Policy Perspectives
Vol. 11/5, October 2023

CSS
ETH Zürich

Making Cyber Attribution More Transparent

Following the example of Intrusion Truth, governments should substantiate their attribution statements that identify the perpetrators of malicious cyber operations with open-source intelligence.

By Eugenio Benincasa

Governments are generally reluctant to attribute malicious cyber operations targeting them to a specific country, entity, or operator. This is often due to technical obstacles, such as limited forensic capabilities, and the risk of exposing classified sources and methods. When states do attribute, they fear publicly disclosing extensive evidence for the same secrecy-related reasons. As a result, public attributions (see glossary on page 2) of cyber incidents often lack transparency. Can this approach be regarded as responsible state conduct in cyberspace?

Back in April 2017, an anonymous online group named Intrusion Truth created an online blog. Since its creation, the blog has exposed the real identities of more than 30 Chinese state-sponsored cyber operatives across several Chinese Advanced Persistent Threat (APT, see glossary) groups. Intrusion Truth also succeeded in connecting these operatives to the Ministry of State Security (MSS) and its regional bureaus in Tianjin, Jinan, Hainan, and elsewhere.

In all its investigations, Intrusion Truth relies primarily on Open-Source Intelligence (OSINT, see glossary) and publishes a comprehensive account of its findings in a step-by-step manner on its blog. Its ability to leverage OSINT reliably and consistently for evidence-based, high-confidence public attribution purposes remains unique.

It is unknown who is behind Intrusion Truth and how many people are involved. While it is possible that intelligence or law enforcement officers actively collaborate with or form a part of Intrusion Truth, no evidence has yet emerged to support such a claim.

Key Points

- Many states hesitate to attribute malicious cyber operations targeting them due to technical limitations, and the risk of exposing classified sources and methods. Those that do attribute often withhold extensive evidence for the same secrecy-related reasons.
- Since 2017, Intrusion Truth has identified more than 30 Chinese intelligence cyber operatives by primarily using open-source intelligence. They also publish their findings on their blog.
- Intrusion Truth's revelations have had tangible operational consequences, inflicting technical and socio-economic costs on Chinese threat actors and influencing policy discussions.
- Government agencies should integrate open-source intelligence into their public attribution processes and statements to overcome overarching challenges.

Nevertheless, the activities of Intrusion Truth offer valuable lessons for government agencies. By optimizing the use and integration of OSINT in attribution processes and statements, government agencies can help address transparency issues in public attribution.

Public Attribution involves identifying the perpetrator of malicious cyber operations and sharing this information publicly for diplomatic and deterrence purposes. In contrast, private, bilateral attribution is a discreet process where the victim communicates directly with the perpetrator to request or compel the cessation of malicious activities.

Advanced Persistent Threat (APT) refers to a covert cyberattack on a computer network where the attacker gains and maintains unauthorized access to the targeted network and remains undetected for a significant period.[1]

Open-Source Intelligence (OSINT) is defined as intelligence produced by collecting, evaluating, and analyzing publicly available information with the purpose of answering a specific intelligence question.[2]

Untransparent Public Attribution

The process of attributing cyber operations involves three key considerations: technical, political, and legal. First, the attributing state must understand the attack on a technical level through forensic analysis, often requiring both signals intelligence and human intelligence to trace the actors responsible. Second, after identifying the source of the operation, the state faces political decisions about whether and how to publicly hold the responsible actor accountable, considering domestic and international factors. Finally, from a legal perspective, attribution means assigning responsibility by connecting the offense to an offender according to applicable rules, potentially legitimizing future responding measures. The current body of international law does not require evidence disclosure for cyber attribution.

The issue of untransparent public attribution arises when states publicly blame another for an attack without sufficient evidence. This problem is exacerbated by political motivations, technical limitations, and the absence of legal constraints.

States disagree on the necessity of disclosing evidence to support accusations. Since 2015, China, Russia, and other nations have been arguing that accusations should be supported by evidence. In contrast, the United States, the United Kingdom, and other European nations contend that, since international law does not mandate the disclosure of evidence to validate accusations, states should act based on what they deem reasonable given the circumstances. This view is informed by the position that disclosing evidence could expose classified information. States holding this position base their public attribution of attacks on the tools and methods used, as well as the broader geopolitical context.

Untransparent public attribution has several negative implications. These include its ineffectiveness as a deterrent; inability to help forge alliances; potential to leave room for misinterpretation; and susceptibility to political manipulation. It also contributes little to norm-building in cyberspace due to the lack of clear rules.

The Modus Operandi of Intrusion Truth

Intrusion Truth's investigations primarily draw upon publicly available information, demonstrating that threat actors leave important traces behind that can be accessed without requiring security clearance.

There are notable instances of such traces discovered by Intrusion Truth. These include public postings on discussion forums and online job advertisements which displayed a significant interest in malware development and password-cracking skills. Importantly, these public postings and advertisements contained identifiers, such as usernames, email addresses, and phone numbers. The exposed individuals often used the same identifiers for covert malicious operations, such as registering front companies. When identifiers and data linked to covert activity appeared in public sources like threat intelligent reports, analysts were able to establish connections to specific individuals and APT groups.

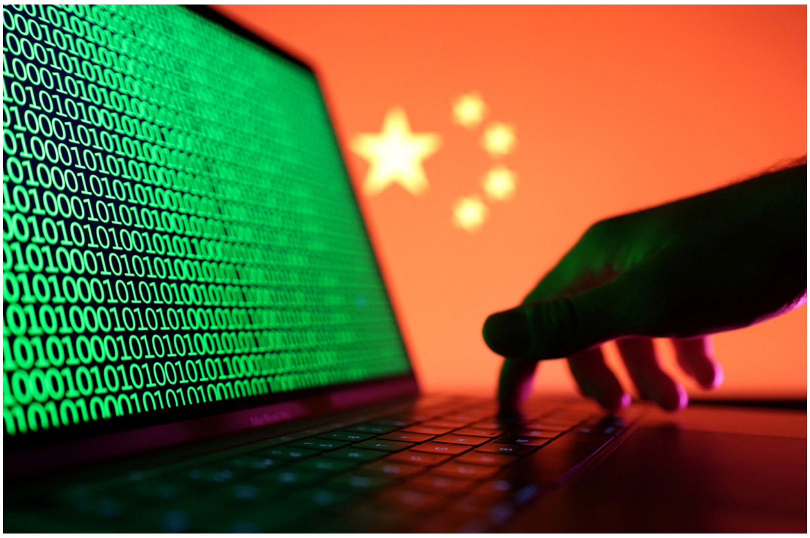
In some cases, identifiers such as distinctive names, phone numbers, or email addresses employed in both overt and covert malicious activities were also utilized for private use, such as on social media and mobile apps. This enabled Intrusion Truth to uncover personal identifiable information, including private photos and user locations. With this data, analysts were able to establish connections between the hackers' personal information, the malicious activities they were involved in, the front companies they operated, and their affiliations.

Intrusion Truth also sometimes relied on sources other than OSINT to bridge information gaps, including anonymous tips and data contributed by others. More specifically, examples of this included tip-offs from anonymous analysts, credit card statements, Uber ride receipts, private cloud storage accounts, and frequent flyer membership information.

Operational Impact of Intrusion Truth

The operational impact of Intrusion Truth's revelations can be broken down into two categories.

First, the disclosure of the hackers' identities and their connection to specific tools, tactics, and behaviors likely imposed heightened technical costs on the exposed groups by providing information for potential victims and



Since April 2017, an anonymous group calling itself Intrusion Truth has exposed the identities of individuals working for Chinese intelligence. *Dado Ruvic / Reuters*

making malicious activities harder to conceal. This effect is amplified when considering that, over the last few years, numerous cyber threat actors from China have collaborated and shared resources.

Second, Intrusion Truth's revelations have imposed socio-economic costs on individual Chinese operators. These include individual travel constraints and disincentives to continue their work following exposure by Intrusion Truth. Public identification diminishes the operators' prospects of finding work outside China and exposes them to potential charges by foreign law enforcement agencies. The revelations also potentially erode trust between the MSS and the hackers, and even among the hackers themselves. For instance, Intrusion Truth's actions may cause anxiety and apprehension in APTs by revealing hackers' careless mistakes or their presence on social media, something that is typically frowned upon by intelligence agencies. This exposure may instill distrust and fear among non-exposed individuals within a group, including about their potential apprehension by foreign law enforcement agencies. This is a valid concern, especially considering that the US Department of Justice (DOJ) has, in some cases, issued indictments not only against exposed hackers but also affiliated individuals not specifically mentioned in Intrusion Truth's blog posts.

Policy Impact of Intrusion Truth

The revelations published by Intrusion Truth have been corroborated by leading cyber threat intelligence companies, including Crowdstrike[3] and Recorded Future.[4] Moreover, they have had a discernable impact. They have led to the disappearance of APTs and US DOJ indictments of individuals affiliated with Chinese threat groups, such as APT3, APT10, and APT40. More generally, Intrusion Truth's revelations have exposed the inner workings and coordination of China's multifaceted offensive ecosystem, which is characterized by strong interlinkages between private companies, intelligence agencies, and academia.

Intrusion Truth's findings have contributed to policy discussions surrounding how to address the potential threats arising from China's expanding offensive cyber ecosystem. This year, Bart Groothuis, Member of the European Parliament, indirectly referenced the work of Intrusion Truth while urging EU member states to limit the participation of Chinese high-risk equipment suppliers in their 5G networks. In an interview on the matter, he affirmed the following: "Past open source reporting has shown an overlap between Huawei personnel and Chinese spies. In the case of hackers group APT3, Boyusec and CNITSEC, Huawei vulnerabilities have been exploited against European targets."[5]

It is difficult to determine the extent to which Intrusion Truth's findings influenced the US DOJ's public denunciations of Chinese APTs, including those mentioned earlier. However, the group has raised awareness about China's state-sponsored threat actors and contributed to a larger debate on cyber attribution.

Transparent Cyber Attribution by States

Government agencies employ OSINT in attribution processes to differing degrees, contingent on their technical capabilities. However, OSINT is rarely, if ever, included in states' attribution statements. In addition to OSINT, the attribution process often requires classified sources and methods, such as signals intelligence and human intelligence, to trace the actors responsible for cyber operations. At the same time, institutional factors often create biases that favor the use of classified information and internal datasets over publicly available information. This leads to an underestimation and neglect of the value and volume of non-classified sources, ultimately resulting in their underutilization.

The case of Intrusion Truth provides several valuable lessons for government agencies. First, OSINT can in certain cases provide a cost-effective alternative to traditional classified intelligence methods, meeting national security needs efficiently and providing a comprehensive threat landscape understanding. Second, OSINT can complement or validate classified information, protecting sensitive sources and methods during public attribution. Third, OSINT is particularly useful for initial intelligence gathering, helping analysts move from suspicion to identifying culprits and gaining a preliminary understanding of their activities. However, OSINT may not always provide conclusive evidence, and traditional intelligence disciplines may be necessary to fill gaps in such cases.

Further Reading

Kim Zetter, **“Unmasking China’s State Hackers,”** *Zero Day,* 29.03.2022.
Provides additional insights on Intrusion Truth’s history, mindset, and motives.

Intrusion Truth, https://intrusiontruth.wordpress.com.
Intrusion Truth’s official WordPress blog, which serves as a repository for all their detailed analyses and investigations.

Fan Yang, **“The Problem With Ill-Substantiated Public Cyber Attribution: A Legal Perspective,”** in: Ariel E. Levite / Lu Chuanying / George Perkovich / Fan Yang (eds.), *Managing U.S.-China Tensions Over Public Cyber Attribution* (Washington DC: Carnegie Endowment for International Peace, 2022), 6–13.
In-depth legal analysis of the risks and concerns stemming from ill-substantiated public attribution.

In addition, the public release of information by intelligence agencies is limited by strict bureaucratic regulations and oversight. While government agencies may occasionally disclose information for strategic reasons, they typically err on the side of excessive classification. This approach can vary based on circumstances. In late 2021, the US employed declassified intelligence in an information campaign against Russia. Despite lacking absolute certainty about the intelligence, the US used it to disrupt Moscow’s plans of invading Ukraine. This process involved frequent and extensive intelligence disclosures that required extra safeguards for sources and methods during declassification. When feasible, a similar approach could be applied to cyber attribution. This could help tackle the challenges linked to untransparent attributions.

Finally, transparent cyber attribution is vital for democratic decision-making and political legitimacy, particularly in areas that deal with malicious cyber activity concerns like insurance, criminal proceedings, and national security. A lack of transparency leaves societies uninformed about state decisions, leading to accountability gaps and misunderstandings.[6]

Selected sources

1. Cisco, *What Is an Advanced Persistent Threat (APT)?,* cisco.com.
2. Ritu Gill, “What Is Open-Source Intelligence?” *SANS,* 23.02.2023.
3. Adam Kozy, “Two Birds, One STONE PANDA,” *Crowdstrike Blog,* 30.08.2018.
4. Insikt Group, “Recorder Future Research Concludes Chinese Ministry of State Security Behind APT3,” *Recorded Future Blog,* 17.05.2017.
5. Alexander Martin, “EU States Told to Restrict Huawei and ZTE from 5G Networks ‘Without Delay,’” *The Record,* 16.06.2023.
6. Florian J. Egloff / Andreas Wenger, “Public Attribution of Cyber Incidents,” *CSS Analyses in Security Policy* 244 (2019).

Eugenio Benincasa is Senior Researcher in the Cyber Defense Project at the Center for Security Studies (CSS) at ETH Zurich.

Policy Perspectives is published by the Center for Security Studies (CSS) at ETH Zürich. The CSS is a center of competence for Swiss and international security policy.

Series Editor: Oliver Thränert
Issue Editor: Boas Lieberherr
Layout: Miriam Dahinden-Ganzoni

Feedback welcome: PolicyPerspectives@sipo.gess.ethz.ch
More issues and free online subscription:
css.ethz.ch/en/publications/css-policy-perspectives

Most recent editions:
Satellite Imagery for Disaster Resilience (11/4)
Mind the E-Waste: A Case for Switzerland (11/3)
Securing Europe’s Supply of Rare Earths (11/2)
Navigating Stormy Seas in US-China Relations (11/1)
The Ukraine Drone Effect on European Militaries (10/15)
Minsk’s Signals: Belarus and the War in Ukraine (10/14)

ISSN: 2296-0244; DOI: 10.3929/ethz-b-000637147