2020 Policy brief

Three tales of attribution in cyberspace

Criminal law, international law and policy debates

Dennis Broeders, Els De Busser and Patryk Pawlak

THE HAGUE PROGRAM for Cyber Norms

EU CYBER DIRECT

Universiteit Leiden

Suggested citation:
Broeders, D., E. De Busser and P. Pawlak. (2020). *Three tales of attribution in cyberspace: Criminal law, international law and policy debates.*
The Hague Program For Cyber Norms Policy Brief. April 2020.

Three tales of attribution in cyberspace: Criminal law, international law and policy debates

Introduction[1]

Attribution can be broadly defined as the process of assigning responsibility for a (malicious) cyber activity to a specific actor on the basis of the available evidence, including all-source intelligence, forensic investigation, and taking into account the political context. Given the sensitive nature of such evidence and the implications that a decision about attribution might have on bilateral relations between the accuser and the accused, states maintain their exclusive right to attribute (or not) a cyber operation based on their own methods, procedures and political interests.

Attribution strengthens the ability of an actor to identify those responsible for malicious activities in cyberspace and potentially hold them accountable. The capacity of a state to attribute is a key element in curtailing impunity in cyberspace and ensuring justice for the victims. But attribution is not a silver bullet and should not be an aim in itself. A decision to attribute a cyber operation to another actor should be strictly linked to a broader policy objective(s) that a state or a group of states wishes to achieve. Depending on the overall goal, the process of attribution embodies several concrete choices and dilemmas concerning the level of certainty for arriving at such decision, the quality of the evidentiary standards, or the concrete instruments available to a state in response to such malicious activities, ranging from issuing a statement to criminal prosecution or imposing restrictive measures.

The process leading to attribution is often lengthy and encompasses several stages with clearly identified thresholds:

- **Suspicion:** any malicious cyber operation that is discovered leads automatically to the question 'who did it?' and ensuing speculations about potential perpetrators. This is how an actor – state or non-state – becomes a suspect. The notion of 'suspicion' and 'suspect', however, have different implications across disciplines. Whereas in criminal law suspicion has to be substantiated by an investigation (i.e. collecting and assessing the evidence) and follow well-established doctrines (e.g. the presumption of innocence until proven guilty), such strict rules are not always applied in international politics where suspicion could rest on solid (albeit sometimes covert) evidence but also on a hunch and personal convictions. The substantiation of suspicion also plays a role in whether or not states decide to act upon their suspicion.

1 Exploring the different approaches to attribution from an international law, a criminal law and a policy perspective, The Hague Program for Cyber Norms and EU Cyber Direct convened 20 international experts in a workshop in The Hague on 24 May 2019. The participating experts were from Europe, Asia, North and South America and had a background in international law, criminal law or policy. This policy brief includes ideas and opinions expressed during this workshop.

- **Accusation:** the shift from suspicion to accusation is the step that has attracted most attention, both in legal and in policy terms. When and on what basis does an actor decide to accuse another actor of a malicious activity – executed or planned – in cyberspace? Contrary to suspicion, an accusation is likely to bring about more scrutiny on both the accuser and accused. While accusation and attribution are frequently treated as the same, we make a clear distinction between these two processes whereby attribution is a step that may lead to accusation. This distinction is important to make. While victims of a cyber-attack (that are technologically capable to do so) will often go through the process to come to an attribution of an attack – if only to improve their own defence mechanisms in the future – not all victims will decide to (publicly) accuse another state.[2] In our view, therefore, most of the discussions about public attribution are de facto discussions about accusations by one actor against another. Especially public attribution opens up the discussion about the quality and transparency of the evidence leading to the accusation.

- **Consequences:** there is a general expectation that accusation leads to concrete actions and responses aimed at enforcing international or national law, often in the form of retorsion, countermeasures, or other types of sanctions foreseen in the (inter)national legal order. Such an approach is very reductionist and ignores other potential benefits that an accusation can bring, such as the right to respond and the obligation to assist in solving the problem. For instance, the attribution of an attack to the *territory* of a state gives the victim state the right to request assistance or compensation from that state. It is not unreasonable to expect that such request should be responded to in good faith and in the spirit of cooperation between states. Should such efforts fail and malicious intent of the other state become clearer, potential punishment could ensue. The debate about consequences is also the one where evidentiary standards matter most as it raises the issues of the legality, legitimacy and proportionality of the response.

Interacting with all these stages is the fact of **disclosure** or revelation. Disclosure can be part of the attribution process as a deliberate decision by the wronged state, but it can also be external: other actors like cyber security companies, private companies under attack, and other states can also disclose the fact that an attack is occurring or has occurred. In some cases it can even be the attacking state that reveals the attacks, often through proxies. The timing of disclosure also affects the options of the wronged state: a disclosure may compel a state to act – even though it perhaps preferred not to. In short: disclosure can be the result of an attribution process or it can externally interact with it.

In the relatively short history of attribution of cyber-attacks, states have used different paths (see Figure 1). While many initiatives hardly ever see daylight and are hidden from the scrutiny of public opinion, some states have also opted for more public forms of attribution through indictments under criminal law and political attributions – albeit with a very limited reference to international law and norms that have been violated. These different legal and political tales however all have their own internal logic and rules. The following sections of this paper aim to shed some light on how these different stages of the attribution process are addressed in the areas of criminal law, international law, and international policy.

2 For an analysis of why states do or do not make attribution public or overt see: Gil Baram and Udi Sommer. (2019). "Covert or not Covert: National Strategies During Cyber Conflict", pp. 197-212 in: T. Minarik, S. Alatalu, S. Biondi, M. Signoretti, I. Toolga and G. Visky (eds.) *11th International Conference on Cyber Conflict: Silent Battle*. Tallinn: CCDCOE.

Fig. 1: Attribution of cyber operations: Four cases

The data used to create this visual can be found in the annex on pages 15-18.

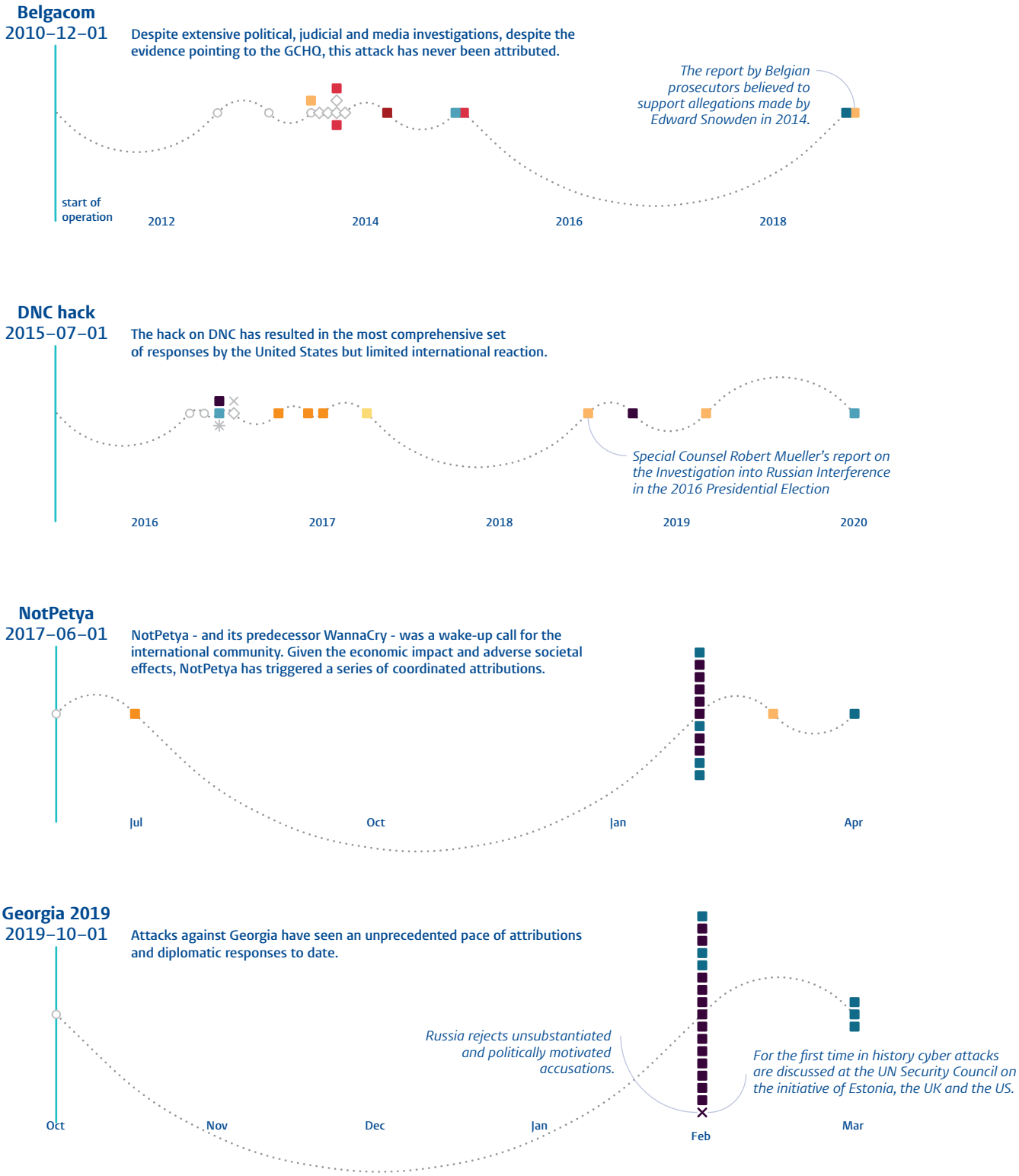
Even though the focus of this paper is on legal, criminal and policy dimension of attribution, this graphic demonstrates the involvement of other actors and tools used to name suspected perpetrators. The cases suggest an increasing willingness of states to call out state sponsored cyber operations.

stage

- Discovery
- Reaction
- Investigation
- Attribution
- Admission

type of attribution

- civil society
- criminal
- diplomatic
- intel
- judicial
- media
- political
- technical
- none



A tale of criminal law

Criminal punishment is considered to be the harshest intrusion into the personal rights of an individual by the state. Coercive measures and sentences, ranging from capital punishment and the deprivation of freedom to the seizing of objects, have a substantial impact on a person's life. For that reason the use of criminal law as a reaction to an unlawful act is curbed in two strongly related ways: a) by means of the basic principle that criminal law should be an *ultima ratio* and b) by building in safeguards protecting the individual(s) involved and ensuring the integrity of the criminal process. These restrictions apply to the use of criminal law in general. We here highlight both the general use of criminal law and the specifics of relying on criminal law for the purpose of attributing unlawful cyber-incidents to a malicious actor. It is essential to first emphasize that a state as such cannot be criminally indicted, so the use of criminal law for the attribution of cyber incidents is exclusively related to acts carried out by individuals, which can be non-state, state-sponsored or even directly employed by a state.

a) Criminal punishment should be used only as an **ultima ratio** or as the last resort: a mechanism that due to its profound impact is activated only when other – less intrusive – mechanisms are inadequate or disappoint. The ultima ratio principle is brought back to the rule of law and the monopoly on the legitimate use of force that lies in the hands of the state. The state is the only actor that can exercise coercive measures and punishment upon its citizens in reaction to unlawful behaviour, although it can delegate some of it to the private sector as has happened for example with botnet takedowns where the state coordinated with the private sector.[3] At the same time, no one is above the law so the way in which coercive measures are conducted by state authorities is restricted by built-in safeguards.

b) The rule of law **protects** the individual against arbitrariness by state authorities from the start of a criminal investigation (even before the individual is informed), throughout the prosecution phase and up to the moment of sentencing. This includes the (cross-border) gathering of evidence, the standard of proof, the admissibility of evidence, fair trial rights and extradition. These rules and principles make up the integrity of the criminal proceedings and are used here to contextualize criminal attribution for cyber incidents.

Attribution of a criminal act to an individual – although this terminology is hardly used in the national criminal law context – is preferably done at an early moment, with the ultimate goal of discovering who is responsible and which individual(s) should be prosecuted. Identifying a first suspect could be called an investigative hypothesis for attribution. This is done at the start of the search for evidence with the purpose of making the search targeted; but attribution is also the process of finding proof to confirm the first suspicion. Attribution at the start of the search for evidence is necessary due to the protection of the individual by a set of **fair trial rights** laid down in the basic human rights instruments such as the European Convention of Human Rights and the EU Charter for Fundamental Rights and Freedoms. The moment of activation of the fair trial rights is the moment of the criminal charge. In other words, from the moment an individual is informed of the criminal charge against him, he can rely on a set of rights safeguarding his position such as the presumption of innocence, the right of access to an independent and impartial tribunal and the rights of defence. These rights

3 Benoit Dupont. (2017). "Bots, cops, and corporations: on the limits of enforcement and the promise of polycentric regulation as a way to control large-scale cybercrime". *Crime, Law and Social Change,* 67 (1): pp. 97-116.

continue to protect the individual for the full length of the criminal proceedings until a decision is made by a final judgment. The **level of certainty** a court should base a judgment on is – due to the mentioned intrusion level of a criminal sentence and the impact it has on an individual's life – high. Most Anglo-Saxon criminal justice systems work with the "beyond a reasonable doubt" threshold, whereas most continental criminal justice systems maintain the "intime conviction" or the reasonable certainty of the judge. The evidence gathered should thus convince the judge or court with reasonable certainty that the individual has committed the criminal act in question. For a criminal attribution the evidence should thus be strong and public. Yet the gathering of evidence for the purpose of a criminal investigation is strongly linked to the availability of **national resources** and is almost exclusively governed by **national laws**. The conditions for obtaining evidence, the admissibility of evidence and the excluding of evidence are all governed by national law. This can be explained by the close connection between a state's (historical, political, cultural, religious, etc.) identity and its criminal justice system. It may therefore result in differences between states in when or how they criminally indict for a similar – or even the same – cyber-incident.

Once evidence is available and admissible, attribution or criminal indictment is unproblematic. The difficulties lie in gathering the evidence in a cross-border – or in the case of a cyber-incident: a global – setting. Traditionally, for the majority of states worldwide the mechanism of **mutual legal assistance** is established for this purpose, whereas for the EU member states the mechanism of mutual recognition is the primary mode of cooperation. However, in the context of cyber-incidents much of the evidence will be of a digital nature. The mentioned cooperation mechanisms are not necessarily adequate for the reliance on **digital evidence**. They lack swiftness in exchanging evidence and are dependent on trust between states. Trust between states is challenging enough when a traditional type of cross-border crime is concerned; it will be an even higher hurdle when relations between states are contentious as is usually the case with cyber-attacks where states are prepared to escalate to the level of a (public) attribution. Although efforts are made to improve this situation – for example the European Commission's e-evidence proposals[4] regulating the obtaining of digital evidence from service providers within EU member states and in third states – focus should be on trying to improve the means and cooperation mechanisms between states on collecting reliable evidence. It should be noted that trust makes the exchange of evidence easier, but attribution tends to come in play when there is a lack of trust.

Accusation in the context of criminal law may also lead to the "**instrumentalisation**" of criminal law for political purposes. For example, the FBI indictments under criminal law pinpoint individuals – and not 'states' – and require evidence that has to stand up in a court of law. Obviously, indictments of individual members of the PLA or the GRU implicate the Chinese and the Russian state respectively and clearly signal political discontent with state behaviour. That may actually be their main purpose, but therein also lies a problem. The use of indictments and criminal law – a trend exemplified by the approach of the United States – should be a goal in itself in terms of law enforcement, instead of instrumentalising criminal law into the service of foreign policy goals. This may happen in cases where a state does not extradite nationals or where no extradition treaty exists between the two states involved. When grounds for refusal of extradition are known in advance or could reasonably be expected, such practice seems to be an intended misuse of criminal attribution. Major actors in international cyber operations such as China, Russia and France have laws in place that forbid

4 Council of the European Union. (2019). "Regulation of the European Parliament and of the Council on European production and preservation orders for electronic evidence in criminal matters – general approach". 10206/19, 11 June 201 9.

the extradition of their own citizens; many other countries as a rule also do not do so in practice. Moreover, when political tensions or human rights concerns stand in the way of extradition to the state that indicted them and the prosecution can therefore never take place, the criminal attribution becomes nearly futile. The purpose of such attribution is not criminal punishment but seems to be political. Additionally, the judicial review of the collected evidence is missing. The instrumentalisation of criminal law is therefore diametrically opposed to the principle of ultima ratio and to the protective mechanisms explained above.

A tale of international law

Like in the case of criminal law, the purpose of attribution in public international law is to avoid impunity and ensure that victims of a cyber-attack or operation – that is deemed either illegal or unfriendly – can exercise their rights as provided for in international law. However, international law is primarily preoccupied with answering whether and which norm of international law has been breached before dwelling into a question of who is responsible. In international law, attribution is a technical step to determine who is responsible for a violation of international law and which appropriate legal framework should be triggered (i.e. ILC Articles on State Responsibility, Law of International Responsibility of International Organizations, or International Criminal Law, if the breached rule triggers not only state responsibility but also individual criminal liability, as in the case of international crimes). In other words, whereas in law enforcement the focus is primarily on establishing responsibility of an individual or an entity, attribution in international law is necessary in order to establish the rights and obligations of states affected by an incident and impose consequences, if appropriate. But establishing facts and circumstances does not automatically imply that there are specific consequences for those violating international law. This is a separate (political) decision.

When it comes to international law, the key question is that of state responsibility and the associated evidentiary standards under international law. There are a few specific challenges to answering this question. Given that malicious cyber activities are perpetrated also by non-state actors who may act as proxies for the state, there is a challenge of establishing a sufficient link between the two. For instance, many of the intelligence reports attribute attacks performed by APT1 to China by claiming that Unit 61398 was sponsored by the government or received its direct support.[5] While this may be sufficient for politicians, only the claims about direct support – as opposed to the claims about 'sponsoring' – would have any legal meaning. At the same time, the existing technology and the loose nature of connections between governments and 'patriotic hackers' or other cyber-groups who may act 'under the instructions' from a government[6] make it very difficult to establish such a link in a convincing manner.[7] With the diffusion of control over the instruments of force in cyberspace, the current interpretation of the principles of international law does not necessarily reflect the complexities of the relationship between state and non-state actors.

5 Mandiant. (2013). "APT1: Exposing One of China's Cyber Espionage Units", February 2013.

6 This is an uncontroversial attribution standard, reflected in Article 8 of the ILC Articles on State Responsibility and confirmed by the ICJ in the Nicaragua Judgment.

7 See for an overview of the (legal) relationship between states and proxies: Tim Maurer. (2016). "'Proxies' and Cyberspace". *Journal of Conflict and Security Law,* 21 (3): pp. 383–403.

The second issue that needs to be addressed is that of the **standard of proof** with clearly stated benchmarks in order to ensure equality before the law. Under what circumstances and on the basis of what criteria do states accuse each other of international law violations? There is a clear distinction between diplomatic complaints and accusations before the Security Council or in another international political forum (e.g. when an action is attributable to a state or the state has failed its due diligence obligations), formal dispute settlement settings before an international arbitration, court or tribunal or unilateral measures adopted by states (e.g. countermeasures adopted in response to a violation of international law). Which of these paths is pursued depends on the state itself, with different evidentiary standards and rules for producing the associated evidence for each (keeping in mind that there are no 'concrete' evidentiary standards in international law, apart from international criminal tribunals).

For instance, **diplomatic complaints** or discussions in the UN Security Council have been sometimes followed by states resorting to **measures of retorsion** (i.e. lawful but unfriendly actions) such as targeted sanctions or expulsion of diplomats. In such cases, under international law, there is no formal obligation to present the evidence in support of such actions: all states are simply allowed to take measures of retorsion. If states decide to apply a formal dispute settlement mechanism, an accusation may require presenting concrete evidence as prescribed by the rules for evidentiary standards of a given dispute settlement body. For instance, Article 38 of the Rules of Court reads that '[t]he application shall specify as far as possible the legal grounds upon which the jurisdiction of the Court is said to be based; it shall also specify the precise nature of the claim, together with a succinct statement of the facts and grounds on which the claim is based'.[8] Outside of the court rooms, international law does not have any specific rules that would answer the question of how much evidence is enough. In its rulings, however, the ICJ has applied the 'clear and convincing standard' approach even though what that means exactly is still debated among legal scholars.[9] Finally, with substantial evidence of a violation of international law in hand, a state may decide to adopt **countermeasures** (i.e. unilateral measures adopted in response to an internationally wrongful act that would otherwise be unlawful under international law). An accusation in such cases needs to be explicit (i.e. the accused state should receive a prior notice and a chance to repair before the application of the countermeasure). The 'clear and convincing evidence' explaining why the injured state thinks it needs to apply a countermeasure for cessation of the conduct and/or reparation might be required by an international tribunal like the ICJ should the case end up before one. It is important to note that such notification does not necessarily need to be public and can be made on a bilateral state-to-state level. According to some states, countermeasures without prior notification may be taken under certain circumstances – for example when notification would undermine the effect of the countermeasures.[10] This is clearly different from the current practice whereby states view public accusation – and shaming the suspected perpetrator – as part of the overall strategy. Such practice, however, makes it difficult for the public to judge whether the adopted countermeasures were indeed legal and justified.

8 Rules of Court (1978) adopted on 14 April 1978 and entered into force on 1 July 1978.

9 See: Marco Roscini. (2015). "Evidentiary Issues in International Disputes Related to State Responsibility for Cyber Operations". *Texas International Law Journal,* 50: pp. 248 ff; Mary Ellen O'Connell. (2006). "Rules of Evidence for the Use of Force in International Law's New Era". *Proceedings of the Annual Meeting (American Society of International Law),* 100: pp. 44-47.

10 The UK and France, for example, hold this position. See for UK: Jeremy Wright. (2018). "Cyber and International Law in the 21st Century". Speech by UK Attorney General Jeremy Wright QC, MP on 23 May 2018. See for France: French Ministry of the Armies. (2019). *International law applied to operations in cyberspace*.

Like in the case of criminal law, the discussion about attribution in international law is meant to provide the ground for reflection about consequences. The 2015 UN GGE report reaffirmed that 'the accusations of organising and implementing wrongful acts brought against States should be substantiated'[11] while the UK Attorney General Jeremy Wright made it clear that 'the victim state must be confident in its attribution of that act to a hostile state before it takes action in response'.[12] However, so far states have been vague in their declarations concerning the level of certainty about attribution required to trigger a response to a cyber-attack and for such a response to be lawful under international law.[13] This is particularly relevant if a state opts for a non-judicial path. When a state resorts to countermeasures, it not only needs to demonstrate another state's wrongdoing but also act in compliance with the **principles of necessity and proportionality**.

In the absence of one universal evidentiary standard for the legitimate use of countermeasures, experts have suggested different solutions. The Tallinn Manual 2.0, for instance, suggests adopting a case-by-case analysis whereby in the absence of a universal standard a state does what any other 'reasonable state' would do in the given circumstances, including the evaluation of rights involved and robustness of response.[14] An alternative solution is to make the requirement of the 'clear and convincing argument' more specific and develop a standard of proof through state practice. Finally, based on the general judicial practice some international lawyers suggest that approaches such as 'balance of probability standard' or 'reasonable standard rule' could provide a valid solution to the dilemmas associated with evidentiary standards.[15]

Despite the availability of extensive legal reasoning on how to apply international law to cyberspace, including the process of attribution and evidentiary standards, the tale of international law is a sad one. To date, no state has attributed a malicious cyber activity by providing a clear reference to a rule of international law that would allegedly have been violated. Quite to the contrary, states remain cautious in making such declarations and resort to broad statements of condemnation. Even in the most recent case of the cyber-attacks against Georgia in 2019, states have used opaque references to 'undermining' Georgia's sovereignty as a political term rather than the international law concept (see Table 1).[16] Ironically, this makes the application of international law and ending impunity in cyberspace more difficult. Furthermore, the absence of references to international law in the existing accusations also diminishes the value of international law as an instrument aimed at preventing conflict in cyberspace. Or, formulated more positively, if states do attribute with reference to international law, it will contribute to 'general practice accepted as law' that may contribute to the formation of customary international law.

---

11 See 2015 UN GGE consensus report: United Nations General Assembly. (2015). "Group of Governmental Experts on the Developments in the Field of Information and Telecommunications in the Context of International Security". A/70/174, 22 July 2015. Also: Sputnik News. (2015). "UN Cybersecurity Report Compromises on Self-Defense Issue – Russian Official". *Sputnik News,* 17 August 2015.

12 Jeremy Wright. (2018). "Cyber and International Law in the 21st Century". Speech by UK Attorney General Jeremy Wright QC, MP on 23 May 2018.

13 Przemysław Roguski. (2020). *Application of International Law to Cyber Operations: A Comparative Analysis of States' Views*. The Hague Program for Cyber Norms Policy Brief. March 2020.

14 Michael N. Schmitt. (2017, ed.). *Tallinn Manual 2.0 on the international law applicable to cyber operations.* Cambridge: Cambridge University Press, pp. 115-116.

15 For instance, the balance of probability standard requires a judge to rule in favour of the accused party if the accuser refuses to present the proof and reserves for them the judgment whether something occurred or not on the basis of probability (e.g. it is more likely that a state that has conducted malicious activities and demonstrated intent to harm another state is indeed behind an attack rather than a victim of a false flag operation that is difficult to mount). See: HHJ Stephen Davies. (2009). "Proof on the balance of probabilities: what this means in practice". *Thomson Reuters Practical Law,* 22 October 2009.

16 See for an analysis of these attributions: Przemysław Roguski. (2020). "Russian Cyber Attacks Against Georgia, Public Attributions and Sovereignty in Cyberspace". *Just Security,* 6 March 2020.

Table 1. Selection of statements issued in relation to the cyber-attacks against Georgia in 2019

Country/organisation	Statement
Australia	Attribution of malicious cyber activity in Georgia by Russian Military Intelligence
Canada	Canada condemns Russia's malicious cyber-activity targeting Georgia
Czechia	Statement on Twitter
Denmark	Statement on Twitter
Estonia	Statement of the Foreign Minister of the Republic of Estonia Urmas Reinsalu
European Union	Declaration by the High Representative on behalf of the European Union - call to promote and conduct responsible behaviour in cyberspace
Lithuania	Statement on Twitter
Netherlands	The Netherlands considers Russia's GRU responsible for cyber attacks against Georgia
New Zealand	New Zealand condemns malicious cyber activity against Georgia
Norway	Statement on Twitter
Poland	Statement of the Polish MFA on cyberattacks against Georgia
Ukraine	Comment of the MFA of Ukraine on cyberattacks committed by the Russian Federation against Georgia
United States	The United States Condemns Russian Cyber Attack Against the Country of Georgia
United Kingdom	UK condemns Russia's GRU over Georgia cyber-attacks

A tale of policy

Policy makers look at attribution through a different lens than legal scholars and legal professionals do. Ideally, the legal profession looks at a cyber-incident and, if its fits within the parameters of international or criminal law, follows the procedural guidelines that the law provides. Policy makers firstly define an incident in terms of a political problem and then see what route may be helpful to deal with that problem. International law and criminal law are merely *among* the options. More importantly, while the nature of cyber operations might be considered the same from the international or criminal law perspective – i.e. it does not matter who the perpetrator is, law is the same for everyone – in the policy tale cyber operations are not necessarily equal. Many are politically ignored, some are addressed behind closed doors and only a few are addressed publicly. The fact that formal, public attributions of cyber operations are a relatively new phenomenon suggests that policy makers have only recently defined them as so problematic that they should be addressed publicly. Moreover, most of the recent public attributions have been championed by the United States and/or its western allies and are built on the rationale that trying to negotiate rules, norms and even laws for responsible state behaviour in cyberspace is pointless if there are no consequences to bad behaviour. There is a fear that *not* confronting malicious cyber-attacks will create a norm that tacitly allows (all) peace time cyber operations.

Attribution is about both sense-making (i.e. establishing what has happened) and meaning-making, i.e. communication with the outside world about the act and attribution in order to have a political effect. *Public* attribution of cyber operations is a meaning-making process that can serve different strategic goals. One goal is 'shaping the operational space'. This covers the signalling of what is and is not allowed in the digital operational space as mentioned above. This is the realm of diplomacy and legal opinion. Another is about creating a 'counter threat strategy', meaning the strategic use of attribution to counter certain specific adversaries (and not others). This is the realm of punishment, of indictments, building attribution coalitions and burning toolsets.[17]

Given the political-strategic nature of public attribution for policy makers, the role of evidence in the policy process differs from legal proceedings. In law the evidence has to meet high standards – especially in criminal law – whereas in politics the standard seems to be more about plausibility and winning the hearts and minds of the audiences that open up with public attribution. Efrony and Shany[18] showed that most public attributions do not reference (violations) of international law at all or, if they do, only in the most general terms and without citing any specific legal principles or prohibitions. There may be two reasons for that. In the first place, it is often not needed as the message to both adversary and allies is foremost a political one. Secondly, and related, the application of international law in cyberspace is simultaneously generally accepted in principle – see the 2013 and 2015 UN GGE consensus reports – as well as highly contested when it comes to specifics.

As the decision to attribute cyber-attacks to state (actors), or to *not* do so, is often more political than legal, there are many differences between states that determine their room and will to manoeuvre on the issue.

- Attribution requires certain technical capabilities to detect and determine a cyber-attack and its perpetrator. These abilities are not evenly distributed around the world, making attribution a possibility for some states but not for others. If we add in the vital role of high end (signals) intelligence capabilities, the possibilities for attribution narrow even further.
- The role of intelligence agencies also complicates the increasing popularity of 'attribution coalitions' in which groups of countries collectively point the finger. Given the fact that there are circles of trust in the western intelligence world (5 eyes, 9 eyes), there will be differences in access to the underlying evidence of an attribution between countries.[19]
- Geopolitics also plays a role. While it is relatively unproblematic for the United States and Europe to call out the activities of, for example, China, Russia and North Korea, the (expected) consequences for neighbouring countries of these states may be enough to refrain from public attribution.
- Consequences sometimes play a different role for countries that are able to attribute or join an attribution coalition, but choose not to. If geopolitical or diplomatic reasons stand in the way of imposing consequences, some countries view an attribution as meaningless or even damaging.

The link between attribution and consequences is an important one. If the aim is change of (bad) state behaviour, then attribution on its own is unlikely to be enough in many cases.

17 Florian J. Egloff. (2019). "The Politics of Public Attribution", Introductory remarks presented at "The dilemmas of attribution in cyberspace: between politics and law" workshop, The Hague, 24 May 2019.

18 Dan Efrony & Yuval Shany. (2018). "A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice". *American Journal of International Law,* 112 (4): pp. 583-657.

19 Dennis Broeders, Sergei Boeke and Ilina Georgieva. (2019). *Foreign intelligence in the digital age. Navigating a state of 'unpeace'.* The Hague Program For Cyber Norms Policy Brief. September 2019.

Naming and shaming are policy instruments in itself but may not deter top tier cyber actors. Again, the relationship between evidence and international law comes to the fore when we look at the consequences that have been imposed on states in relation to attributed cyber-attacks. So far, the consequences mirror the fact that public attributions have not indicated (which) international law has been violated. Therefore, the consequences have not exceeded the level of 'retorsion' which, though considered an unfriendly act, are allowed to states under international law without any real burden of evidence.[20] An open question is whether measures of retorsion – as opposed to 'higher level' reactions such as counter measures or self-defence – actually have any discouraging value. Can they actually shape and support norms of responsible state behaviour?

More worryingly, attribution may have different outcomes and perceived uses depending on the actors and the audiences we look at. Given that attribution is about signalling, there are at least three possible signals:

- **Signalling a norm for appropriate behaviour.** This is the intended signal that attributing states are after. By calling out a transgression, they confirm the norm.
- **Signalling strength.** Some states may welcome being called out because it signals their status as a strong – albeit rogue – actor in cyberspace. Giles and Hartmann suggest that an actor like Russia may actually be appreciative of the publicity that comes with attribution.[21]
- **Signalling impotence.** By calling out a transgression, but not acting on it, states signal that they are unwilling or unable to correct bad behaviour. For some countries this is a reason to not take part in the attribution game.

Like anything in international politics, actions may have foreseen and unforeseen effects. Although it is difficult to say at this moment, when the number of attributions is still rather limited and recent, this is something that needs to be considered in the longer run.

The consequences under criminal law are dependent on an actual trial being held which in many, if not most, cases is unlikely to happen once the indictments have been unsealed.[22] For political purposes an actual trial is not necessary: the court of public opinion is what matters most – both when it comes to judging the response by the state-victim and the initial malicious activity by the perpetrator. If naming and shaming is the foreign policy goal, the indictment may be enough from a policy perspective. However, from the perspective of the (integrity of) criminal law this instrumental use of indictments looks very different. Lastly, the EU has recently embarked on a different route. By putting into place the EU Cyber Sanctions regime, the Union has created a formal instrument that may impose consequences when it is used.[23] If a natural or a legal person gets sanctioned under the new regime the EU in principle opens itself up to legal challenges in the European Court of Justice, which will require substantial evidence to support the charges and sanctions. For now, it remains to be seen if and how the EU27 will use the new instrument.

20 François Delerue. (2019). *International law in cyberspace matters: this is how and why*. Ideas in Focus. Brussels: EU Cyber Direct, May 2019.

21 Keir Giles and Kim Hartmann. (2019). ""Silent Battle" Goes Loud: Entering a New Era of State-Avowed Cyber Conflict, p. 32 in: T. Minarik, S. Alatalu, S. Biondi, M. Signoretti, I. Toolga and G. Visky (eds.) *11th International Conference on Cyber Conflict: Silent Battle.* Tallinn: CCDCOE.

22 Russell Buchan. (2018). *Cyber Espionage and International Law.* London: Hart Publishing, pp. 25-26.

23 See for instance: Patryk Pawlak. (2019). "Guardian of the galaxy: EU cyber sanctions and norms in cyberspace". *Chaillot Paper* 155, October 2019.

Tales of the unexpected: Dilemmas and challenges of attribution

The fact that states can navigate between the three different pathways of criminal law, international law and political attribution – with the possibility of creating shortcuts and connections along the way – results in a number of dilemmas and challenges. As things stand, attribution can be based on different legal grounds – or none at all. Attribution may require different levels and quality of evidence and procedural safeguards depending on the chosen legal framework. Attribution can be covert or public and depending on circumstances, such as geopolitical considerations, and may just expose the act instead of both the act and the actor. But criminal indictments of individuals are not – at least formally – attribution to a state, even when the 'private person' on the FBI poster is wearing a PLA uniform. Similarly, in the EU's view, imposing targeted sanctions against an individual is not considered a formal attribution. Attribution is indeed what states make of it.[24]

Such ambivalent attitude towards attribution has distinct advantages. It gives states a degree of flexibility and manoeuvrability and allows them to engage with adversaries in the cyber domain on a case by case basis. It prevents states from becoming predictable and avoids them getting locked into fixed procedures that do not take all of the relevant context into consideration. It also has distinct disadvantages. Whereas the political process – which is dominant – works well with flexibility, legal processes sit less comfortably with some of the ambiguities in the emerging practice of attribution. Here standards of proof, procedural guidelines and requirements, and in some cases the letter and the spirit of the law are at odds with what makes perfect sense politically. Five dilemmas and challenges with regard to attribution, state responsibility, standards of proof and (un)foreseen consequences of attribution stand out.

1. Finding the best route

The first dilemma is about what 'the best route' is to pursue attribution. The route of criminal law is possible and sends a clear political message but may end up instrumentalising criminal law as there is no real expectation that the accused will actually stand trial in a court of law. The political process offers much flexibility and has lower evidentiary requirements but does not offer much bite in terms of consequences. Most measures adopted by the states to date are at the level of retorsion which brought limited results in terms of warding off a determined adversary. As a matter of fact, states have resorted to retorsions by default since they do not require presenting any concrete evidence, allowing the states to remain relatively vague on the norms of international law they think have been violated. Alternatively, calling out violations of concrete rules or principles of international law could open up the possibility to use countermeasures as a tool to impose consequences but this approach remains largely theoretical. Consequences would be more severe and would be framed in the language of international law, which in itself would also be politically significant. However, states are hesitant to open up Pandora's box of international law, not in the least as it may also limit their own possibilities to conduct cyber operations. Given state practice, attribution and international law is currently a theoretical, academic debate, or in keeping with the title of this policy brief, it is still largely a tale of science fiction.

24 Thomas Rid and Ben Buchanan. (2015). "Attributing Cyber Attacks". *Journal of Strategic Studies,* 38 (1-2): pp. 4-37

2. Living with ambiguity

The second dilemma follows from the previous one. The fact that attribution is a sovereign political decision does not necessarily mean that states have no responsibility to clarify the international legal framework and develop some sort of standard for evidentiary requirements. To a certain extent this is a pastiche on the more general principle that top tier states prefer strategic ambiguity – as it allows room to maneuver – while small states usually prefer a rules-based order that makes the world more predictable. It is both a short-term and long-term dilemma. As long as there are few top tier states their strategic advantage is great and ambiguity is beneficial, but as their ranks begin to swell rules would also start to benefit the larger players. Given that cyber capabilities have a steep learning curve for dedicated actors, 'cyber years' are relatively short, leading some scholars to urge states to stop shunning international law and the related question of evidence: 'Evidentiary issues have legal underpinnings, and the U.S., U.K., and French efforts to block the development of customary international law on attribution are short-sighted'.[25]

3. Accusing responsibly

In contrast to criminal proceedings, the attributions and consequences seen in the international public domain – political or legal – are not subject to established procedures regarding accusation, evidence, prosecution, trial and redress. As states have been reluctant to go down the path of international law, favouring the political process, evidentiary standards are relatively weak. The UK's position, for instance, is that 'there is no legal obligation requiring a state to publicly disclose the underlying information on which its decision to attribute hostile activity is based, or to publicly attribute hostile cyber activity that it has suffered in all circumstances'.[26] In contrast, the procedural integrity of criminal law raises the bar for the accuser but also bestows rights on the accused, for example through fair trial rights. As such, it embodies the rule of law by placing a check on the state when it exercises its monopoly on the legitimate use of force. The accused has rights and there are procedures to challenge the accusation and have possible wrongs righted in a trial. While states do not enjoy similar protection, there is a merit in asking about potential consequences of mistaken attribution that served as the foundations for actions of a state-victim against the suspected state-perpetrator. In that sense, advancing the discussion on the standard of evidence could serve to protect against insufficiently justified and unfair accusations. This would in addition serve as a potential de-escalatory measure in relations between the states.

25 Kristen Eichensehr. (2020). "The Law & Politics of Cyberattack Attribution". *UCLA Law Review,* 67. Forthcoming, available at SSRN.

26 Jeremy Wright. (2018). "Cyber and International Law in the 21st Century". Speech by UK Attorney General Jeremy Wright QC, MP on 23 May 2018.

4. The trouble with state responsibility

The principle of state responsibility is well-established in international law and some states have also taken a stand on the burden and the standard of proof that is required for attribution of cyber operations. Yet, a comparative analysis of states' view on the applicability of international law to cyber operations by Roguski[27] shows that it is unclear what these states mean when they make an argument for acting "reasonable" under the circumstances when gathering the information on the basis of which they draw conclusions. Clarification is needed on whether 'reasonableness' refers to a *duty of care* that supposedly rests on a state making an attribution[28] or whether it refers to the *standard of evidence* a state has to rely upon when determining a certain fact such as the identity of an attacker.[29] Roguski points out that the duty of reasonable care is known in international law with regard to primary rules – such as the duty to prevent transboundary harm – and cannot be easily projected onto secondary rules, such as those of state responsibility and attribution. Finding a standard for state responsibility under International law is already difficult, but matters get even more complicated when the involvement of non-state actors and proxies are concerned. Non-state actors operate in all sectors and the ILC Articles on state responsibility have standards of attribution for the conduct of non-state actors. However, establishing sufficient evidence that proxies act under control or instruction of state is especially difficult in cyberspace operations. The nature and strength of the link between a state and a proxy actor (the threshold of control) can point towards state-sponsored actors but may require developing specific evidentiary standards and the technical means to reliably make an attribution to proxy actors.

5. Living with consequences

The final dilemma relates to the intended and unintended consequences of attribution. States attribute cyber-attacks, and impose consequences, because they want to push back against bad actors and increase national and international cyber security. Naming, shaming and relatively low-level sanctions are unlikely to deter some of the most active and disruptive actors in cyberspace. It is not at all certain that attribution and consequences (at the level of retorsion) make any change to their calculations. It is even possible that attribution produces adverse effects. It may signal impotence to some actors, i.e. instead of strength it signals an unwillingness or incapacity to impose real consequences. Also, some countries may even see attribution as a badge of honour or as a welcome way of advocating their cyber capacities to the world. The question then is 'are we sending the right message?' and the answer to that is perhaps more dependent on the eye of the beholder(s) than on the sender of the message.

27 Przemysław Roguski. (2020). *Application of International Law to Cyber Operations: A Comparative Analysis of States' Views*. The Hague Program for Cyber Norms Policy Brief. March 2020.

28 For example the American position that a state is obliged to "act reasonably under the circumstances when they gather information and draw conclusions based on that information" seems to point to a duty of care. See: Brian J. Egan. (2016). "Remarks on International Law and Stability in Cyberspace". Remarks by Brian J. Egan, Legal Adviser to the U.S. Department of State, on 10 November 2016.

29 For example a standard of proof leaving no room for reasonable doubt, as was referred to by the ICJ in the Case concerning the Application of the Convention on the Prevention and Punishment of the Crime of Genocide, (Bosnia-Herzegovina v. Serbia and Montenegro), Judgment of 26 Feb. 2007, ICJ Rep. 2007, 129, para. 208-209.

Annex: Attribution of Cyber Operations – Four Cases

Cyber attack on Democratic National Committee (DNC) hack

Stage	Attribution type	Date	Facts and actors
Start of operation		Jul 2015	
Operations		Mar 2016	Fancy Bear sends phishing emails
Operations		Mar 2016	WikiLeaks publishes Clinton's emails
Operations		Apr 2016	DNC hack
Discovery		Apr 2016	DNC staff discovers the hack
Discovery		May 2016	Crowdstrike investigation
Attribution	technical	Jun 2016	Crowdstrike report
Attribution	political	Jun 2016	Democratic Party
Admission		Jun 2016	Guccifer 2.0
Operations		Jul 2016	DNCLeaks
Investigation		Jul 2016	FBI report
Reaction		Jul 2016	Russia
Attribution	intelligence	Oct 2016	DHS
Attribution	intelligence	Dec 2016	FBI, CIA
Attribution	intelligence	Jan 2017	CIA, FBI, NSA
Attribution	judicial	Apr 2017	Civil lawsuit by Democratic Party
Attribution	criminal	Jul 2018	Muller report
Attribution	political	Oct 2018	United Kingdom statement
Attribution	criminal	Mar 2019	Advocate General
Attribution	technical	Jan 2020	Crowdstrike

Cyber attack on Belgacom

Stage	Attribution type	Date	Facts and actors
Start of operation		Dec 2010	
Discovery		Jul 2012	Belgacom staff
Discovery		Jan 2013	Microsoft investigation
Discovery		Jun 2013	Fox-IT investigation
Investigation		Jul 2013	Federal Public Prosecutor
Investigation		Aug 2013	Information to Minister of Justice
Investigation		Sep 2013	Cabinet briefing
Investigation		Sep 2013	Investigation by the Privacy Commission
Attribution	media	Sep 2013	De Standaard reporting
Attribution	media	Sep 2013	Der Spiegel reporting
Investigation		Oct 2013	European Parliament hearing
Attribution	criminal	Jun 2013	Belgian Federal Prosecutor
Attribution	civil society	Mar 2014	Report by Edward Snowden
Attribution	technical	Nov 2014	Symantec report
Attribution	media	Dec 2014	De Standaard, The Intercept reporting
Attribution	diplomatic	Sep 2018	Belgian Ministry of Foreign Affairs
Attribution	criminal	Oct 2018	Belgian Federal Prosecutor

NotPetya malware

Stage	Attribution type	Date	Facts and actors
Start of operation		Jun 2017	
Discovery		Jun 2017	Kaspersky Lab
Attribution	intel	Jul 2017	Ukrainian Security Service
Attribution	political	Feb 2018	United Kingdom
Attribution	political	Feb 2018	United States
Attribution	diplomatic	Feb 2018	New Zealand
Attribution	political	Feb 2018	Australia
Attribution	political	Feb 2018	Canada
Attribution	political	Feb 2018	Denmark
Attribution	political	Feb 2018	Lithuania
Attribution	political	Feb 2018	Estonia
Attribution	diplomatic	Feb 2018	Norway
Attribution	diplomatic	Feb 2018	Sweden
Attribution	diplomatic	Feb 2018	Finland
Attribution	criminal	Feb 2018	US Department of Justice
Attribution	diplomatic	Apr 2018	European Union

Cyber attack on Georgia 2019

Stage	Attribution type	Date	Facts and actors
Start of operation		Oct 2019	
Discovery		Oct 2019	
Attribution	political	Feb 2020	Georgia
Attribution	political	Feb 2020	United States
Attribution	political	Feb 2020	United Kingdom
Attribution	political	Feb 2020	Lithuania
Attribution	political	Feb 2020	Netherlands
Attribution	political	Feb 2020	Poland
Attribution	political	Feb 2020	Estonia
Attribution	diplomatic	Feb 2020	Denmark
Attribution	political	Feb 2020	Poland
Attribution	diplomatic	Feb 2020	Norway
Attribution	political	Feb 2020	Czechia
Attribution	political	Feb 2020	Australia
Attribution	political	Feb 2020	Canada
Attribution	political	Feb 2020	New Zealand
Attribution	political	Feb 2020	Ukraine
Attribution	diplomatic	Feb 2020	European Union
Reaction	political	Feb 2020	Russia
Attribution	diplomatic	Mar 2020	Estonia, United Kingdom and United States in the UN Security Council
Attribution	diplomatic	Mar 2020	Georgia at the UN
Attribution	diplomatic	Mar 2020	Belgium at the UN

Authors

Dennis Broeders is Associate Professor of Security and Technology and Senior Fellow of The Hague Program for Cyber Norms at the Institute of Security and Global Affairs at Leiden University. Prior to joining Leiden University, he was a Senior Research Fellow at the Netherlands Scientific Council for Government Policy and Professor of Technology and Society at Erasmus University Rotterdam.

Els De Busser is Assistant Professor Cybersecurity Governance at the Institute of Security and Global Affairs. She is Educational Director of the Executive Master Cyber Security and teaches in several other programmes at Leiden University. Els is also a researcher in The Hague Program for Cyber Norms. She often publishes on issues related to data protection, cyber security and criminal law cooperation and is a member of the Standing Committee of Experts on International Immigration, Refugee and Criminal Law (Meijers Committee).

Patryk Pawlak is the EUISS Brussels Executive Officer. In this capacity, he maintains and develops relations with other Brussels-based institutions. In addition, he is in charge of the cyber portfolio, leading the Institute's cyber-related projects and contributing to its outreach activities. He currently coordinates an EU-funded project 'EU Cyber Direct'. Since June 2017, he is a Co-Chair of the Advisory Board of the Global Forum on Cyber Expertise.

Acknowledgements

This policy brief is based on the discussions in two expert workshops entitled 'The dilemmas of attribution: between policy and law' held in May 2019 in the Hague and September 2019 in Brussels. The authors would like to thank the participants for sharing their ideas and insights. They are grateful to Florian Egloff, Marco Roscini, François Delerue and Tatiana Tropina for their comments on an earlier draft. They would like to thank Corianne Oosterbaan, Nikolay Bozhkov and Fabio Barbero for the support with the organisation of the workshops. Finally, they would like to thank Corianne Oosterbaan for the editorial assistance and Christian Dietrich for designing the visual representation of various attributions of cyber operations.

All of the activities and publications of The Hague Program for Cyber Norms are supported by a grant of the Dutch Ministry of Foreign Affairs. The activities of the EU Cyber Direct project are funded by the European Union.

Contact information

The Hague Program for Cyber Norms
E-mail: info@thehaguecybernorms.nl
Website: https://www.thehaguecybernorms.nl
@HagueCyberNorms

Address
The Hague Program for Cyber Norms
Faculty of Governance and Global Affairs
Leiden University
Hague Campus
Turfmarkt 99
2511 DP The Hague
The Netherlands

EU Cyber Direct
Website: https://www.eucyberdirect.eu
@EUCyberDirect

Address
EU Institute for Security Studies
100, Avenue de Suffren
75015 Paris
France

Colofon
Published April 2020.
Visual timeline of attribution of cyber operations by Christian Dietrich/EU ISS.

Graphic design: www.pauloram.nl

THE HAGUE
PROGRAM
for Cyber Norms

EU
CYBER
DIRECT

Universiteit
Leiden