‘Unpacking’ technical attribution and challenges for ensuring stability in cyberspace

Submission to 2021–2025 UN Open-Ended Working Group (OEWG) on security of and in the use of information and communications technologies

Authors: Anastasiya Kazakova (Kaspersky), Ivan Kwiatkowski (Kaspersky), Julia Ryng (LSE IDEAS), Kenddrick Chan (LSE IDEAS)

May 2022

Contents

Introduction

When reports of a cyberattack appear in the headlines, questions abound regarding who launched it and why. Even if an attacker has what are to it perfectly rational reasons for conducting such an attack, these reasons are often known only to them. The rest of the world, including the victims of

the attack, must often engage in some degree of speculation to explain the events and devise ways to protect themselves accordingly. Knowing the technical aspects of an attack may allow victims to build stronger defences, patch gaps and increase their cyber-resilience. This is why both policymakers and industry leaders are usually eager to have this knowledge as a possible 'cure' to mitigate or prevent such cyberattacks from happening again.

A constant challenge in such an endeavour is that the cyber context, in all its complexity and interconnectedness, remains a dark, unknown forest for many decision-makers. How then can they find out who was behind an attack and why?

Attribution of a cyberattack is not 'magic'. It is a complex process where *technical, legal* and *political* discussions[1] intertwine to produce as complete a narrative as possible – with as many plausible answers as possible (though not always comprehensive ones). *Technical* attribution relates to a technical investigation to identify who was behind a cyberattack or cyber operation. *Legal* attribution assesses if there has been a breach of international law. Finally, *political* attribution implies the political decision to publicly or privately announce those assessments and tie them to a particular state or private actor.

> ***Technical* attribution** relates to a technical investigation to identify who was behind a cyberattack or cyber operation.
>
> ***Legal* attribution** assesses if there has been a breach of international law.
>
> ***Political* attribution** implies the political decision to publicly or privately announce those assessments and tie them to a particular state or private actor.

Security researchers and private cybersecurity companies can typically analyse cyber incidents from a technical standpoint and cluster them into groups, which they then tie to particular threat actors. However, the only actors that deliver the *entire* narrative of a cyberattack – discussing accountability and international law – are nation states.[2] The decision-making of states is highly complex in nature, often involving multiple considerations – from domestic issues to foreign affairs. Publicly attributing cyberattacks, meaning announcing who is thought to be responsible,[3] is therefore often not straightforward, and states might not always be willing to make such announcements.[4]

Within this piece, we, a collection of policy scholars and industry experts, discuss how technical attribution – identifying *who* is behind a cyberattack – can become more transparent and better understood by the wider public. Our key discussion points include: How is technical attribution carried out? What are the key challenges in conducting reliable technical attribution? And finally, how can this be more accessible to the multitude of stakeholders who are operating in cyberspace and/or have interests there?

Below are our reflections on these questions, divided into several parts. Firstly, we discuss *technical* attribution and options to make it more transparent and accessible; next we reflect on how, given existing limitations within multilateral initiatives, the international community might make incremental improvements towards ensuring that *technical* attribution is made more transparent and more accessible and ensure the stability and security of cyberspace.

---

[1] https://unidir.org/publication/non-escalatory-attribution-international-cyber-incidents
[2] https://front.un-arm.org/wp-content/uploads/2021/08/A-76-136-EN.pdf
[3] https://www.tandfonline.com/doi/full/10.1080/01402390.2021.1895117
[4] E.g., Estonia has expressed that "attribution remains a national political decision based on technical and legal considerations regarding a certain cyber incident or operation. Attribution will be conducted on a case-by-case basis, and various sources as well as the wider political, security and economic context can be considered". https://front.un-arm.org/wp-content/uploads/2021/08/A-76-136-EN.pdf

Why would anyone want to know details of technical attribution?

> *"Attribution capabilities are a crucial element in building a deterrence strategy and in deterring malicious behavior."*

Cyber attribution is a necessary step to accountability in cyberspace.[5] It serves as a basis for a response: for states in accordance with international law, and for the private sector (if it owns or manages attacked infrastructure) in accordance with national applicable laws. Cyber attribution is a necessary precursor to retaliation – technical, legal or political. And if a state has advanced attribution capabilities, it is in a better position to understand what has happened and appropriately react to a cyberattack. Attribution capabilities are therefore a crucial element in building a deterrence strategy against malicious behaviour.

Besides states and security researchers, why would anyone else be interested in conducting technical attribution and finding out all possible details?

Given the multistakeholder nature of cyberspace, a state's sovereign decision on cyber attribution – whether public or private – may have far-reaching consequences for other stakeholders. The increasingly relevant discussions, from scholars, academia and civil society,[6] further signal interest in greater transparency of and accessibility to, at least, technical attribution.

> *"Given the multistakeholder nature of cyberspace, a state's sovereign decision on cyber attribution – whether public or private – may have far-reaching consequences for other stakeholders."*

These decisions impact geopolitical realities, and in this regard the more information other decision-makers (e.g., owners of ICT infrastructure that procure tools and services for critical functions and sectors) have, the better they are informed about these geopolitical realities to make decisions that would, in turn, impact users of such ICT infrastructure.

In addition, accessibility to information about technical attribution provides third parties with the ability to assess the results of technical analysis and investigation. Third parties may spot gaps and inconsistencies in the evidence presented and thus help increase the credibility and quality of technical attribution.

In practice, some States[7] have clearly communicated that there is no obligation under international law to disclose underlying evidence of attribution. Scholars also highlight the significant security risks that public (technical) attribution brings and thus argue that "*public attribution is not always better*."[8] Nonetheless, the idea of an international attribution mechanism has been floated by several experts and organizations, proposing an independent mechanism for impartial analysis and decision-making in cyber attribution to complement and assist in states'

[5] E.g., Germany has expressed that "Attributing a cyber incident is of critical importance as a part of holding States responsible for wrongful behavior and for documenting norm violations in cyberspace." https://front.un-arm.org/wp-content/uploads/2021/08/A-76-136-EN.pdf

[6] E.g., submissions from some multistakeholders to the 2019-2021 UN OEWG highlight the need for a "multistakeholder approach which engages all relevant stakeholders to build strong, impartial and verifiable verification mechanisms that build trust and confidence" (https://front.un-arm.org/wp-content/uploads/2020/04/cs-coordination-perspectives-on-oewg-pre-draft.pdf) and support to "support multistakeholder, independent and coordinated attribution efforts" (https://front.un-arm.org/wp-content/uploads/2020/04/oewg-pre-draft-gpd-response-final.pdf).

[7] https://front.un-arm.org/wp-content/uploads/2021/08/A-76-136-EN.pdf

[8] https://www.tandfonline.com/doi/full/10.1080/01402390.2021.1895117

sovereign decisions.[9] [10] While states' reservations regarding calls for greater transparency in cyber attribution and the risks it may carry for strategic competition are valid and do have merit,[11] the benefits of making technical attribution more accessible and better understood by the wider international community should also be explored.

Following on from the previously discussed key thought that cyber attribution as states' sovereign prerogative has wider impacts on non-state actors, we hope that more transparent and accessible technical attribution would be important for states to develop better and fact-based assessments, and thus contribute to greater stability and predictability in cyberspace. After discussing how technical attribution is conducted and current difficulties and limitations with it, we reflect on suggestions for the continuing negotiations in the United Nations Open-Ended Working Group on Developments in Information and Telecommunications in the Context of International Security (UN OEWG) and for the international community broadly.

How the pie gets made: steps in conducting technical attribution

While the idea of cyber attribution – i.e., determining who is responsible for an attack – is generally understandable by anyone, its technical underpinnings usually rely on domain-specific knowledge. In almost all cases, with the exception of attribution provided through human intelligence (or HUMINT), it is based on careful analysis of available technical information. The end result of this process is *technical* attribution: intelligence that informs readers about the identity of the attackers. But in this specific context, 'identity' should not be understood in the traditional sense: this type of attribution is not aimed at pointing to a door that law enforcement can then kick down. This would in fact require leaps that are usually impossible based on the information available. Instead, the objects crucial to the process of technical attribution are threat actors and attack campaigns. Such 'objects', as referred to here, point to things such as malware and hijacked servers, which, when put together and 'manipulated', inform the technical attribution process. The process produces clusters that represent malicious cyber activity, which can be grouped together based on identifiable characteristics of the attack and previous attacks.

Technical attribution may lead the conclusion that, e.g., "APT 41 is responsible for this attack", where APT 41 is a term used to consistently designate a specific attacker in the context of different incidents. Figuring out who the people behind APT 41 are would be the prerogative of *political,* rather than *technical,* attribution.[12] Technical attribution is only concerned with understanding what characterizes APT 41 as an attacker, and which cyberattacks they are responsible for. But how does this process take place?

To understand this, one must first understand what information is used in the technical attribution process and where such information comes from. Taking control of an IT system implies a lot of interaction with it: using vulnerabilities to acquire privileges on the machine, deploying programs to exert control over it, and so on. All these operations affect the target computer or device in very specific ways: files are created, log entries are written, network traffic is generated, and so on. Despite the best efforts of the attackers to leave the smallest possible footprint, it is almost impossible to erase all traces of an attack. The main reason for this is that some of the technical

[9] Mueller, M. et al, (2019) 'Cyber Attribution: Can a New Institution Achieve Transnational Credibility?', *The Cyber Defense Review*, 4(1): 107-122; https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse244-EN.pdf.

[10] https://ict4peace.org/wp-content/uploads/2019/08/ICT4Peace-2019-Submission-UN-Open-Ended-Working-Group.pdf; https://front.un-arm.org/wp-content/uploads/2021/02/WILPF_zero-draft_23Feb2021.pdf.

[11] E.g., through signaling their own capabilities and technical advances to adversaries who could use this as an additional advantage.

[12] https://www.fbi.gov/wanted/cyber/apt-41-group

information generated during the interaction is not stored in the compromised machine, or even in the target network (i.e., within network activity logs collected by the Internet Service Provider (ISP), etc.). Below are some examples of the type of data collected and strategies of collection and analysis during the technical attribution process.

Tooling

Most readers will be familiar with 'backdoors', 'Trojans' and the like – computer programs used to send arbitrary commands to a compromised device, but where do these tools come from? Some of them can be purchased as commercial products, others are open-source and freely available. Some of these tools have even been created by threat actors themselves.

In the latter case, discovering the same unique malware family in two separate cyber-incidents is a strong indication that they share the same perpetrator. A significant part of the work that cyberthreat intelligence teams perform is meticulously indexing known and unknown attack software, and keeping track of which entities use it.

Infrastructure

The tools described in the previous paragraph and deployed in victims' networks do not function in a vacuum. The stolen data needs to be exfiltrated somewhere, and the backdoor needs a place (like a dead-drop)[13] where it obtains the commands to execute. Servers and domain names purchased online serve this function and are collectively seen as the 'infrastructure'.

When the same server is used simultaneously in the context of two apparently uncorrelated attacks, conventional wisdom suggests that whoever owns it must therefore be responsible for them both. Technical attribution can take the investigation process further, often by noting attacker habits – such as which resellers they favour, the specific way they configure their machines, and so on. All of this makes it possible to tie together incidents that do not involve the exact same servers (i.e., servers with different IP addresses), or sometimes even discover the infrastructure of a given threat actor before it is used in an operation.

Attacker procedures

Finally, it is sometimes possible to obtain a clear picture of what the attackers do once they are inside a network: this encompasses the deployment of additional offensive tools and utilities, but also the commands they type. Due to the number of members of attack groups, some of them have to put strict and repeatable procedures in place – allowing for various operators to fill in for one another. Identifying such recurrent patterns, such as a sequence of information gathered in a specific order from a new machine, can allow defenders to recognise a threat actor across multiple incidents.

Beyond this, other aspects that can hint at the identity of attacks include more trivial elements – such as which encryption algorithm, network protocol or attack vector an attacker favours.

The above strategies form a key aspect of the technical attribution process: they transcend individual incidents and aim at building knowledge that can be useful within a larger context. However, no vendor or intelligence service has full visibility over all cyber incidents taking place: defenders only know about what is going on inside their network; incident responders are only aware of the incidents they are asked to remediate; security vendors only obtain limited telemetry from their customer base.

It follows that no single entity can be successful at attribution alone: it is only by sharing information about threat actors (i.e., through whitepapers, conferences and blog posts) that the

[13] https://www.wired.com/story/what-is-dead-drop/

industry's knowledge has allowed us to keep track of the hundreds of threat actors identified over the years.

What are the difficulties, uncertainties and limitations of technical attribution?

Even assuming that it is fully available to those performing cyber attribution, technical information is limited, and does not always allow for robust answers. The obvious blind spot is that private cybersecurity vendors often have no investigative powers. This precludes everyone in the industry from 'following the money'. In the case of attacker infrastructure, servers and domain names are not obtained for free: one way or another, threat actors must find a way to pay for them. Traces generated at this stage are simply unavailable at the technical level.

Tool-based attribution (i.e., grouping together attacks that leverage the same unique malware families) has also been getting more difficult over the years – for two reasons. Firstly, a number of attackers have eschewed homemade backdoors to solely rely on open-source software. These publicly available backdoors can be used by anyone and cannot characterize a single threat actor unless some significant and unique modifications have been made to them. Secondly, Kaspersky has also observed the tendency to share tools and procedures between close yet distinct threat actors, e.g., instance hacker groups from the same region, working for the same sponsor, but going after different target verticals (e.g., the education, energy, or fintech sectors).

Adding to the uncertainty of technical attribution are two important issues that need mentioning. The first is that a number of attackers are acutely aware of the technical attribution process and will therefore attempt to impede it by misleading analysts. The issue is further compounded by the fact that attackers interact with each other and may steal tools from one another – how would defenders then be able to distinguish between the original owner and the copycat? Threat actors may also try to purposefully leave behind 'false flags' – indicators that incriminate other groups.[14] Such false flags may only be discovered through very careful analysis and are likely happening more frequently than the industry is aware of.

The second issue is that the technical information analysts work with is very ambiguous. High-profile victims may be compromised by several attackers at once: each of them deploying their tools and generating a muddled footprint. This means that those performing technical attribution are unable to clarify whether it is one, two, or even more distinct groups that are responsible for the activity they are investigating. The risk that a tool would be attributed to the wrong group always exists, with the implication of poisoning the global knowledge-well for years.

What are the obstacles to a transparent technical attribution process?

The lack of a global database and the inherent ambiguity of the attribution process imply that cooperation and verifiable procedures are necessary to bolster existing technical attribution efforts. But arguments against global information sharing and transparency in this sphere suggest that a call for such procedures is not a straightforward solution.

One key objection in this regard stems from the possibility of attackers gaining access to such a global knowledge pool of information. Attribution reports contain precise indicators (e.g., file hashes, IP addresses, domain names, and so on), which would allow attackers to see which incidents led to their discovery and how defenders tied the activity to them. The immediate

[14] https://securelist.com/the-devils-in-the-rich-header/84348/

consequence of transparent attribution is that it provides resources for attackers to tap into to further hone their methodologies and cover up the most characteristic aspects of their operations. There is therefore arguably value in protecting methods used to track threat actors from interception, even at the expense of wider cooperation.

A further argument against global information sharing and transparency in *technical* attribution processes is that disclosing how an attack is attributed also provides information about the capabilities of the defender. Such information might involve trade/industry secrets or even sensitive and/or classified information. Government actors use their signals intelligence (SIGINT) capabilities to provide invaluable data for the cyber attribution process (including *political* and *legal*) and would rather not attribute an attack publicly at all than have to disclose their SIGINT capabilities and the extent of their visibility. In a limited number of instances, covertly discovering who the attacker is allows analysts to engage in a post-discovery reconstruction of an alternate trail of technical data. By covert means, we refer to signals intelligence, illegal wiretapping and sometimes even plain hacking. But this process – called 'parallel construction' – cannot always be performed, sometimes leaving only a choice between unsubstantiated attribution or no attribution at all.

Paths forward?

Norm 13 (b) of the UN GGE report

"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."

There is clearly a cat-and-mouse aspect to *technical* attribution. As attackers update their methodologies in order to avoid blame, defenders look for new sources of information to help them produce intelligence. The first observation is that tool-based attribution is on the decline. The global availability of sophisticated and free cyber offensive programmes means that attackers will not need to create their own anymore. On the other hand, new capabilities are offered to defenders in response. For instance, there are now offerings for private SIGINT capabilities, where defenders can purchase raw network data collected from the whole world, allowing them to see operators connecting to their attack servers.[15]

Given all the limits discussed above, what could be a way forward for *technical* attribution that is both more transparent and more accessible?

Building upon existing discussions already taking place at various multilateral fora, such as the first iteration of the UN OEWG or the UN Group of Governmental Experts on advancing responsible State behaviour in cyberspace in the context of international security (UN GGE), this paper suggests some areas through which technical attribution can be made more transparent and accessible. These generally focus on issues pertaining to norm implementation (i.e., norm 13(b) of the UN GGE report concerning cyber attribution)[16] and creating more clarification and guidance for policymakers. Both areas are outlined below:

[15] https://team-cymru.com/
[16] https://reachingcriticalwill.org/images/documents/Disarmament-fora/other/icts/gge/documents/gge-report-adopted.pdf

Building consensus amongst states regarding (technical) attribution

The lack of consensus amongst states regarding the necessity of *technical* attribution and its associated processes remains to be addressed. For example, regarding Norm 13(b), the UN GGE report has noted that “attribution is a complex undertaking" and that “a broad range of factors should be considered before establishing the source of an ICT incident”.[17] While this acknowledges the complexity of attribution (including technical) that practitioners have raised, it leaves various questions unanswered. Given that attribution is complex and involves many factors, what is the "agreed-upon baseline" for such *technical* attribution to occur?

Based on documents submitted by states (to both the UN OEWG and UN GGE), it appears that the international community is divided along the lines of whether providing the technical details of ICT incidents (i.e., part of technical attribution) ought to be made compulsory. Some states, such as Russia, have called for legally formalising the need for technical attribution.[18] Others, such as China,[19] while highlighting that states should “demonstrate genuine, reliable and adequate proof” when attributing ICT incidents, hold back from making the provision of evidence a mandatory requirement – note the distinction between use of the term “states ***should***...” and “states ***must***...”.

Additionally, paragraph 24 of the GGE report highlighted the need for states who have suffered cyberattacks to "include the incident’s technical attributes...including the incident’s bearing on international peace and security; and the results of consultations between the States concerned". However, there remains much potential for future discussions on the topic (such as at the second iteration of the UN OEWG) to take the discussion further to specify or outline what such technical data collection actually entails, as well as the processes for sharing information and consulting with other concerned states.

Fostering mechanisms for multistakeholder cooperation at the regional and international levels

While the UN GGE (July 2021) aimed to promote "cooperative measures to address existing and potential threats" in the ICT sphere, support for mechanisms that promote such cooperation remains lacking. Although states recognise the importance of information-sharing and the value that exchanging best practices could bring,[20] it concedes that considerations on how cooperation regarding attribution can actually occur will have to be addressed in future discussions.

Additionally, the call for increased regional and international cooperation is limited to national-level representatives such as Computer Emergency Response Teams (CERTs) or Computer Security Incident Response Teams (CSIRTs) and national ICT authorities. Considering that private sector partners (e.g., cybersecurity vendors) have a significant amount of cybersecurity expertise, the involvement of such private sector partners in international cooperation efforts could significantly assist the international community in gaining more insight into cyberattacks, and aid in attribution processes. It must be noted that expanding the ‘playing field’ to bring in private sector entities would entail significant national security considerations. Successfully addressing this issue would allow private sector expertise to be utilised and further elevate the technical attribution capabilities that states currently possess. Identifying the channels and mechanisms for private

[17] Ibid, paragraph 22.
[18] https://front.un-arm.org/wp-content/uploads/2021/02/RF-Revised-consensus-aimed-OEWG-draft-report-ENG.pdf
[19] https://front.un-arm.org/wp-content/uploads/2021/02/Chinas-Contribution-on-the-Zero-Draft-of-the-OEWG-Substantive-Report.pdf
[20] UN GGE Report 2021 (n 17), paragraph 27 and 28.

sector involvement seems therefore critical for future discussions on building trusted and verifiable technical attribution.

As demonstrated, there is a clear need for greater communication, transparency, and accessibility to information amongst states, while paying some degree of consideration to national security concerns. However, this is not the first time the international community has faced such transnational problems, which require capacity and expertise-building, as well as cooperation among actors from both the public sector and private industry, in order to solve them. As outlined in the Annex, examples from such disparate areas as piracy, nuclear non-proliferation, and space offer us lessons as to how technical cooperation can operate globally. Information-sharing, a key barrier to international cooperation due to national security concerns, can be effectively implemented once the necessary supporting structures (i.e., agreed-upon definitions and norms) are in place. Potential cooperation can also initially take the form of 'minilaterals' or technical ad-hoc groups, where good practices are first shared across a small number of states before being scaled up. Private sector expertise can also be shared with national-level agencies via an array of carefully crafted private-public partnerships (PPP).

Lastly, capacity building to help the multitude of stakeholders as well as states (with less cyber capacities) to learn complexities of technical attribution should be another critical element in ongoing international efforts. Examples of this could be security training sessions, roundtable discussions (e.g., such as those organized by UNIDIR),[21] gamified virtual exercises (e.g., the *Cyber Stability Games* developed by Kaspersky with the support of DiploFoundation),[22] amongst others.

Conclusion

It is hoped this piece has shed more light on the nuances and caveats of technical attribution, which could be used for further research and analysis by other actors. As no one cybersecurity vendor or any other actor in cyberspace has comprehensive visibility into the threat landscape, closer cooperation among security researchers and cybersecurity companies is necessary for building fact and evidence-based technical attribution as well as public research. Greater dialogue between security researchers, diplomats, and academia is necessary to avoid their 'worlds' existing in silos. Furthermore, technical attribution and its nuances need to be better understood and more accessible for both diplomatic negotiations within the bodies like the UN First Committee (which is responsible for dealing with disarmament and international security matters) and evidence-based academic research (which may also inform the former).

An international attribution mechanism could be a solution to greater transparency in, and accessibility of technical attribution in an ideal world. However, the likelihood of this being set up in the near future remains relatively low. The lack of political will of states to tie themselves to formal legal obligations in cyberspace means that an effective information-sharing mechanism resembling that which exists in the Somali piracy context is highly unlikely, at least for the near-term. The UN and International Atomic Energy Agency's nuclear information-sharing mechanisms further point to the institutional limits of any such international body. A more feasible alternative would be the building of technical ad-hoc groups, or various mini-lateral groupings, following the examples of from the nuclear and space policy realm. Such groups, represented by a diverse security research community and academics, could serve as a technical consultative tool for intergovernmental negotiations taking place within various international fora, such as the UN First Committee. Leadership efforts of a few states, coupled with a global recognition of the danger the

[21] https://unidir.org/events/pol-tech-legal-aspects
[22] https://www.diplomacy.edu/blog/whos-behind-a-cyberattack/

lack of information sharing mechanisms creates, is therefore urgently required for any such group to be effectively set up.

Annex: Lessons from past global information-sharing initiatives

Information-sharing on Piracy in Somalia

Although it exists in a totally different domain, the case of piracy off the coast of Somalia can give us insights into how information-sharing to tackle a common threat might actually occur.[23] A series of UN Security Council (UNSC) Resolutions have bolstered several informal information-sharing mechanisms, aiming to aid in the direct enforcement of international law and the prosecution of piracy crime.[24] The Contact Group on Piracy off the Coast of Somalia (CGPCS) is one such mechanism. An international forum bringing together more than 60 States and international organisations, the CGPCS meets in plenary sessions and various issue-based working groups to share data and enforce coordination strategies. Piracy-related information-sharing mechanisms have been praised as instrumental in lowering rates of Somali piracy over the past two decades.[25] Why then has this area proven so fertile for functional and effective information-sharing regimes and how might this measure up in the case of technically attributing cyberattacks?

First and foremost, the piracy context enjoys a well-established customary legal practice in international law. The nature of the crime means it is carried out in 'international waters', removing jurisdictional conflicts, giving any State the right to seize and penalise pirate ships in high seas. Secondly, the information that is shared among States and organisations for prosecution purposes rarely relies on data protected under the umbrella of 'national security'. This is not to say that counter-piracy information-sharing mechanisms do not face obstacles. Investigators and prosecutors use similar techniques to cyber attributors of 'following the money' and mapping data on group activities and group characteristics. However, unlike in cyber attribution cases, piracy prosecutions centre on relatively unambiguous sets of perpetrators (i.e., Somali pirates), a shared public venue for apprehension activities (i.e., international waters), and less sensitive data required to prosecute piracy (e.g., GPS-based location data, photographs of attacks on vessels). Drawing upon such criteria, technical attribution would therefore require a mechanism to unambiguously identify the sets of perpetrators (i.e., cyberattackers/attack groups), a shared venue that is clearly outlined (i.e., public vs private cyberspace), and data that is both valuable yet falling short of the 'classified' threshold. All the above need to be established within the international 'cyber environment' via clear and widely-accepted cyber norms. Their relative absence is therefore indicative of the fact that the success of CGPCS and other piracy-related information sharing mechanisms may be difficult to replicate in the cyber context.

Nuclear non-proliferation and space cooperation as possible PPP models

Nuclear non-proliferation, or nuclear weapons disarmament, is another issue of international concern where, like in cyber attribution cases, information sharing is recognisably important yet swarmed with political and security apprehension. The widely ratified Treaty on the Non-Proliferation of Nuclear Weapons ('NPT') governs the international efforts to prevent the spread of nuclear weapons and to promote cooperation in the peaceful uses of nuclear energy. Article IV of the Treaty specifically states that parties undertake to facilitate and have the right to participate in the fullest possible exchange of information, with the International Atomic Energy Agency ('IAEA') being entrusted with key verification responsibilities under the NPT. Additional Protocols to the IAEA's Statute have, over the years, improved the Agency's ability to verify the integrity of information provided by states regarding their nuclear activities. Replicating this system in the

[23] McLaughlin, R. and Paige, T. (2015) 'The Role of Information sharing in Counter Piracy in the Horn of Africa Region: A Model for Transnational Enforcement Operations', *Journal of International Law and International Relations,* 12(1).
[24] http://www.piracylegalforum.org/about/background-2/
[25] http://www.allaboutshipping.co.uk/wp-content/uploads/2014/07/2014-Q2-IMB-Piracy-Report-ABRIDGED.pdf

cyber context would be difficult primarily because of the lack of a treaty that comprehensively regulates state behaviour in cyberspace.

Indeed, there is little unified political will for any such international agreement in the foreseeable future.[26] Despite the limited powers that the IAEA has over sovereign states, it nonetheless has the authority to conduct inspections, gather data and share information because signatory state parties (to the NPT and IAEA Statue) have willingly given up some of their sovereign rights for these purposes. The existence of the NPT, as a formal source for state obligations, establishes expectations that states can be held to, and provides any mechanisms stemming from it with a degree of authority and political weight. Furthermore, this makes ad-hoc and informal mechanisms in the nuclear context easier to establish and find global support for. The International Partnership for Nuclear Disarmament Verification (IPNDV)[27] is one example of a public-private partnership that brings global actors together to identify and solve technical challenges in monitoring and verifying nuclear disarmament that formal state agreements are not equipped to solve. The fact that states have legal obligations to participate in information sharing means that research opportunities, funding and solutions to information protection issues are also more likely.

The realm of nuclear non-proliferation, where a comprehensive treaty and a slew of associated organisations and bodies support it, is unlike the cyber domain where the quantity of agreements is lower and less comprehensive in terms of issue-area coverage. Yet it is also worth pointing out that the lack of an international treaty does not preclude actors from working together. Initiatives undertaken by just a few states (termed 'minilaterals') can lead to the development of good practices, which can be scaled up and tweaked to accommodate additional members. An example of such an initiative is the Space Situational Awareness ('SSA)' Sharing Program set up by the US Air Force Space Command in recognition that space situational awareness is critical to avoiding unintentional collisions and detecting and attributing attacks to space assets.[28] Initially, the Program suffered from severe asymmetries of information among the interested parties, with the US Air Force having access to an internal catalogue with detailed information on all tracked objects, while the publicly accessible catalogue contained only basic information on a subset of space assets. Such an approach, justified through national security concerns, showed its limitations in 2009, when a commercial communication satellite and a defunct Russian Cosmos satellite collided without advanced warning to the commercial operators. Through series of multistakeholder agreements in 2019 between the US Strategic Command, 19 states, two international organisations, and more than 77 commercial satellite owners, operators and launchers, data that is of a significantly higher-quality has begun to be shared in a more systematic manner between all parties. Such an outcome can perhaps offer us some insight, not just to the benefits of private-public partnerships (PPP), but how such PPPs can benefit all actors that operate within the realm of both space and/or cybersecurity. The increased frequency and impact of cyberattacks targeted at governmental infrastructure over the past years[29] has to some extent pushed the international community to explore such coordinated responses. Whether or not these events will have a sufficient impact for a coordinated effort like with the SSA remains to be seen.

[26] https://globalinitiative.net/analysis/un-cybercrime-treaty-debate/
[27] https://www.ipndv.org/
[28] https://spp.fas.org/military/commission/executive_summary.pdf
[29] https://www.weforum.org/agenda/2021/10/protecting-critical-infrastructure-from-cyber-pandemic/

Authors

Ivan Kwiatkowski

Ivan Kwiatkowski is a Senior Security Researcher in the Global Research and Analysis Team (GReAT) at Kaspersky. He is also an OSCP and OSCE-certified penetration tester and malware analyst, and delivers Kaspersky's reverse-engineering training in Europe.

Anastasiya Kazakova

Anastasiya Kazakova is Senior Public Affairs Manager at Kaspersky where she coordinates the company's worldwide public policy with a focus on cyber diplomacy/global projects and Kaspersky's Global Transparency Initiative.

Julia Ryng

Julia Ryng is the Project Coordinator of the Digital International Relations project at LSE IDEAS. She is also an incoming doctoral candidate at University College London.

Kenddrick Chan

Kenddrick Chan is the Deputy Head of the Digital International Relations Project at LSE IDEAS. He has authored/co-authored several reports and papers on the digital aspects of international relations.