WHODUNIT IN CYBERSPACE:
THE ROCKY ROAD FROM ATTRIBUTION
TO ACCOUNTABILITY

December 2023
Background Paper No. 18

BY ANDREAS KUEHN, DEBRA DECKER, AND KATHRYN RAUHUT
INTRODUCTION

Attribution describes the highly complex process of investigating, identifying, and publicly disclosing the threat actors responsible for a cyber operation or attack that disrupted, denied, degraded, destroyed, or manipulated computers and networks or led to the theft or extortion of data. The malicious actors in cyberspace include states, state-sponsored hackers, and criminal groups, all with varying motivations and overlapping capabilities. States are increasingly working with private sector cybersecurity experts to identify these actors and then hold them to account. Thus, the problem of attribution involves technical, legal, and political assessments. Attribution is seen as a critical element to strengthen accountability in cyberspace. Going forward, governments should be transparent in their attribution processes, providing clear evidence and justifications for attributing cyber operations. This transparency can build trust and legitimacy in the eyes of the international community and enhance prospects for greater accountability.

Cybersecurity firms do the technical forensic work to identify how and what customer vulnerabilities were exploited in an incident, but some also attribute cyber operations to help their customers develop better defenses, and in some cases obtain compensation from their insurers. States go further—and can perform a forensic investigation but also use their intelligence sources to understand an incident. They help critical infrastructure operators and others focus on strengthening their security posture and resilience. In addition, states can prosecute criminal actors as well as respond to state actions through sanctions, diplomatic measures, and other tools of statecraft. In many cases, states may choose not to publicly attribute or pursue legal cases in malicious cyber operations conducted by states or state-supported actors.

States view public attribution as a political tool through a prism of geopolitical considerations. Public disclosure is a strategic choice. There is no absolute measure of success in making attribution public, as it depends on the specific goals the attributing state aims to achieve by publicly disclosing or concealing the origin of and responsibility for a cyberattack. State attribution with a high level of confidence is a precondition to effectively hold attackers accountable for their malicious actions, either through criminal indictments, sanctions, or other measures. Investigators assess so-called indicators of compromise (IOCs) to help determine the source and origin of a malicious action within the larger context and history of malicious actors’ patterns. The analysis supports the technical, operational, and strategic levels of an attribution investigation. However, in most cases, attackers are not caught red-handed and real-time observations of network activities may not be available. Thus, the collected IOCs indicate, ideally with a high degree of confidence, how a cyber incident unfolded; what tactics, techniques, and procedures (TTPs) attackers deployed; and how they match to known advanced persistent threat (APT) actors. This helps support attribution claims to establish the linkage between a threat actor and a state’s political or military leadership.

Some cybersecurity firms, scholars, and think tanks have called for more cooperative approaches, including establishing levels of transparency and standards for “evidentiary processes” in attribution. Indeed, new approaches and mechanisms are needed to strengthen attribution and eventually accountability. This is especially true today as growing geopolitical tensions and conflicts play out prominently in cyberspace while at the same time the use of emerging technologies, such as artificial intelligence (AI), have the potential to elevate risk beyond cyberspace and affect international security, posing new, yet unsolved challenges to cyber attribution.

The practice of attributing responsibility for malicious acts to nation-states, such as physical, kinetic military attacks, is well established in international relations. Determining attribution of cyber operations is more complex due to the inherently anonymous nature of the digital environment and the techniques that are employed by the attackers to conceal their identity. Cyberspace’s architecture with billions of interconnected systems and devices led to calls in the early days to reengineer the Internet for better and faster identification and geolocation of malicious actors. Today’s uncertainty is less technical in nature but is more concerned with establishing the links between perpetrators and the government they work for. But also competing views on cyberspace governance and rules of the road have led to sharp rejections of and tensions over public attribution. Liberal democracies support an open, free, secure, and safe cyberspace and are confronting others, especially autocratic regimes, such as China and Russia, who pursue a domineering approach to state-sovereignty over data and government control of cyberspace.

While attribution may be recognized by like-minded and co-attributing states, accused states tend to reject such accusations as shared, mutually agreed evidentiary standards do not exist. For instance, both China and Russia have refuted blame for malicious actions originating within their borders. In response to NATO’s Hafnium attribution to China, a spokesperson for the Chinese Foreign Ministry stated, “The so-called technical details released by the U.S. side do not constitute a complete chain of evidence. In fact, the United States is the world’s largest source of cyber attacks.” In a similar manner, Russian president Vladimir Putin shirked responsibility for Russian activities: “We have been accused of all kinds of things: election interference, cyber attacks and so on and so forth. And not once, not one time did they bother to produce any kind of evidence or proof.” While rejections by China and Russia may not surprise anyone, evidence in attribution decisions is generally not available to the public. Lack of transparency and evidence combined with low confidence levels are valid grounds for criticism, so attribution declarations must reflect high standards of rigor, transparency, and evidence. Tackling this challenge of building accountability through attribution, requires, first, a brief review of the historical context for attribution followed by an analysis of public attribution as a political tool of states, then a rundown on building accountability through attribution, requires, first, a brief review of the historical context for attribution, followed by an analysis of public attribution as a political tool of states, then a rundown on the procedure of attributing an operation through investigation and political deliberation, and finally an assessment of policy ideas on how to move the needle on accountability.

A BRIEF HISTORY OF CYBER ATTRIBUTION

Malicious cyber incidents are ubiquitous, while public attributions of incidents are few and far between. What explains the low numbers? Governments have to make principled determinations of their equities in what instances and when they make a public attribution. Their objectives may be better served, for instance, in responding covertly and not risking the disclosure of sensitive sources and methods that could be revealed as a consequence of public attribution.

Early cases of attribution were responses to cyber espionage campaigns like Moonlight Maze (Russia, starting in 1996) and Titan Rain (China, starting in the early 2000s) targeting U.S. government systems as foreign actors started to develop sophisticated cyber capabilities in the mid-to-late 1990s. The growing nation-state cyber capabilities, especially by Russia and China, necessitated capabilities to manage attribution -- public disclosure or joint responses was not a chief concern at the time. The 2010 Stuxnet cyber operation was significant in that it degraded physical infrastructure, a uranium enrichment operation, located at Iran’s Natanz facility. It is one of a few instances that has been attributed to, but not formally acknowledged by, the United States and Israel, and is believed to be part of a joint operation to derail Iran’s build-up of nuclear capabilities. In most cases China, Russia, North Korea, and Iran are considered to be responsible for most malicious cyber activities and thus are often the subject of state attribution by Western governments.

More recently, the United Kingdom, the United States, and others, as well as firms, attributed the

---

WannaCry ransomware that impacted almost 150 organizations in 2017 to North Korea.\(^{11}\) With the Colonial Pipeline incident in 2021, ransomware became a national security priority as President Biden of the United States mentioned the incident in his remarks following the issuance of the Executive Order on Improving the Nation’s Cybersecurity (EO 14028).\(^{12}\) Since 2018, public attributions by states have become increasingly frequent and coordinated among democratic, like-minded states. The EU Cyber Diplomacy Toolbox, adopted in 2017, facilitates joint diplomatic responses, including sanctions, to malicious cyber activities based on member states’ attribution decision.\(^{13}\) In 2021, a large collective of states including the United States, the European Union, the United Kingdom, Australia, Canada, Japan, and New Zealand, as well as NATO, publicly attributed the exploitation of Microsoft Exchange Server vulnerabilities to conduct cyber espionage to Chinese state actors.\(^{14}\) Earlier in the same year, the United States imposed sanctions on the Russian government for the SolarWinds supply chain attack.\(^{15}\) More than 30 partners, including the United Kingdom and Australia issued a public attribution statement condemning Russia’s actions.\(^{16}\) In 2022, following the attribution to Iran, Albania severed diplomatic ties with the country that orchestrated a cyber operation that pummeled its digital government services.\(^{17}\) The importance of co-attribution and coordination amongst like-minded states to strengthen attribution cannot be understated, especially in times of growing geopolitical tensions.

### PUBLIC Attribution as a Political Tool

Cyber attribution may be aimed at states as well as at a broader audience in international relations. Public attribution can also serve to strengthen and reaffirm agreed international norms, rules, and principles for responsible behavior of states in cyberspace. It also seeks to clarify individual State’s understanding of how international law applies to cyberspace.\(^{18}\) Since 2019, United Nations (UN) negotiations on a Cybercrime treaty have been taking place to expand agreement on what constitutes criminal behaviors in cyberspace, with fears from Western states that a new treaty would criminalize too many acts and infringe on human rights.\(^{19}\)

UN discussions on responsible state behavior in cyberspace have taken place over two decades through Groups of Governmental Experts (six UN GGEs, 2004-2021) and multi-stakeholder Open-Ended Working Groups (two OEWGs, since 2019). Although the norms, rules, and principles for responsible state behavior in cyberspace resulting from these processes are formally non-binding, adhering to them is generally understood as critical for stability in cyberspace and international security. For example, norm 13 (c) proclaims that states should not knowingly allow their territory to be used for wrongful acts in cyberspace, whereas norm 13 (f) directs states to not damage or impair critical infrastructure through their conduct or support of information and communication technology (ICT) activities. To be able to uphold and ensure compliance with those norms, an effective and credible attribution process is necessary. Norm 13 (b) directly addresses attribution by recognizing its importance, its inherent challenges,


\(^{19}\) The existing Council of Europe Convention on Cybercrime or Budapest Convention to combat cybercrime has been ratified by 68 states. The new UN cybercrime treaty currently under negotiation has extensive issues. See Cynthia Brumfield, “New UN cybercrime convention has a long way to go in a tight timeframe,” CSG, January 31, 2023, https://www.csoonline.com/article/574447/new-un-cybercrime-convention-has-a-long-way-to-go-in-a-tight-timeframe.html.
and the risk of escalation that cyber incidents pose to international peace and stability.20

Unfortunately, and not surprisingly, the states considered major cyber offenders currently appear not to adhere to these norms. Ongoing geopolitical tensions, especially given war in Ukraine and Russia’s use of cyber capabilities, plus increased ideological competition over an open versus closed internet, and a surge in ransomware and other cybercrime are likely factors that further undermine efforts towards cyberspace accountability. Russia and China have repeatedly called out Western states’ attributions of malicious cyber operations as hypocritical and led purely by a political agenda to promote their open vision for cyberspace. The attributors and the accused often fall along the fault lines of the international cyber norms discussions, who have an interest in promoting their “rules” or “understanding” of normative behavior and standards in cyberspace. Effectively assessing attribution thus requires deeper knowledge of the mechanisms and procedures for building a case.

UNDERSTANDING THE ATTRIBUTION PROCESS

Attribution is a multi-level process that determines responsibility for a malicious cyber incident. The attribution process can be broken down into three parts.21 First, from what computers and where did the incidents originate and how did they transit. In this first step, the machines that carried out the operations are identified, using technical forensics. Second, who are the operators behind the screens and on the keyboards who executed the attack? In this next step, the human operators of the machines are identified, using operational forensics. Third, who is the entity ultimately responsible for the attack? In this last, and most challenging step, attribution attempts to identify the operators’ relationship with a responsible state and its role, if any, in the malicious act. To express the degree of certainty in the result of the investigation, attributing states and private cybersecurity firms often do and should assign a level of confidence to the attribution. This is necessary when asking others to join a public attribution and possibly support punitive measures against the perpetrators.

Once a significant incident has been discovered, the government or private firm will begin an investigation with the objective of establishing attribution to a threat actor. Commonly starting at a technical level, the investigation tries to collect technical artifacts, relevant data regarding the network intrusion, malicious activities, and data exfiltration among others to gather information about the scope and type of the incident and the threat actor. Patterns of behavior, for example, may help identify and match the perpetrator to a known threat actor. Sophisticated threat actors may try to lead investigators in the wrong direction by falsifying or placing misleading evidence in “false flag” operations. A faulty investigation undermines trust among allied partners, including in the private sector, and can call into question the legitimacy and reliability of the current and future attribution efforts.22 Trust among government partners, backed by a high level of confidence in the investigation’s outcome through well-defined technical, procedural, and legal standards, is critical for attribution.

Although technical and operational forensics can supply the “where,” “when,” and “how,” puzzle pieces to the attribution, they may play only a minor role in the final political judgment to make a public attribution. It is the “why” puzzle piece that aids in concluding the ultimate party responsible and requires political contextualization and understanding of motive. There is an important distinction between (a) identifying intrusion sets and assigning them to an adversary or threat group and linking this to a non-state actor or a state, and (b) then further determining a state’s role and legal responsibility under international law in actively or passively supporting the malicious act, simply not seeing it, or falling short of the capacity to stop it.23 The investigation of this last layer of responsibility is the most complex one, with states and their technical, legal, and intelligence agencies best placed to investigate this.

Cybersecurity firms engaged in threat intelligence services provide important capabilities and intelligence for attribution. Companies like CrowdStrike, Symantec, Microsoft, and others make available threat intelligence and attribution reports, often faster than governments. Knowledge resides also in academia and think tanks, which can help with driving impartial, transnational attribution and develop evidentiary standards.24

Finally, a state has to decide whether and how to publicly disclose the determination. Increasingly, like-minded states have coordinated their public attribution, such as in the cyber operations that exploited vulnerabilities in SolarWinds and Microsoft Exchange Server noted earlier, increasing the pressure and cost on attributed cyber perpetrators. States’ public attribution supports specific policy objectives, such as reaffirming behavioral and normative standards important to the attributing state or deterring cyber offenders from engaging in hostile actions in cyberspace.25 However, numerous factors may weigh against public disclosure: a government may not have sufficient confidence in the attribution decision, sources and methods that have enabled intelligence collection might be compromised, or a state decides it is not in its broader interest to publicly accuse. Attributing without overtly or covertly acting can undermine norms whereas establishing credible costs and deterring through measures and actions can discourage attackers and enhance accountability.

MOVING THE NEEDLE ON ACCOUNTABILITY

Holding malicious states and actors accountable for their actions remains a challenge. For one, states accused of malicious cyber offenses reject accusations and question the integrity of the attribution and its political intent. To help address this, a shared and agreed upon, transparent, defensible process is needed to strengthen the credibility and effectiveness of attribution. The private sector can help drive and collaborate in such a process. As the 2023 U.S. National Cybersecurity Strategy notes, “The private sector has growing visibility into adversary activity. Its body of insight is often broader and more detailed than that of the federal government, due in part to the sheer scale of the private sector and its threat hunting operations, but also due to the innovation into tooling and latent capabilities.”26 To get closer to solving accountability, societies need both -- the insights and capabilities from private sector firms and the government authorities and the means to act and hold perpetrators ultimately accountable.

Partnerships are developing to support better information sharing with transparent processes that aid attribution. Some stakeholders are collaboratively developing information on cyber incidents to assist in identifying and tracking malicious actors. Examples include MITRE’s move to extend its ATT&CK framework into charting adversary behaviors and the World Economic Forum’s Cybercrime Atlas.27 28

Second, the toolbox to hold bad actors accountable has appeared limited thus far. Diplomatic, economic, and other overt measures, including sanctions, have fallen short of being effective and in some cases can have unintended side effects.29 The United States, for example, has indicted Russian government employees for their roles in global hacking campaigns. However, this largely remains a symbolic act.30 While attribution, if done right, can help identify attackers and degrade or halt their operations at least for some time if coupled with some defensive actions or punitive measures, there is little evidence that suggests that major perpetrators would change their behavior through public “naming and shaming” that attribution affords. The latter holds especially true in times of heightened geopolitical tensions where cyber operations are a low-cost tool to engage in gray-zone confrontation.

Some governments are not waiting for an incident to occur but are adopting a proactive posture and employing new strategies and operational approaches. The United States through its strategy of defend forward and persistent engagement, for instance, seeks constant engagement with the adversary to shape the cyberspace environment actively and persistently. It means the U.S. government is working with allies to intervene in potential adversaries’ plans to prevent, stop, or mitigate the effects of an incident. This approach not only helps to learn about threat actors and their TTPs, which can be beneficial for attribution in other instances, but is also intended to thwart some cyber attacks thus reducing the need for attribution of cyber operations that did not unfold.

Recently, many are also starting to call for accountability and responsibility for cybersecurity of technology firms and service providers more broadly. The European Union, for instance, through amendments to the EU Cybersecurity Act, proposes heightened security requirements to providers of managed security services and critical infrastructure operators, and the United States aims to shift more cybersecurity responsibilities to technology firms and service providers. Rebalancing the burdens of cybersecurity away from end-users to the most capable in the ecosystem, strengthening product and service cybersecurity by default, facilitating stronger cyber defenses, and increasing investments in cyber resilience provide important complementary approaches to shape adversaries’ behaviors and add important policy measures to the tool box.

While institutional and policy changes take time, the issues of identifying “who did it” in cyberspace and keeping perpetrators accountable will remain a key challenge. The current toolbox needs to be reviewed, revised, and expanded with effective, evidence-based policies and measures for attribution and accountability. In the long run, a transparent, robust, and broadly accepted attribution process is needed with the goal of increasing joint attribution statements and coordinated responses.

ACKNOWLEDGEMENTS

The authors acknowledge Allison Pytlak and Herbert Lin for their review, comments, and suggestions on an earlier draft of this paper. This background paper reflects the personal research, analysis, and views of the authors, and does not represent the position of either of the institution, its affiliates, or partners.
The Observer Research Foundation America (ORF America) is an independent, non-partisan, and nonprofit organization in Washington DC dedicated to addressing policy challenges facing the United States, India, and their partners in a rapidly changing world.

ORF America produces research, curates diverse and inclusive platforms, and develops networks for cooperation between the developed and developing worlds based on common values and shared interests. Its areas of focus are international affairs, technology, climate and energy, and economics. Established in 2020, ORF America is an overseas affiliate of the Observer Research Foundation (ORF), India’s premier non-government think tank.

Observer Research Foundation America
1100 17th St. NW, Suite 501, Washington DC 20036

www.orfamerica.org