the leading entities in the investigation efforts,[53] the process should be inclusive and would benefit from the engagement of representatives of the political and policy, diplomatic and legal communities. Investigation efforts could also benefit from the insight of the private sector.[54] Some States have officially recognized the value of the private sector and have pledged to collaborate in attribution efforts.[55] States should also consider engaging with academia, which can support the incident investigation with advanced and impartial attribution advice.[56] To enable a multi-stakeholder approach to incident investigation, States should enact the relevant enabling domestic legal and policy frameworks.

c) Moreover – and as suggested by the additional layer of understanding in the 2021 GGE report – States "are encouraged to consult among relevant competent authorities".[57] This may contribute to the prevention of potentially escalatory misattribution and facilitate trust among States. To enable such communication, States should consider establishing cooperative mechanisms with international partners, which will facilitate cooperation when the incident occurs.

d) Technical or factual investigation and analysis of the cyber incident should indicate the degree of confidence in the conclusions. This will enable the decision makers to make an informed decision regarding the communication of the attribution claim and its format. It will also assist them as they consider the response to ICT incidents.

52 UNGA (2015a, para 21) does encourage States to assist in capacity-building, although attribution is not mentioned.
53 "After an attack has been detected, technical agencies have the responsibility to come up with an assessment of the nature of the incident." G7 (2019).
54 While attribution process may indeed benefit from the insights of the non-State actors, States wishing to attribute a cyber operation should bear in mind that "the motivations for private attributions and governmental attributions may differ." Eichensehr (2019, 213–217).
55 See, for example, United States Department of Defense (2015, 12): "The Defense Department will continue to collaborate closely with the private sector and other agencies of the U.S. government to strengthen attribution."
56 Egloff (2019).
57 UNGA (2021a, 13b, para 23).

4. Legal aspects of attribution

While the technical or factual investigation seeks to indicate the responsible natural person or the territory of origin and thus to establish responsibility in fact, the legal analysis aims to establish attribution in law. The aim of attribution as a normative operation is to establish the legal responsibility of a State.

Until special rules are agreed upon or the progressive development of international law gives rise to a context-specific framework, the customary international law of State responsibility is considered to be the applicable regime governing the legal attribution of the wrongful cyber operations. The applicability of this particular international legal regime to wrongful cyber operations has been confirmed by several States.[58]

According to the international law of State responsibility, conduct attributable to a State and amounting to a breach of international obligations results in State responsibility.[59] This triggers the emergence of the secondary obligations of cessation, non-repetition and reparation.[60] When the State responsible for the internationally wrongful act fails to comply with the resulting secondary obligations, the injured State is legally empowered to take countermeasures.[61]

Indeed, legal attribution is an important legitimizing factor for reparations as well as, potentially, countermeasures, both of which inflict costs on the responsible State and have an impact on more than its reputation. According to the international law of State responsibility, any State responsible for internationally wrongful cyber operations is under an obligation to cease the non-compliant conduct and to provide reparations to the injured State.[62]

Not all cyber operations and not all responses dictate legal considerations of attribution. In the context of attribution, the 2021 GGE report distinguishes between two types of malicious cyber operation targeting States: cyber operations that amount to a breach of international obligations of States and cyber operations that do not. The normative expectations imply that the former type of operation necessitates legal considerations in the process of attribution.[63] But whether the targeted State should consider the international law of attribution further depends on the desired response to the cyber operation. If the State is considering taking measures of retorsion in response to a cyber operation[64] – whether or not it amounts to an internationally wrongful act – legal considerations of attribution are not necessary. In contrast, if a State is contemplating taking countermeasures,[65] the law of State responsibility prescribes an attribution framework that is designed to prevent the escalation of tension.

58 Brazil, for instance, argued that "[i]n the absence of any lex specialis for cyberspace, the customary norms concerning the attribution of conduct to a State are also applicable to the State's use of ICTs." UNGA (2021b, 21). See also United States (2020, 2); Canada (2019, 2); Australia (2021, Annex B, 1); Romania (2021, 3); Finland (2020, 3). Note that some States maintain reservations about the aspects of the international law of State responsibility in the context of cyber operations. See, for example, Ministry of Foreign Affairs of Japan (2021); China (2020); United Nations General Assembly (2021c, 60).

59 UNGA (2002, Annex, art 2). See the concurring national positions of, for example, Brazil and Norway (UNGA, 2021b).

60 UNGA (2002, Annex, Arts 30 & 31).

61 "In certain circumstances, the commission by one State of an internationally wrongful act may justify another State injured by that act in taking non-forcible countermeasures in order to procure its cessation and to achieve reparation for the injury." ILC (2001, Art 22, cmt 1).

62 "The Group reaffirms that States must meet their international obligations regarding internationally wrongful acts attributable to them under international law." UNGA (2021, para. 71(g)).

63 Cf. UNGA (2015a, paras 28(f) and 13(b)).

64 See section 2.3.

65 See section 2.

One of the elements of the law of State responsibility intended to prevent escalation is the obligation of the injured State to invite the responsible State to return to compliance and to notify it of its intention to take countermeasures. In the Case Concerning the Gabčíkovo-Nagymaros Project, the International Court of Justice (ICJ) argued that, "the injured State must have called upon the State committing the wrongful act to discontinue its wrongful conduct or to make reparation for it".[66] The expectations of engagement in dialogue prior to a reaction to a wrongful cyber operation is repeated in the additional layer of understanding found in the 2021 GGE report.[67]

4.1. The nexus between a State and a natural person

To proceed with the legal attribution analysis of a cyber operation, the injured State would normally need to be in possession of factual evidence implicating the natural person perpetrating the operation. This is not a requirement of the relevant international law or the voluntary norms of responsible State behaviour in cyberspace. It is purely a practical consideration – as the Permanent Court of International Justice put it in 1923, "States can act only by and through their agents and representatives",[68] that is, persons in flesh and blood.[69]

Attribution in law and the responsibility of a particular State cannot be automatically established based on the nationality of the natural person perpetrating the act or the territorial origin of the cyber operation.[70] To establish attribution in law, the injured State should be able to establish a nexus between a natural person perpetrating the cyber operation or the territory of origin and a particular State.

Cyber operations conducted by a State organ, even if not legally empowered to conduct malicious cyber operations, can be legally attributed to a particular State. The same is true for cyber operations conducted by a foreign organ placed at the disposal of another State and operations conducted by entities empowered to exercise government authority. According to the law of State responsibility, an entity being instructed, controlled or directed by a State can also result in State responsibility.[71] The jurisprudence of the ICJ indicates that an act conducted by a non-State actor and later adopted by a State organ as its own can equally amount to the international responsibility of a State.[72]

It remains unclear what degree of control over a non-State actor would constitute a sufficient attribution nexus and allow for the invocation of State responsibility. Jurisprudence suggests two possible interpretations of the law.

On the one hand, the ICJ has argued that States are responsible for the conduct of a non-State actor only when the State in question is found to have directed or enforced the perpetration of the specific acts contrary to international law.[73] A number of States argued for the adoption of this control standard in relation to the legal attribution of cyber operations.[74]

66 ICJ (1997, para 84).
67 UNGA (2021a, paras 24–25).
68 PCIJ (1923).
69 Lauterpacht (1968, 40).
70 "[T]he indication that an ICT activity was launched or otherwise originates from the territory or the ICT infrastructure of a State may be insufficient in itself to attribute the activity to that State." UNGA (2021, para 71(g)).
71 UNGA (2002, Annex, ch II).
72 See UNGA (2002, Annex, art 11); ICJ (1981).
73 ICJ (1986, para. 115).
74 See, for example, Norway, Netherlands and Brazil in UNGA (2021b, 71, 61 & 20).

On the other hand, it may be possible to establish international responsibility of the State when it is considered to have exhibited overall control over the non-State actor. While deliberating on the question of international criminal responsibility, the International Criminal Tribunal for the former Yugoslavia (ICTY) also touched upon State responsibility. In a 1999 judgement, the tribunal argued that the overall control test would be satisfied (so giving rise to State responsibility) when a State is found to have been "not only … equipping and financing the group, but also … coordinating or helping in the general planning of [the malicious] activity".[75] However, overall control is considered to be sufficient only for the purpose of establishing State responsibility in the event of cyber perpetrators being found to have acted in an "organized and hierarchically structured"[76] manner. Despite these limitations to applicability, scholars such as Shackelford and Tsagourias have argued against disregarding the overall control standard in the context of cyberspace.[77] States currently remain hesitant to endorse the use of this standard in the context of cyber operations.

4.2. Suggestions for operationalization

Considering this, implementation of the following suggestions will facilitate non-escalatory attribution of ICT incidents and contribute to the progressive development of the relevant international law.

a) If the cyber operation in question is deemed to be internationally wrongful, the injured State should approach the attribution in law in accordance with the provisions of the international law of State responsibility. By following the letter of the law, the potential for the injured State to take wrongful countermeasures and thus contribute to aggravating international relations is reduced.

b) States should develop interpretation of the international law of State responsibility prior to the occurrence of any malicious ICT incident. In particular, each State should consider legal standards emerging from the existing jurisprudence and develop a national interpretation of the relevant international law of attribution. Note, however, that outside the judicial setting, the standard of control sufficient to establish a nexus between the natural person perpetrating a cyber operation and a State remains a political decision.

c) To facilitate transparency and predictability of international relations, States should share their interpretations of the international law of State responsibility with the wider international community.[78] This would contribute to the progressive development of the international law applicable to State conduct in cyberspace.

d) Before reacting by way of retorsion or countermeasures, the injured State should engage in dialogue with all the States involved. This is an expectation of the norm and also a requirement of the law of State responsibility.[79] The dialogue will provide the opportunity for the allegedly responsible State or States to challenge the attribution claims and provide evidence in rebuttal.

75 ICTY (1999, para 131).
76 Cassese (2007).
77 Tsagourias (2012); Shakelford (2010).
78 They could do this, for example, by using the Cyber Policy Portal (www.cyberpolicyportal.org), a confidence-building tool and a repository of national policies and legislation of all the United Nations Member States. The Portal has been officially recognized by the 2021 consensual reports of both the Open-ended Working Group and the GGE.
79 UNGA (2021(a), 13(b)); UNGA (2002, Annex, art 52(1)).

5. Political aspects of attribution

Political aspects of attribution materialize in the act of allocating responsibility. Indeed, some scholars have labelled this as "an inherently political act".[80] In fact, States are under no obligation or expectation to attribute a malicious cyber operation, and assigning responsibility remains a sovereign prerogative of every State.[81]

5.1. Format of attribution

One political decision related to attribution is the choice of the format of attribution. Depending on the purpose of attribution, it can be made in private or in public.[82] Public attribution relies on publicity,[83] which may impose reputational costs on the accused State, and overtly signal a particular normative interpretation or technical capability for attribution. On the other hand, public attribution may carry a risk of deterioration of international relations. In these cases, States injured by a malicious cyber operation may wish to communicate the attribution claims in private, which is likely to inflict less strain on the relationships. As is the case with public attribution, private attribution signals attribution capacities and an interpretation of the norm. However, since the audience is limited to the allegedly responsible party, private attribution can reduce the reputational impact. Moreover, by communicating attribution in private, the parties can present evidence and confront arguments surrounding attribution.[84]

In recent years, public collective attribution claims have emerged as the preferred attribution format for some coalitions of States.[85] Collective action may raise the perceived level of confidence in the attribution and therefore increase the reputational costs inflicted. However, it is possible that not all coalition members share the same attribution capabilities and so may not be able to independently verify the assertions of the State that claims to have proof of attribution or has initiated the collective attribution. In such a case, the attribution process rests on trust.[86] Additionally, collective attribution requires the alignment of strategic interests and may be slowed by the alignment process needed for the coalition's collective attribution.[87]

5.2. Confidence in the attribution

An important part of the relevant political considerations is confidence in the factual underpinning of the attribution claims. Attribution based on proof of dubious quality may carry the risk of misattribution, leading to deterioration of international relations or worse. Whether or not a cyber operation was in violation of the rights of a State, the quality of proof and consequentially attribution confidence should be proportional to the risk carried by the planned response.

80 Egloff (2019); Global Commission on the Stability of Cyberspace (2019, 24).
81 Individual State positions, such Australia's, have also argued that assigning responsibility to a particular State is a political decision; "Australia will, *in its sole discretion, and based on its own judgement*, attribute unlawful cyber activities to another State." Commonwealth of Australia, Department of Foreign Affairs and Trade (2021, 100) [emphasis added]. Note also the position of the G7 (2019, 2) countries considering "that attribution is sovereign political decision, taken on a case-by-case basis with due consideration for all relevant information".
82 Public versus private is a simplified dichotomy for the purpose of this paper. See, for example, UNITAR (n.d.); Berridge (2015); Collins and Packer (2006).
83 Collins and Packer (2006, 10).
84 Collins and Packer (2006, 11).
85 European Parliament (2021, para 33). "The imposition of consequences will be more impactful and send a stronger message if it is carried out in concert with a broader coalition of like-minded States." United States (2018). See also Stilgherrian (2018).
86 Ivan (2019).
87 Ivan (2019).

The burden of the decisions related to confidence and therefore the adequate standard of proof substantiating the aforementioned legal nexus between a State and a perpetrating entity is on the State targeted by a malicious cyber operation. Standards of proof supporting legal attribution outside a judicial setting are not exact; this has been reiterated by a number of States.[88] However, this may very well change in the future; during the OEWG 2021, "some States highlighted the importance of *genuine, reliable and adequate proof* in this context"[89] while other individually argued in favour of adequate, convincing[90] and sufficient evidence.[91] Until then, whether the evidence available amounts to a credible attribution is a matter of political deliberation and subject to self-imposed standards of proof.

The standard of proof should be proportionate to the gravity of the alleged wrongdoing.[92] It also depends on the nature of the malicious cyber operation and the intended reaction of the victim State. If the cyber operation qualifies as internationally wrongful conduct and the injured State wishes to employ countermeasures, the self-imposed standards of proof related to attribution claims should be as high as possible in order to avoid misattribution and unlawful countermeasures, which could lead to the deterioration of international relations, including negative impacts on international peace and security. Particularly high self-imposed standards should be considered when the malicious cyber operation reaches the threshold of an armed attack and the targeted State is contemplating self-defence measures. Conversely, more lenient standards of proof can be followed in the event that the reaction of the targeted State is within the realm of merely unfriendly acts.

In establishing attribution for the purpose of invoking State responsibility, the injured State is under no obligation to provide evidence and to prove its attribution claims.[93] Nevertheless, in line with the 2021 GGE report, the State targeted by an internationally wrongful cyber operation "should substantiate"[94] the attribution claims. In consideration of a non-escalatory response and the aforementioned challenges of technical attribution, substantiation of the attribution claims would offer the accused State an opportunity for a rebuttal and to scrutinize the evidence presented. Attribution claims substantiated by credible evidence would also facilitate dialogue between the States engaged in an exchange about the ICT incident, which is also a recommendation provided by the additional layer of understanding provided by the 2021 GGE report.[95]

88 See Finland, United States and the Netherlands in UNGA (2021b).
89 UNGA (2021d) [emphasis added]. Similar argument advanced by Islamic Republic of Iran (2021).
90 Consider, for example, the position of Finland (2020, 6) arguing for an "adequate proof of the source of the operation and convincing evidence of the responsibility of a particular State." See also Government of the Netherlands (2019, 9); Global Commission on the Stability of Cyberspace (2019, 24).
91 United States and Germany in UNGA (2021b, 141 & 40).
92 "Claims against a State involving charges of exceptional gravity must be proved by evidence that is fully conclusive" ICJ (1949).
93 "[T]here is no international legal obligation to reveal evidence on which attribution is based prior to taking an appropriate response." Council of the European Union (2017b).
94 "[A]ccusations of organizing and implementing wrongful acts brought against States should be substantiated." UNGA (2021a, 71(g)). Note that some States consider the expectation to substantiate attribution claims to be an international obligation. See, for example, the position of Brazil (UNGA (2021b, 21): "difficulties must not serve as a justification to lower the bar for determinations on attribution, which must be substantiated." [emphasis added]
95 UNGA (2021a).

Much of the evidence collected by socio-political and behavioural methods of inquiry into attribution[96] can be considered as circumstantial. The 2010 GGE report argued that the perpetrating party can often be established through a combination of circumstantial evidence and reasoning by inference.[97] Similar appreciation for circumstantial evidence can be found in ICJ jurisprudence. In fact, in the Corfu Channel Case, the ICJ allowed for "a more liberal recourse to inferences of fact and circumstantial evidence"[98] given that the direct evidence of the alleged wrongdoing was only available on territory under exclusive foreign jurisdiction, preventing the injured State from pursuing effective investigation. However, in subsequent cases, the court has shown prudence when accepting circumstantial evidence and reasoning with inference.[99] To avoid erroneous, and thus potentially escalatory, claims of attribution, States should also exhibit caution when asserting attribution based on circumstantial evidence.

5.3. Suggestions for operationalization

In light of the arguments above, implementation of the following suggestions related to the political aspect of attribution will facilitate non-escalatory attribution of ICT incidents.

a) States should decide on the format of attribution based on the desired effects of the attribution, taking into account the benefits and potential pitfalls of public and private attribution. When attribution is jointly made by a group of States, each State should critically assess the evidence before joining a declaration of attribution, just as they would in cases where attribution rests on claims reported by private enterprises. To strengthen a joint attribution, some scholars have suggested greater transparency by the involved partners.[100]

b) States should aim to substantiate their attribution claims, regardless of whether the cyber operation constitutes a breach of international law. This will not only facilitate the stability of international relations and the dialogue between the involved States, but would also allow the allegedly responsible State to review the evidence and possibly provide counter-arguments.

c) Factual investigation is frequently supplemented by outcomes of socio-political attribution methods or even considered to be an alternative to computer science. In asserting attribution, States may resort to circumstantial evidence. While these considerations are not without merit, their utility remains limited as the outcomes of socio-political methods of attribution should not be the sole basis for any unfriendly reaction by the injured State. Circumstantial evidence can only serve as a supplement to evidence collected by way of technical forensic analysis. In any case, State should strive to rest the attribution claims on multiple sources of reliable and objective evidence, prioritising outcomes of technical forensic analysis, to be supplemented by circumstantial evidence.

96 See section 3.1.
97 "Often, the perpetrators of such activities can only be inferred from the target, the effect or other circumstantial evidence." UNGA (2010).
98 ICJ (1949, 18).
99 Teitelbaum (2007, 157).
100 "To be sure, some States and other constituencies may credit even unsupported attributions, such as those made by their allies. But such 'trust us' attributions are unlikely to fully persuade anyone besides close allies. Building broader coalitions to accept attributions and ultimately condemn the underlying cyberattacks will require greater transparency." Eichensehr (2020).

d) The self-imposed standard of proof should be commensurate to the gravity of the malicious cyber operation and the planned reaction by the injured State. However, it is advisable for the State making the attribution claims to aim for the highest standard possible to avoid potential misattribution and any consequent straining of international relations.[101] Additionally, States should continue to engage in international discussions related to the standard of proof expected in the context of attribution.

e) States should consider resorting to various confidence-building measures and share their standards of proof with the international community. This would contribute to defining the limits of acceptable attribution practice and to the development of the customary international law.

101 "[D]ifficulties must not serve as a justification to lower the bar for determinations on attribution." Brazil in UNGA (2021b, 21).

6. A possible international attribution mechanism?

In response to these challenges, several scholars and civil society organizations have proposed the establishment of a special international attribution mechanism. Such a mechanism draws inspiration from fact-finding mechanisms like those envisioned under the Organisation for the Prohibition of Chemical Weapons (OPCW), the International Atomic Energy Agency (IAEA) and the Comprehensive Nuclear-Test-Ban Treaty (CTBT).[102]

The proposal envisions an independent and impartial body, be it ad hoc or regular, that complements sovereign attribution prerogatives. The mechanism would have investigative and assistance roles and would not amount to international arbitration. Accordingly, the mechanism should not make any claims of a breach of the primary international obligations of States or international responsibility.

It is envisioned that the utility of this mechanism would be attractive notably to States with less advanced attribution capabilities. The independence of the mechanism could assist with attribution upon request when a State injured by a cyber operation does not have sufficient capacity to attribute or when its attribution claims remain contested by the accused party. The conclusions of the attribution body's analysis could also be used outside the bilateral setting or by international adjudication entities when requested to adjudicate a conflict related to a cyber operation.[103]

The mandate of such a body would therefore be limited to the collection and appraisal of the technical, factual evidence; evaluation would be conducted by a multi-stakeholder body, and the process would allow for a rigorous peer review process. Additionally, transparency of the investigation process would be imperative as other investigation mechanisms, such as those of the OPCW, have been openly criticized by some States for their alleged lack of transparency,[104] have been seen as biased,[105] and their findings described as "unreliable and technically unconvincing".[106]

Several questions surrounding the proposed mechanism remain. There may be hesitation on the part of States to allow the fact-finding body to investigate the networks or systems on their sovereign territory. There may also be limited willingness to cooperate over concerns related to national security or sovereign prerogatives.

International discussions about such a mechanism appear to have been limited and relevant official State positions hardly ubiquitous. Some reservations have already been voiced.[107] This perhaps warrants additional research as well as an international debate to scrutinize the established fact-finding mechanisms and evaluate whether the potential issues can be ironed out.

---

102 Healey et al. (2014, 10–12); Mueller et al. (2019); Eichensehr (2019, 213–217); Tsagourias and Farrell (2020); Shany and Schmitt (2020); Egloff and Wenger (2019); ICT for Peace Foundation (2019); Reaching Critical Will (2021); Charney et al. (2016).

103 Tsagourias and Farrell (2020).

104 People's Republic of China (2020).

105 UNSC (2020, 11).

106 UNSC (2020, 10).

107 Japan, for example, argued that it "has reservations to the idea of establishing a new international mechanism for attribution." Permanent Mission of Japan to the United Nations (2021, 2). Other States seem to have been more receptive to the proposal; Pakistan, for instance, took the position that "[d]eveloping a common approach to attribution in a universal setting under the UN auspices remains the most effective way forward in this regard". UNGA (2021d, 20).

7. Conclusion

Attribution – the process of assigning responsibility for a malicious cyber operation – seeks to deter such operations or to legitimize the imposition of costs for the wrongdoing in order to promote accountability in international relations. Attribution is a process with several aspects – factual, legal and political. This paper reviews these three aspects in the context of the operationalization of the norm of responsible State behaviour aimed at ensuring non-escalatory attribution.[108]

These three aspects of attribution are intertwined: technical or factual investigations underpin the legal analysis and relevant political considerations. Therefore, factual attribution is a crucial element of accurate, non-escalatory attribution that facilitates stability in international relations. Legal examination of the malicious cyber operation establishes the legal responsibility of a State. Attribution in accordance with the rules and principles of the customary international law of State responsibility allows the injured State to employ reaction beyond retorsion. Political aspects of attribution take into account the outcomes of the factual and legal analyses and include decision-making related to the format of attribution and standards of proof.

However, factual attribution remains challenging, which hinders attribution as a normative operation and prevents States from invoking credible State responsibility. It is perhaps because of this that attribution claims "often remain contested"[109] and "lack both transparency and verifiability".[110]

This is not to say that questions remain only related to factual aspects of attribution. In fact, several aspects of attribution could benefit from further research and international discussions. This should include assessing the idea of an international attribution mechanism, which is envisioned to assist States with the evidence collection and analysis elements of the attribution process.

Further clarity is also needed in relation to the legal aspects of attribution. Further research of opinio juris and State practice can contribute to the understanding of the degree of control required by the law of State responsibility to establish a legal nexus between the de facto responsible non-State actor and the de jure responsible State in the context of cyberspace.

Moreover, States should continue discussing and sharing the interpretation of the standards of proof deemed sufficient to attribute a cyber operation to a particular State. Although these standards are subject to political considerations, relevant international legal doctrine can provide guidance.

Finally, further consideration of the role of the attribution claims made by non-State actors would perhaps be in order.

108 UNGA (2015b, para. 13(b)).
109 Egloff (2019).
110 Egloff and Wenger (2019).


35.94
22.34
40.07
49.48
59.11
95.09
97.61
37.12
75.87
97.04
18.48
42.43
39.27
69.77
28.85
13.82
97.61
18.48
39.27
97.61
27.82
98.53
35.94
85.37
98.05
40.07
42.43
22.34
63.66
35.94
13.82
85.37
58.23
54.66
69.77
59.11
85.37
58.23
97.61
43.31

References

Advisory Council on International Affairs, the Netherlands. 2011. *Cyber Warfare*. Advisory Committee on Issues of Public International Law, No 77, AIV/ No. 22, CAVV December.

Australia. 2021. *Australia's International Cyber and Critical Tech Engagement Strategy*.

Baezner, Marie and Patrice Robin. 2017. *Hotspot Analysis: Stuxnet*. Center for Security Studies, ETH Zürich, October. https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Cyber-Reports-2017-04.pdf

Bencsáth, Boldizsár, Gábor Pék, Levente Buttyán & Márk Félegyházi. 2011. *Duqu: A Stuxnet-Like Malware Found in the Wild*. Technical Report, Budapest University of Technology and Economics, Department of Telecommunications, Laboratory of Cryptography and System Security (CrySyS). https://www.crysys.hu/publications/files/bencsathPBF11duqu.pdf

Berridge, Geoff R. 2015. Diplomacy: *Theory and Practice* (5th edition). Basingstoke: Palgrave-Macmillan.

Boebert, William Earl. 2011. 'A Survey of Challenges in Attribution' In National Research Council. 2010. *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy*. Washington, DC: National Academies Press.

Burt, Tom. 2021. 'Russian Cyberattacks Pose Greater Risk to Governments and Other Insights from Our Annual Report'. Microsoft, 7 October. https://blogs.microsoft.com/on-the-issues/2021/10/07/digital-defense-report-2021

Canada. 2019. *Canada's implementation of the 2015 GGE norms*. https://unoda-web.s3.amazonaws.com/wp-content/uploads/2019/11/canada-implementation-2015-gge-norms-nov-16-en.pdf

Case concerning the Air Service Agreement of 27 March 1946 between the United States of America and France. UNRIAA XVIII (9 December 1978) 417.

Cassese, Antonio. 2007. 'The Nicaragua and Tadić Tests Revisited in Light of the ICJ Judgment on Genocide in Bosnia'. *European Journal of International Law* 18(4): 649–668.

Charney, Scott, Erin English, Aaron Kleiner, Nemanja Malisevic, Angela McKay, Jan Neutze and Paul Nicholas. 2016. *From Articulation to Implementation: Enabling progress on cybersecurity norms*. Microsoft, June. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/REVmc8

China. 2020. *China's Contribution to the Initial Pre-Draft of OEWG Report*. https://ceipfiles.s3.amazonaws.com/pdf/CyberNorms/UNGGE/China%E2%80%99s+Contribution+to+the+Initial+Pre-Draft+of+OEWG+Report.pdf

China. 2021. 中国代表团在联合国信息安全开放式工作组: 首次会议关于国际法适用的发言. https://documents.unoda.org/wp-content/uploads/2021/12/Statement-of-China_ICT-OEWG-7th-plenary-meeting_international-law_DEC-16-AM_CHN.pdf

Collins, Craig, and John Packer. 2006. *Options and Techniques for Quiet Diplomacy.* Stockholm: Folke Bernadotte Academy. https://www.corteidh.or.cr/tablas/r31305.pdf

Commonwealth of Australia, Department of Foreign Affairs and Trade. 2021. *Australia's International Cyber and Critical Technology Engagement Strategy*. https://www.international-cybertech.gov.au/sites/default/files/2021-05/21066%20DFAT%20Cyber%20Affairs%20Strategy%202021%20update%20Internals%201%20Acc.pdf

Council of the European Union. 2017a. *Draft Council Conclusions on a Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities ("Cyber Diplomacy Toolbox") – Adoption*. 7 June, 9916/17. https://data.consilium.europa.eu/doc/document/ST-9916-2017-INIT/en/pdf

–––. 2017b. *Draft implementing guidelines for the Framework on a Joint EU Diplomatic Response to Malicious Cyber Activities – approval of the final text*. 9 October, 13007/17. https://data.consilium.europa.eu/doc/document/ST-13007-2017-INIT/en/pdf

Egloff, Florian J. 2019. 'Contested Public Attributions of Cyber Incidents and the Role of Academia.' *Contemporary Security Policy* 41(1): 55–81.

–––. 2020. 'Public Attribution of Cyber Intrusions'. *Journal of Cybersecurity* 6(1): 1–12.

Egloff, Florian J. and Andreas Wenger. 2019. 'Public Attribution of Cyber Incidents'. *CSS Analyses in Security Studies*: ETH Zurich, 244, May. https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse244-EN.pdf

Eichensehr, Kristen E. 2019. 'Decentralized Cyberattack Attribution.' *American Journal of International Law Unbound* 113: 213.

–––. 2020. 'Cyberattack Attribution and International Law'. *Just Security*, 24 July. https://www.justsecurity.org/71640/cyberattack-attribution-and-international-law

European Parliament. 2021. *State of EU cyber defence capabilities*, Resolution, TA(2021)0412, 7 October.

Federal Council. 2018. *National Strategy for The Protection of Switzerland Against Cyber Risks (NCS) 2018-2022*, April, 10. https://www.ncsc.admin.ch/dam/ncsc/en/dokumente/strategie/Nationale_Strategie_Schutz_Schweiz_vor_Cyber-Risiken_NCS_2018-22_EN.pdf.download.pdf/Nationale_Strategie_Schutz_Schweiz_vor_Cyber-Risiken_NCS_2018-22_EN.pdf

Finland. 2020. *International law and cyberspace – Finland's national positions*. 15 October. https://um.fi/documents/35732/0/Cyber+and+international+law%3B+Finland%27s+views.pdf/41404cbb-d300-a3b9-92e4-a7d675d5d585?t=1602758856859

FireEye. 2014. *Digital Bread Crumbs: Focusing Seven Clues To Identifying Who's Behind Advanced Cyber Attacks*. Mandiant, Security Reimagined.

–––. 2018. APT28: *A Window into Russia's Cyber Espionage Operations*. Mandiant, Security Reimagined. https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-apt28.pdf

Fischerkeller, Michael P. 2021. 'Initiative Persistence and the Consequence for Cyber Norms'. *Lawfare*. 8 November. https://www.lawfareblog.com/initiative-persistence-and-consequence-cyber-norms

Foreign and Commonwealth Office. 2019. *Non-Paper on Efforts to Implement Norms of Responsible State Behaviour in Cyberspace, as Agreed in UN Group of Government Expert Reports of 2010, 2013 and 2015*. https://unoda-web.s3.amazonaws.com/wp-content/uploads/2019/09/uk-un-norms-non-paper-oewg-submission-final.pdf

G7. 2019. *Cyber Norm Initiative Synthesis of Lessons Learned and Best Practices*. https://www.diplomatie.gouv.fr/IMG/pdf/_eng_synthesis_cyber_norm_initiative_cle44136e.pdf

Global Commission on the Stability of Cyberspace. 2019. *Advancing Cyberstability*. Final Report, November. https://cyberstability.org/report

Government of the Netherlands. 2019. *Letter to the Parliament on the International Legal Order in Cyberspace*. Appendix: International Law in Cyberspace, 5 July.

Guzman, Andrew T. 2002. 'A Compliance-Based Theory of International Law'. *California Law Review* 90(6): 1823–1887. https://www.jstor.org/stable/pdf/3481436.pdf?refreqid=excelsior%3Ae2e8702b19945bdf2ada22b820ae3970

Hackdig. 2021. *Operation MKLG: Analysis of Suspected Attacks in the Middle East for Years*. 10 March. http://www.hackdig.com/03/hack-293629.htm [unofficial translation]

Healey, Jason, John C. Mallery, Klara Tothova Jordan and Nathaniel V. Youd. 2014. *Confidence-Building Measures in Cyberspace a Multistakeholder Approach for Stability and Security*. Atlantic Council, November. https://www.atlanticcouncil.org/wp-content/uploads/2014/11/Confidence-Building_Measures_in_Cyberspace.pdf

House of Representatives. 2010. *Statement of Ed Giorgio*. Hearing Before the Subcommittee on Technology and Innovation Committee on Science and Technology House of Representatives One Hundred Eleventh Congress Second Session, 15 July. https://www.govinfo.gov/content/pkg/CHRG-111hhrg57603/html/CHRG-111hhrg57603.htm

Huth, Paul K. 1997. 'Reputations and deterrence: A theoretical and empirical assessment.' *Security Studies* 7(1): 72–99.

ICT for Peace Foundation. 2019. *ICT4Peace Submission to the UN Open Ended Working Group (OEWG) on ICT and International Security*. August. https://ict4peace.org/wp-content/uploads/2019/08/ICT4Peace-2019-Submission-UN-Open-Ended-Working-Group.pdf

International Court of Justice (ICJ). 1949. *Corfu Channel case*. Merits, ICJ Reports 1949: 4.

–––. 1981. *Case Concerning United States Diplomatic and Consular Staff in Tehran (United States of America v. Iran)*. Judgement. ICJ Reports 1980: 3.

———. 1986. *Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States of America)*. Merits, Judgment, ICJ Reports 1986: 14.

———. 1997. *Case Concerning the Gabcikovo-Nagymaros Project (Hungary-Slovakia)*. Judgement, ICJ Reports 1997: 7.

International Criminal Tribunal for the former Yugoslavia (ICTY). 1999. *Prosecutor v. Dusko Tadic*. Judgement, IT-94-1-A, 15 July.

International Law Commission (ILC). 2001. *Draft articles on Responsibility of States for Internationally Wrongful Acts, with commentaries*. UN Document A/CN.4/SER.A/2001/Add.1 (Part 2). II(2) Ybk of the ILC 312. https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf

Islamic Republic of Iran. 2021. *Statement by H.E. Mr. Majid Takht Ravanchi Ambassador and Permanent Representative of the Islamic Republic of Iran to the United Nations Before the United Nations Security Council On "Maintenance of international peace and security: Maintaining international peace and security in cyberspace"*. New York, 29 June.

Ivan, Paul. 2019. *Responding to cyberattacks: Prospects for the EU Cyber Diplomacy Toolbox*. Discussion Paper, European Policy Centre, 18 March. http://aei.pitt.edu/97071/1/pub_9081_responding_cyberattacks.pdf

Joyner, Christopher C. 2006. *Coercion*. Max Planck Encyclopedia of International Law, December. https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1749?prd=MPIL#law-9780199231690-e1749-div1-2

Kaspersky. 2017. *From Shamoon to Stonedrill. Wipers Attacking Saudi Organizations and Beyond*. Version 1.05, 3 July. https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07180722/Report_Shamoon_StoneDrill_final.pdf

———. 2021a. 'The power of threat attribution.' *Kaspersky*. https://media.kaspersky.com/en/business-security/enterprise/threat-attribution-engine-whitepaper.pdf

———. 2021b. 'What is IP spoofing?' *Kaspersky*. https://www.kaspersky.com/resource-center/threats/ip-spoofing

———. 2021c. 'What is Spoofing – Definition and Explanation.' *Kaspersky*. https://www.kaspersky.com/resource-center/definitions/spoofing

Kastelic, Andraz. 2020. 'International Cyber Operations: National Doctrines and Capabilities'. International Cyber Operations Research Paper Series No.1. Geneva: United Nations Institute for Disarmament Research. https://unidir.org/sites/default/files/2021-05/International%20Cyber%20Operations%20Series%20-%20Paper%201.pdf

Keromytis, Angelos. 2016. *Enhanced Attribution*. 25 April. https://www.enisa.europa.eu/events/cti-eu-event/cti-eu-event-presentations/enhanced-attribution

Khraisat, Ansam, Iqbal Gondal, Peter Vamplew, & Joarder Kamruzzaman. 2019. 'Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges'. *Cybersecurity* 2(20): 1–22. https://doi.org/10.1186/s42400-019-0038-7

Kingdom of Belgium. 2021. *Written contribution of Ambassador Philippe Kridelka, Permanent Representative High-Level Open Debate on "Maintaining international peace and security in cyberspace"*. New York, 29 June.

Kingdom of the Netherlands. 2021. 'Maintaining International Peace and Security in Cyberspace.' Security Council Open Debate Written Statement by H.E. Ms. Yoka Brandt the Kingdom of the Netherlands to the United Nations New York, 29 June 2021. [on file with the author]

Lauterpacht, Hersch. 1968. *International Law and Human Rights*. Hamden: Archon Books.

Lin, Herbert. 2016. 'Attribution of Malicious Cyber Incidents. From Soup to Nuts'. Stanford University, Hoover Institute. Working Group on National Security, Technology, and Law. Aegis Series Paper No. 1607. https://www.hoover.org/sites/default/files/research/docs/lin_webready.pdf

Mandiant. 2013. APT1 *Exposing One of China's Cyber Espionage Units*. https://www.mandiant.com/media/9941/download

Marczak, Bill, Ali Abdulemam, Noura Al-Jizawi, Siena Anstis, Kristin Berdan, John Scott-Railton, and Ron Deibert. 2021. *From Pearl to Pegasus Bahraini Government Hacks Activists with NSO Group Zero-Click iPhone Exploits*. University of Toronto, The Citizen Lab, Munk School of Global Affairs and Public Policy, 8 November. https://citizenlab.ca/2021/08/bahrain-hacks-activists-with-nso-group-zero-click-iphone-exploits

Ministry of Foreign Affairs of Japan. 2021. *Basic Position of the Government of Japan on International Law Applicable to Cyber Operations*. 28 May. https://www.mofa.go.jp/files/100200935.pdf

Mission of Brazil to the UN New York. 2021. 'Open debate on maintaining international peace and security in cyberspace'. 29 June. [on file with the author]

MITRE. 2021. 'Defense Evasion.' *MITRE ATT&CK*. https://attack.mitre.org/tactics/TA0005

Mueller, Milton, Karl Grindal, Brenden Kuerbis and Farzaneh Badiei. 2019. 'Cyber Attribution: Can a New Institution Achieve Transnational Credibility?' *Cyber Defense Review* 4(1): 107–122.

National Institute of Standards and Technology. n.d. *Tactics, Techniques and Procedures* (TTP). Information Technology Laboratory, Computer Security Resource Center. https://csrc.nist.gov/glossary/term/tactics_techniques_and_procedures

Nicholson, Andrew, Tim Watson, Peter Norris, Alistair Duffy & Roy Isbell. 2012. *A Taxonomy of Technical Attribution Techniques for Cyber Attacks*. 11th European Conference on Information Warfare and Security, ECIW 2012, 188–197.

People's Republic of China. 2020. *Statement by the Delegation of the People's Republic of China on the Issue of the OPCW Investigation and Identification Team*. The Hague, 7 July. https://www.opcw.org/sites/default/files/documents/2020/07/Statement%20by%20China%20on%20the%20Issue%20of%20IIT%20%281%29.pdf

Permanent Court of International Justice (PCIJ). 1923. *Case of Certain questions relating to settlers of German origin in the territory ceded by Germany to Poland*. Advisory opinion, PCIJ Series B, No 6.

Permanent Mission of Japan to the United Nations. 2021. *Statement by Mr. Akahori Takeshi, Ambassador for United Nations Affairs and Cyber Policy of the Ministry of Foreign Affairs of Japan, at the United Nations Security Council Open Debate on Cyber Security*. 29 June. https://www.un.emb-japan.go.jp/itpr_en/akahori062921.html

Reaching Critical Will. 2021. *Women's International League for Peace and Freedom: Responses to the Zero Draft of the Final Report of the UN OEWG on developments in the field of information and telecommunications in the context of international security*. 23 February. https://front.un-arm.org/wp-content/uploads/2021/02/WILPF_zero-draft_23Feb2021.pdf

Romania. 2021. *RO delegation interventions*. OEWG, 18–22 February. https://front.un-arm.org/wp-content/uploads/2021/02/Romania-delegation-interventions-to-OEWG-18-22-feb-2021-final.pdf

Romanosky, Sasha and Benjamin Boudreaux. 2021. 'Private-Sector Attribution of Cyber Incidents: Benefits and Risks to the U.S. Government'. *International Journal of Intelligence and Counterintelligence* 34(3): 463–493.

Seitz, Amanda. 2019. 'FBI Reviewed Cybersecurity Firm's Evidence in 2016 DNC Election'. *AP News*, 26 September. https://apnews.com/article/archive-fact-checking-7657130451

Shakelford, Scott J. 2010. 'State Responsibility for Cyber Attacks: Competing Standards for a Growing Problem'. In: C. Czosseck and K. Podins (Eds.) *Conference on Cyber Conflict Proceedings*, 2010 CCD COE Publications. https://ccdcoe.org/uploads/2018/10/Shackelford-State-Responsibility-for-Cyber-Attacks-Competing-Standards-for-a-Growing-Problem.pdf

Shany, Yuval and Michael N. Schmitt. 2020. 'An International Attribution Mechanism for Hostile Cyber Operations.' *International Law Studies* 96: 196–222.

Skopik, Florian and Timea Pahi. 2020. 'Under False Flag: Using Technical Artifacts for Cyber Attack Attribution.' *Cybersecurity* 3(8): 1–20. https://doi.org/10.1186/s42400-020-00048-4

Stilgherrian. 2018. 'Blaming Russia for NotPetya Was Coordinated Diplomatic Action'. *ZDNet*, 12 April. https://www.zdnet.com/article/blaming-russia-for-notpetya-was-coordinated-diplomatic-action

Teitelbaum, Ruth. 2007. 'Recent Fact-Finding Developments at the International Court of Justice'. *The Law & Practice of International Courts and Tribunals* 6(1): 119–158.

Tran, Delbert. 2018 'The Law of Attribution: Rules for Attributing the Source of a Cyber-Attack'. *Yale Journal of Law and Technology* 20(1): 376–441.

Tsagourias, Nicholas. 2012. Cyber Attacks, Self-Defence and the Problem of Attribution. *Journal of Conflict and Security Law* 17(2): 229–244. https://www.jstor.org/stable/pdf/26296228.pdf?refreqid=excelsior%3Ae7f848186171 9f7668fecdab380eb2ba

Tsagourias, Nicholas and Michael Farrell. 2020. 'Cyber Attribution: Technical and Legal Approaches and Challenges'. *European Journal of International Law* 31(3): 941–967.

Tsimperidis, Ioannis, Cagatay Yucel, and Vasilios Katos. 2021. 'Age and Gender as Cyber Attribution Features in Keystroke Dynamic-Based User Classification Processes'. *Electronics* 10(7): 835. https://doi.org/10.3390/electronics10070835

United Nations General Assembly (UNGA). 2002. *Responsibility of States for Internationally Wrongful Acts*. UN Document A/RES/56/83, 28 January, Annex.

–––. 2010. *Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*. UN Document A/65/201, 30 July 2010.

–––. 2015a. *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*. UN Document A/70/174, 22 July.

–––. 2015b. *Developments in the field of information and telecommunications in the context of international security*. UN Document A/RES/70/237, 30 December.

–––. 2021a. *Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security*. UN Document A/76/135, 14 July.

–––. 2021b. *Official compendium of voluntary national contributions on the subject of how international law applies to the use of information and communications technologies by States submitted by participating governmental experts in the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security established pursuant to General Assembly resolution 73/266*. UN Document A/76/136, 13 July. https://front.un-arm.org/wp-content/uploads/2021/08/A-76-136-EN.pdf

–––. 2021c. *Compendium of Statements in Explanation of Position on the Final Report*. A/AC.290/2021/INF/2, 25 March. https://front.un-arm.org/wp-content/uploads/2021/04/A-AC.290-2021-INF-2.pdf

–––. 2021d. *Chair's Summary: Open-ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security, Third substantive session*. UN Document A/AC.290/2021/CRP.3, 10 March. https://front.un-arm.org/wp-content/uploads/2021/03/Chairs-Summary-A-AC.290-2021-CRP.3-technical-reissue.pdf

–––. 2013. *Report of The Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*. UN Document A/RES/68/243, 27 December.

United Nations Institute for Training and Research (UNITAR). n.d. Public Diplomacy. *What It Is and How to Do It*. https://unitar.org/public-diplomacy-what-it-and-how-do-it

United Nations Office for Disarmament Affairs. 2021. *Remarks by Ms. Izumi Nakamitsu, UN High Representative for Disarmament Affairs at a UN Security Council Open Debate on Cybersecurity*. 29 June.

United Nations Secretary-General (UNSG). 2018. *Securing Our Common Future: An Agenda for Disarmament*. UN Office for Disarmament Affairs, New York. https://www.un.org/disarmament/wp-content/uploads/2018/06/sg-disarmament-agenda-pubs-page.pdf

–––. 2018. *Strategy on New Technologies*, 8–9. https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf

–––. 2021. *Our Common Agenda: Report of the Secretary General.* United Nations, New York, 62. https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf

United Nations Security Council (UNSC). 2020. *The situation in the Middle East*. 8764th meeting, UN Document S/PV.8764, 5 October 2020.

United States, Department of Defense. 2015. *Department of Defense Cyber Strategy*. April. https://www.hsdl.org/?view&did=764848

–––. 2018. *Nuclear Posture Review*. February. https://media.defense.gov/2018/Feb/02/2001872886/-1/-1/1/2018-NUCLEAR-POSTURE-REVIEW-FINAL-REPORT.PDF

United States. 2018. *National Cyber Strategy of the United States of America*. September. https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf

–––. 2020. *Comments on the Chair's Pre-draft of the Report of the UN Open Ended Working Group (OEWG)*. https://ceipfiles.s3.amazonaws.com/pdf/CyberNorms/UNGGE/United+States+Comments+on+the+Chair%E2%80%99s+Pre-Draft+of+the+Report+of+the+UN+Open+Ended+Working+Group+(OEWG).pdf

Wan, Wilfred, Andraz Kastelic & Eleanor Krabill. 2021. 'The Cyber–Nuclear Nexus: Interactions and Risks'. Nuclear Risk Reduction, Friction Points Series No. 2. Geneva: United Nations Institute for Disarmament Research.

Wheeler, David A. and Gregory N. Larsen. 2003. *Techniques for Cyber Attack Attribution*. Institute for Defense Analysis. https://apps.dtic.mil/sti/pdfs/ADA468859.pdf

Zhang, Xiaolu, Oren Upton, Nicole Lang Beebe, Kim-Kwang & Raymond Choo. 2020. 'IoT Botnet Forensics: A Comprehensive Digital Forensic Case Study on Mirai Botnet Servers'. *Forensic Science International: Digital Investigation* 32: 51–60.

Non-Escalatory Attribution of International Cyber Incidents

Facts, International Law and Politics

Attribution – the process of allocating responsibility for a malicious cyber operation – is comprised of three distinct and intertwined aspects: factual or technical, legal and political. This paper analyses these three aspects through the prism of the normative expectations of responsible State behaviour in cyberspace. As a result, the paper makes a number of suggestions of how to consider the challenges of attribution and how to operationalize norm B of the 2015 Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.

UNIDIR UNITED NATIONS INSTITUTE FOR DISARMAMENT RESEARCH

@unidirgeneva @UNIDIR un_disarmresearch