

OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE



A Guide to Cyber Attribution

LEADING INTELLIGENCE INTEGRATION

14 September 2018

(U) KEY TAKEAWAY

***Scope Note:** This memo explains the key concepts the US Intelligence Community (IC) uses to identify the perpetrators of malicious cyber activities. This memorandum was prepared under the auspices of the National Intelligence Officer (NIO) for Cyber Issues. It was drafted by the NIO and the National Intelligence Manager for Cyber.*

Establishing attribution for cyber operations is difficult but not impossible. No simple technical process or automated solution for determining responsibility for cyber operations exists. The painstaking work in many cases requires weeks or months of analyzing intelligence and forensics to assess culpability. In some instances, the IC can establish cyber attribution within hours of an incident but the accuracy and confidence of the attribution will vary depending on available data.

To help with this process, the IC has identified several key indicators to evaluate and determine responsibility for an attack. We also have identified best practices for assessing cyber attribution and presenting our related assessments. A common approach to attribution can help to standardize communications with policymakers on cyber attribution and facilitate timely sharing of data and analytic collaboration.

Attribution: Difficult First Step for National Response

Russia, China, Iran, North Korea, and malign actors all use cyber operations as a low-cost tool to advance their interests, and we assess that unless they face clear repercussions for such actions will continue to do so. Cyber attribution, or the identification of the actor responsible for a cyber attack, therefore is a critical step in formulating a national response to such attacks.

Every kind of cyber operation—malicious or not—leaves a trail. Our analysts use this information, along with their knowledge of previous events and the tools and methods of known malicious actors, to attempt to trace these operations back to their sources. Analysts compare the new information to existing knowledge, weigh the evidence to determine a confidence level for their judgments, and consider alternative hypotheses and ambiguities to produce cyber attribution assessments.

There is no simple technical process or automated solution to determine responsibility for cyber operations. This painstaking work in many cases requires several weeks or months of analyzing intelligence and forensics. In some instances in which analysts can determine responsibility for a cyber attack within hours of an incident the accuracy and level of confidence is likely to vary depending on the available data.

- Analysts can assess responsibility for a cyber attack in three ways: the point of origin, such as a specific country; a specific digital device or online persona; or the individual or organization that directed the activity.
- This third category often is the most difficult to assess because we have to link malicious cyber activities to the specific individuals and assess the sponsor and motivators of these individuals.

Key Indicators That Enable Attribution

Attributing an attack to a particular country or actor requires collecting as much data as possible to establish connections to online actors, individuals, and entities. Because this often results in hundreds of conflicting indicators, we identified key indicators to guide us in seeking timely, accurate attribution. The primary

indicators are **tradecraft, infrastructure, malware, and intent**. We also rely on **indicators from external sources**, such as open-source reports from the private cybersecurity firms.

- **Tradecraft:** Behavior frequently used to conduct cyber attack or espionage. This is the most important indicator because habits are more difficult to change than technical tools. An attacker's tools, techniques, and procedures can reveal attack patterns, but these unique tradecraft indicators diminish in importance once they become public and other actors can mimic them.
- **Infrastructure:** The physical and/or virtual communication structures used to deliver a cyber capability or maintain command and control of capabilities. Attackers can buy, lease, share, and compromise servers and networks to build their infrastructure. They frequently establish infrastructure using legitimate online services, from free trials of commercial cloud services to social media accounts. Some are loath to abandon infrastructure, while others will do so because they can rebuild it within hours. Some routinely change infrastructure between or even within operations to impede detection.
- **Malware:** Malicious software designed to enable unauthorized functions on a compromised computer system such as key logging, screen capture, audio recording, remote command and control, and persistent access. An increasing number of cyber actors can modify some malware indicators within minutes or hours of suspected compromise, and some routinely change malware between or within operations to impede detection and attribution.
- **Intent:** An attacker's commitment to carry out certain actions based on the context. Covert, deniable cyber attacks often are launched against opponents before or during regional conflicts or to suppress and harass enemies of the state.
- **Indicators from External Sources:** We also use reports from the private industry, the media, academia, and think tanks to provide such data or share hypotheses about the perpetrators.

Best Practices for Determining Attribution

Identifying these key indicators requires rapid and careful work. We identified three practices that can aid in the identification of cyber attackers.

- **Looking for Human Error.** Almost all cyber attribution successes have resulted from discovery and exploitation of the attackers' operational security errors. Cyber intruders have often made mistakes related to tradecraft and the use of cyber infrastructure. Our adversaries have sought to minimize these errors with varying degrees of success.
- **Timely Collaboration, Information Sharing, and Documentation.** Attribution efforts benefit from combining the expertise of regional, political, and cybersecurity analysts and the collaboration of network defenders, law enforcement, private cybersecurity firms, and victims. Acquisition, documentation, and recovery of data within twenty-four hours of a cyber incident also is critical because data-deletion cyber attacks can erase the log data necessary for forensics, advance malware dissipates in computer memory, and adversaries may abandon cyber infrastructure within hours of its discovery.
- **Rigorous Analytic Tradecraft.** Analysts may start with a set of plausible actors in mind, based on the nature of the cyber incident, the targets, and the context but must be careful to avoid cognitive bias. To

minimize this risk, analysts can use techniques such as Analysis of Competing Hypotheses, which helps to evaluate multiple competing hypotheses based on the observed data and uncover data that might reveal other potential actors.

Best Practices for Presenting Attribution Analysis

Our best practices for presenting analysis related to cyber attribution include **de-layering the attribution assessment, providing the confidence level, and identifying gaps**. Our attribution assessments typically include a series of judgments that describe whether the event was an isolated incident or not, the likely perpetrator, possible motivations, and whether a foreign government played a role.

De-layer the Judgment. A statement of attribution should include a clear distinction among the following: the physical location where the activities originated, the individual actors or groups involved, and whether leadership sponsorship or direction could be determined.

Provide Confidence Level. Our analysts evaluate three components when assigning probabilistic language and confidence levels: the timeliness and reliability of the evidence, the strength of the logic linking the evidence, and the type of evidence (direct, indirect, circumstantial, or contextual). In many cases, analysts also consider competing hypothesis in order to uncover possible alternative actors.





























- **High Confidence.** This level of confidence is used when analysts judge the totality of evidence and context to be beyond a reasonable doubt with no reasonable alternative. For example: “The Xandi Cyber Force (XCF) almost certainly is responsible for the destructive cyber attack on the Terran oil company. We have high confidence in this assessment because XCF operators discussed how they compromised the oil company and the steps they took to damage the company’s systems.”
- **Moderate Confidence.** This level of confidence is used when analysts judge the totality of evidence and context to be clear and convincing, with only circumstantial cases for alternatives. For example: “Xandi security services are very likely responsible for hacking the e-mail accounts of several Terran human rights activists. We have moderate confidence in this judgment because the hacking operations are linked to known Xandi intelligence infrastructure and the victims are also the Xandi’s priority targets.”
- **Low Confidence.** Analysts use this level of confidence when they judge that more than half of the body of evidence points to one thing, but there are significant information gaps. For example: “Terra probably was responsible for the data deletion attack on a Xandi bank last week after Xandi sanctions were imposed on multiple Terran companies. We have low confidence in our judgment because the actor used publicly available tools, which although previously associated with Terran intelligence, also are used by criminals.”

Identify Gaps. In cases where analysts do not have enough data for a judgment or confidence statement because there are insufficient indicators, they should state this explicitly. For example, “We do not yet have enough information to assess who is responsible for the disruptive cyber attack on the Xandi energy company. We suspect the attackers used a botnet originating from Terra. The attack did not coincide with any bilateral tension between the Xandi and known adversaries.”

Cyber Attribution Examples

The chart below shows how we use analysis of competing hypotheses in combination with the key attribution indicators to show what data we have to link the cyber incident to the actor.

Data to associate with incident:  Sufficient  Limited

CYBER INCIDENT		ADVERSARY	KEY INDICATORS FOR ATTRIBUTION				
			Tradecraft	Infrastructure	Malware	Intent	External Sources
2017	MARCH Major Compromises of Global IT Firms	RUSSIA					
		CHINA*					
		NORTH KOREA					
		IRAN					
		NON-STATE					
	MAY Wannacry Attacks	RUSSIA					
		CHINA					
		NORTH KOREA*					
		IRAN					
		NON-STATE					
	JUNE NotPetya Attacks	RUSSIA*					
		CHINA					
		NORTH KOREA					
		IRAN					
		NON-STATE					
	DECEMBER Saudi Petrochemical Facility Attack	RUSSIA					
		CHINA					
		NORTH KOREA					
		IRAN					
		NON-STATE					

* We highlight the actor we assess to be responsible for the cyber incident when we have a sufficient body of information to link the actor's tradecraft, infrastructure and/or malware to malicious cyber activities.