INTELLIGENCE · CONCEPTS AND HISTORY

Cyber Intelligence Collection and Attribution Challenges

Working out who is behind a cyber attack is one of the hardest problems in modern intelligence. This article explains how cyber intelligence is collected, how analysts build a case for attribution, why false flags and uncertainty make it difficult, and what organisations should take from it.

ModQu research article · 1,686 words · 18 references (APA 7th edition)

When a company's network is breached, a hospital's systems are encrypted for ransom or an election campaign's emails are leaked, two questions follow immediately: what happened, and who did it? The first is a technical investigation. The second, attribution, is an intelligence problem of unusual difficulty. Attackers hide behind compromised computers in third countries, reuse tools that others also use, and sometimes plant false clues to implicate someone else. Yet attribution matters: it shapes how victims defend themselves, whether governments respond and whether anyone is held to account. This article explains how cyber intelligence is collected, how attribution works, why it is so hard and what lessons organisations can draw.

Sources of cyber intelligence

Cyber intelligence draws on an unusually wide range of sources. Lowenthal (2022) notes that cyberspace is both a domain in which intelligence is collected and a target of intelligence in its own right.

- **Victim data.** Logs, memory images, malicious files and network traffic from affected systems are the foundation of any investigation. They show what the attacker did, which tools were used and which infrastructure the malware contacted.
- **Infrastructure analysis.** Domain registrations, hosting records and the reuse of servers or certificates across incidents can link separate attacks. Caltagirone et al. (2013) describe how analysts "pivot" from one known element, such as an IP address, to others connected with it.
- **Malware analysis.** Code contains clues: programming habits, language settings, compile times, shared libraries and reused components. Families of malware can be traced over years.
- **Vendor and community reporting.** Security companies publish analyses of threat groups, and organisations share indicators through industry groups. Lemay et al. (2018) surveyed hundreds of public reports on advanced persistent threat groups and showed how much detail is now openly available, and how inconsistent naming across vendors complicates its use.

- **Government collection.** Signals intelligence, human sources and access to attackers' own infrastructure can provide evidence that no private investigator could obtain. Rid and Buchanan (2015) note that the most confident attributions often rely on such sources, which governments are reluctant to reveal. Landau (2013), analysing the 2013 disclosures about American signals intelligence, shows how extensive such access to global communications networks can be, and why its existence is closely guarded.

Hutchins et al. (2011) showed that analysing each stage of an intrusion, from reconnaissance to the attacker's final actions, reveals patterns that persist across campaigns. Strom et al. (2018) describe how the MITRE ATT&CK framework catalogues these behaviours, providing a common language for describing how particular groups operate.

How attribution works

Rid and Buchanan (2015) argue that attribution is not a single technical answer but a process of judgement, closer to intelligence analysis than to forensic science. They describe it as operating at several levels. At the **tactical** level, investigators work out how an intrusion occurred and what was taken. At the **operational** level, they connect the intrusion to a wider campaign and a recognisable group with characteristic tools and methods. At the **strategic** level, decision-makers judge who ultimately directed the activity and why, which may mean identifying a state or organisation behind a technical group. Finally, attribution must be **communicated**: decision-makers must decide what to say publicly and how much evidence to disclose.

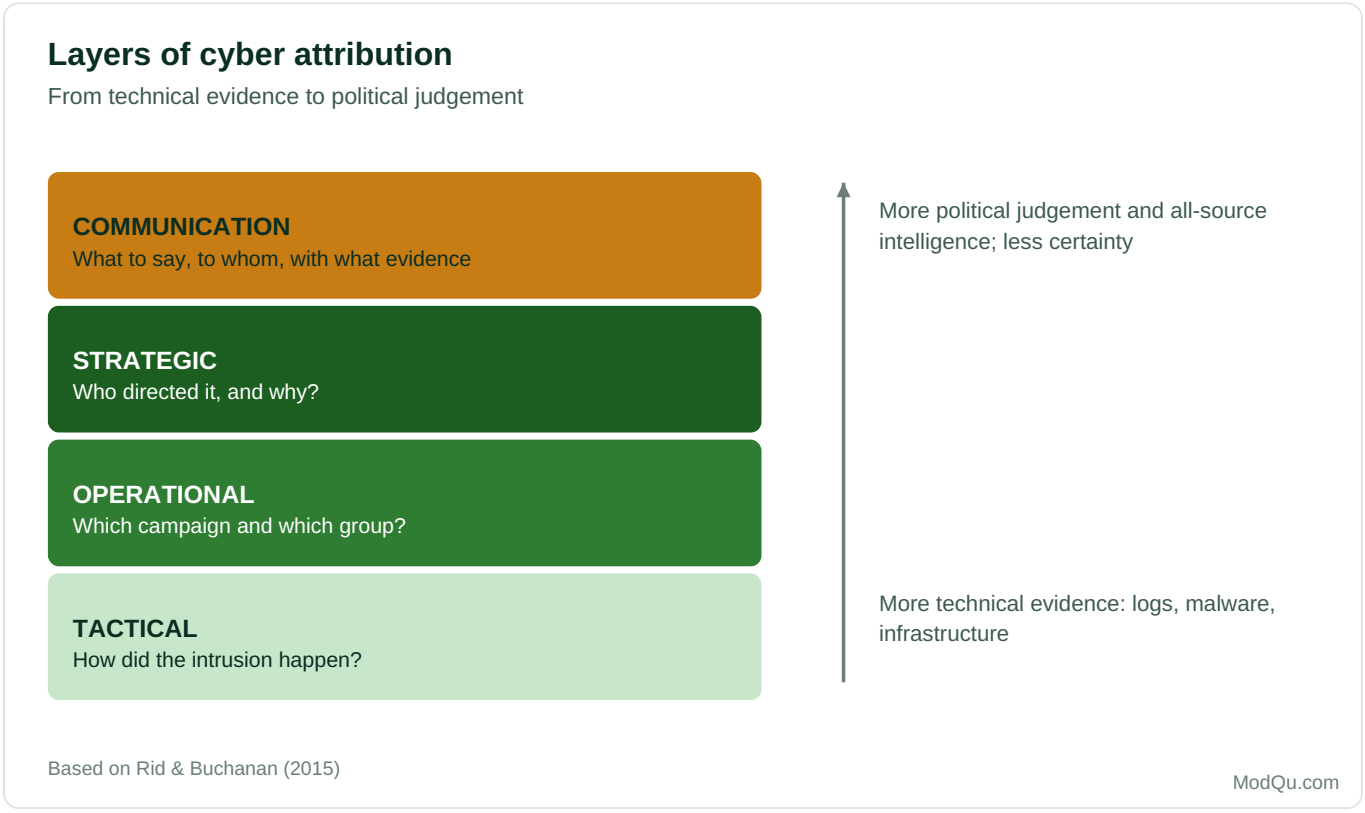


Figure 1. Layers of cyber attribution. Technical evidence is strongest at the bottom; questions of sponsorship and intent at the top depend increasingly on all-source intelligence and political judgement. Based on Rid and Buchanan (2015); diagram by ModQu.

Each level draws on different evidence and carries different uncertainty. Technical indicators can establish with reasonable confidence that two intrusions used the same infrastructure. Linking a technical group to a particular government usually requires context: the targets chosen, the timing in relation to political events, the information sought and, often, intelligence from non-technical sources.

A well-known public example is the 2013 report by the security company Mandiant, which linked a long-running espionage campaign against dozens of organisations to a specific unit of China's People's Liberation Army, citing infrastructure, the location of operators and the pattern of targets (Mandiant, 2013). The report was unusual in naming a state unit and in publishing much of its evidence.

Why attribution is hard

Indirection

Attackers rarely operate from their own computers. They route activity through compromised machines, rented servers and anonymising networks in several countries. The computer that delivered an attack may belong to an innocent third party.

Shared tools

Many groups use the same publicly available tools, purchased malware and common techniques. Similarity of method is weaker evidence of identity than it first appears.

False flags

Sophisticated attackers can plant misleading clues: comments in another language, code borrowed from other groups or infrastructure associated with someone else. Buchanan (2020) describes cases in which state actors appear to have deliberately imitated other groups to confuse investigators. The possibility of deception means that every clue must be weighed for the chance that it was planted.

Time pressure and secrecy

Victims and governments often need to act before investigations are complete. The most decisive evidence may come from intelligence sources that cannot be disclosed without compromising them, so public attributions may appear thinly supported even when the underlying judgement is sound.

Libicki (2009) argued that these difficulties undermine deterrence in cyberspace: if an attacker believes it will not be identified with enough confidence to justify a response, the threat of retaliation loses force.

Weighing the evidence: an illustration

Suppose a manufacturing company discovers that design files were copied from its servers over several months. Its investigators find that the attackers entered through a phishing email, used a widely available remote-access tool, and sent data to a server rented through a hosting company in a third country. The server's address appears in a security vendor's report about a group that has targeted the same industry before, and the phishing email was written in fluent English but contained a time zone setting associated with another region.

How much does this show? The shared server is a meaningful link to the known group, but servers are sometimes reused by different actors or seized and repurposed. The remote-access tool is common and proves little. The time zone setting is easy to change and may be deliberate misdirection. The choice of target and the type of data taken fit the group's past interests, which strengthens the link. A careful analyst would conclude that the intrusion is probably connected to the known group, with moderate confidence, and would say clearly that the evidence does not establish who sponsors the group or whether the clues were planted.

That conclusion is enough for most practical purposes. The company can study the group's known techniques to look for other intrusions, strengthen defences against them and share indicators with peers. Naming a sponsoring state, if it ever happens, is a decision for governments with access to wider intelligence.

Confidence and competing hypotheses

Because attribution evidence is ambiguous, analysts benefit from methods that force them to consider alternatives. Listing several possible perpetrators, including the possibility of a false flag, and asking which pieces of evidence are consistent or inconsistent with each, helps prevent premature closure on the first plausible suspect. Heuer (1999) developed this approach, analysis of competing hypotheses, for exactly the kind of problem in which evidence is incomplete and deception is possible. Recording the reasoning also allows later reviewers to see how the judgement was reached and to revise it if new evidence appears.

Attribution as a political act

Egloff (2020) argues that public attribution is not simply the release of a technical finding but a political decision. Governments choose whether to name an attacker, when, with what evidence and with which allies, considering diplomatic consequences, the protection of sources and the message sent to other potential attackers. In recent years many governments have shifted from silence towards coordinated public attributions, often made jointly by several states.

International law adds a further dimension. Under the law of state responsibility, a state can be held responsible for cyber operations conducted by its organs or by groups acting under its control, but establishing that link requires attribution to an appropriate standard. Schmitt (2017), in the second edition of the Tallinn Manual, sets out how international law experts interpret these rules for cyber operations.

Attribution also intersects with information operations. Rid (2020) describes how stolen material has been leaked selectively to influence public debate, so that analysts must consider not only who hacked a target but who is spreading the results and to what end.

Lessons for organisations

Most organisations will never attribute an attack to a state. But the principles of cyber attribution apply to everyday threat intelligence:

1. **Collect and preserve evidence.** Good logging, retained over time, makes it possible to reconstruct what happened. Without it, both defence and attribution are guesswork.
2. **Focus on behaviour, not just indicators.** Tounsi and Rais (2018) note that technical indicators such as IP addresses change quickly; an adversary's preferred techniques change slowly and are more useful for defence.
3. **State confidence explicitly.** Attribution judgements should carry clear levels of confidence and acknowledge alternative explanations, as in other intelligence analysis (Heuer, 1999).
4. **Consider deception.** Ask what evidence an adversary would plant if it wished to mislead, and whether the case depends on clues that would be easy to fake.
5. **Share carefully.** Wagner et al. (2019) show that sharing threat information multiplies its value but requires trust and clear rules. Industry groups and national cyber security centres provide channels.
6. **Remember the purpose.** For most victims, knowing how the attacker operates and what it wanted matters more for defence than knowing exactly who it was.

Conclusion

Cyber attribution sits at the meeting point of technical forensics, intelligence analysis and politics. Clark (2020) argues that intelligence analysis should build an evolving model of its target; in cyber attribution that model links technical artefacts to campaigns, groups and sponsors through layers of increasingly uncertain inference. Omand (2010) emphasises that public trust in intelligence depends on honesty about its limits. Attribution can be done well, but rarely with certainty, and the best practice is to be clear about the evidence, the confidence it supports and the possibility of deception.

References

Buchanan, B. (2020). *The hacker and the state: Cyber attacks and the new normal of geopolitics*. Harvard University Press.

Caltagirone, S., Pendergast, A., & Betz, C. (2013). *The diamond model of intrusion analysis* (Technical Report ADA586960). Center for Cyber Intelligence Analysis and Threat Research.

Clark, R. M. (2020). *Intelligence analysis: A target-centric approach* (6th ed.). CQ Press.

Egloff, F. J. (2020). Public attribution of cyber intrusions. *Journal of Cybersecurity, 6*(1), Article tyaa012. https://doi.org/10.1093/cybsec/tyaa012

Heuer, R. J., Jr. (1999). *Psychology of intelligence analysis*. Center for the Study of Intelligence, Central Intelligence Agency.

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research, 1*(1), 80–106.

Landau, S. (2013). Making sense from Snowden: What's significant in the NSA surveillance revelations. *IEEE Security & Privacy, 11*(4), 54–63. https://doi.org/10.1109/MSP.2013.90

Lemay, A., Calvet, J., Menet, F., & Fernandez, J. M. (2018). Survey of publicly available reports on advanced persistent threat actors. *Computers & Security, 72*, 26–59. https://doi.org/10.1016/j.cose.2017.08.005

Libicki, M. C. (2009). *Cyberdeterrence and cyberwar*. RAND Corporation.

Lowenthal, M. M. (2022). *Intelligence: From secrets to policy* (9th ed.). CQ Press.

Mandiant. (2013). *APT1: Exposing one of China's cyber espionage units*. Mandiant.

Omand, D. (2010). *Securing the state*. Hurst.

Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies, 38*(1–2), 4–37. https://doi.org/10.1080/01402390.2014.977382

Rid, T. (2020). *Active measures: The secret history of disinformation and political warfare*. Farrar, Straus and Giroux.

Schmitt, M. N. (Ed.). (2017). *Tallinn manual 2.0 on the international law applicable to cyber operations*. Cambridge University Press. https://doi.org/10.1017/9781316822524

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). *MITRE ATT&CK: Design and philosophy* (Technical Report MP180360). The MITRE Corporation.

Tounsi, W., & Rais, H. (2018). A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Computers & Security, 72*, 212–233. https://doi.org/10.1016/j.cose.2017.09.001

Wagner, T. D., Mahbub, K., Palomar, E., & Abdallah, A. E. (2019). Cyber threat intelligence sharing: Survey and research directions. *Computers & Security, 87*, Article 101589. https://doi.org/10.1016/j.cose.2019.101589