A Study of Modern Cyber Attribution & Focus on Theoretical Improvements

CSIS 461

Professor Bansley

James Riley Dorough

Liberty University

October 31st, 2021

**Abstract**

With the rapid growth of cyber-attacks sourced from nation-state-supported threat actors; the proper attribution and documentation of actors’ tactics, techniques, and procedures (TTP) is essential. When the United States suffers a physical attack (i.e., terrorism) we immediately begin collecting evidence, information, and forensics to determine the acting organization responsible. Defense against cyber-attacks should not be and generally is not much different. However, the digital battlefield offers a fog of war unlike any available in the physical domain. Bad actors can mask activity, reroute the flow of information, falsify geographic location, shield themselves under sanctions of various governments, and even mimic signatures and TTPs of other nation-state actors. As tensions rise over the digital frontier defenders must constantly improve their threat attribution capabilities to ensure that entities are not falsely accused and minimize the chance of mitigation or retaliation disproportionate to the original malicious activity.

**Introduction**

One of the most difficult and extensive processes in defensive cybersecurity and threat actor research is proper, accurate attribution. Much like physical crime or assault, the attacking entity should be held responsible. However, unlike physical crime, digital crime, warfare, and hacktivism all possess multiple layers of abstraction. Malicious actors can mask their activity or spoof other attributes of different actors to disrupt an investigation. This stresses the importance of ethical, legal, and effective means of investigation into the responsible group.

Of these groups, many fall under the classification of advanced persistent threat (APT). These groups are delineated from other hacker groups due to the level of their capabilities and the resources provided to them. Many APTs are nation-state sponsored. They are funded and supported by organized governments often in pursuit of compromising non-allied nation’s businesses, infrastructure, or governmental structure. Their capabilities are often enhanced by either high technical proficiency through selective recruitment or ease of target material for research and development of cyber weapon payloads. Because these groups are such a threat in modern cyberspace, equal or superior defensive teams are required. Attribution is a skill that must be homed in these defensive groups. This work delves into the modern methods of threat recognition and attribution, current progress being made to improve these capabilities, and upcoming or theoretical tools that can give defenders an edge on their competition in the digital battlefield.

**Modern methods of attribution**

Two of the most common methods of attribution in industry today are data analysis and digital forensics. A broad understanding of forensics is to trace effects on target to recordable TTPs and from here the best estimate of the threat actor or APT responsible for an attack. This method does not always guarantee 100% certainty who the responsible actor is. However, great progress has been made in this field with the inclusion of better reporting and more sources of information pertaining to compromise. In this age of security, we are now witnessing private vendors and the U.S. government begin to understand the severity of insecure systems, poor event logging, and inaccurate threat attribution of the source.

Through a rise in cybercrime and its realized ability to hide a threat actor, the Cybersecurity Information Sharing Act (CISA) was passed through congress. The CISA ensures threats are handled by the discovering organization and/or its security vendor and are exposed either to the government or from the government to private entities (Noor et al., 2019, p. 228). This bill provides vendors the authority to defend their networks and, with written consent, the networks of other government or private entities. Classified documents are permitted to be shared only with authorized private vendors while unclassified indicators of compromise (IOC) are permitted to be shared publicly (U.S. Congress, 2015). To improve attribution and cyber defense, many open source and publicly accessible threat information, attack signature, and TTP repositories should be utilized by organizations to recognize probable APTs (Maglaras et al., 2018). Many organizations currently partner with security firms who have experience and access to known IOCs and TTPs from threat actors.

Some organizations handle their security and incident response without outsourcing it as a service. When an attack or breach occurs, this method of in-house security requires an incident response team to handle reporting, mitigation, and attribution themselves. Because of this, unclassified reports of an incident can be found in unstandardized or unstructured reports from vendors in the form of white papers (Noor et al., 2019, p. 228). This should only be expected if such an organization has contained or stopped the threat and their documentation does not expose any classified information from the organization.

The more recent understanding of the importance of digital security has flipped industry from a mindset of good enough security and high accessibility to a mindset of least privilege for end users. However, we often find that critical infrastructure is not in a viable position to simply upgrade or tear out and replace with a more secure system design. One such example is an Industrial Control System (ICS). ICSs are often the target of cyberattacks due to their focus on availability and consistency rather than security which can mildly hamper efficiency or availability (Maglaras et al., 2018, p. 2). Because of this weakness and lethal result of compromise, effective attribution is important to understand adversaries and their TTPs.

Models such as Maglaras’s vulnerability assessment cycle have been created To assist in decision making and information gathering efforts to assist attribution.

Diagram

Description automatically generated

Fig. 1. Procedures of vulnerability assessment of cybersecurity for critical infrastructures (2018)

Using a threat model and understanding an organization’s vulnerabilities and hosted services (e.g., SCADA, ICS, Web Hosting, or Finances) can provide a narrower scope of possible malicious threat actors (Maglaras et al., 2018, p. 2). Thusly, we can better expect a nation-state actor or APT to target industrial systems and public infrastructure such as gas lines. We wouldn’t expect ransomware actors to go after these systems. Instead, it would be much more likely that a financially motivated actor would target the companies and businesses supplying fuels to these ICS pipelines or the banks housing the finances of these organizations. The point is that an understanding of reasonable means and opportunities of an organization should be one of if not the first step in illuminating possible adversaries while ruling out unlikely candidates.

**Current improvements and near-future changes**

One tool that is becoming more prevalent in some specific instances of threat modeling is the Structured Threat Information Expression (STIX) model. This model allows for the compilation of threat information from a network to ascertain threat actors and log the events of a breach. Pictured below is an example of attribution through STIX-reported forensics (Noor et al., 2019, p. 229).



Fig. 2. DNC email hack threat actor attribution based on common patterns in high-level IOC. (2019).

The abstraction STIX provides to information retrieval and processing is a great help to responders attempting to implement fixes quickly with C-suite members and shareholders being able to understand how the threat applies to the future operation of their organization. However, the current implementation of STIX is often utilized manually by analysts. This means that reports of similar attacks from the same actor can result in different reports depending on the skill level and accuracy of the reporting entity. Thusly, STIX is a useful new model for simplifying and connecting actions to threats. However, it does not apply to all networks equally and the abstraction of information can result in inconsistent or incomplete cyber threat incident (CTI) reports (Qamar et al., 2017, p. 36). More work in the normalization of reporting needs to be done to allow STIX to apply across various network structures and domains.

As with the development and improvement of STIX, there also exists improvement in the publicly accessible side of CTI gathering and storage. We have seen proprietary logging, network security, and monitoring as a service. However, recently, there has been a great increase in commonly accessible threat information, TTP, and breach reports made publicly available through many sources. Some groups hosting this information are open source while others are business vendors offering some information free while offering paid services to act on this information within an organization. Regardless, an increase of sharing CTI pertaining to IOCs can be the difference in an organization changing their network monitoring and intrusion detection system (IDS) rules to catch a possible attack early enough to mitigate it. This process still relies on a human’s perspective of the importance or severity of threat vectoring and security holes in a system. Because incident handing is often a highly stressful process, it is prone to many instances where human error, a form of bias, or early decision making may impact the end belief, reporting, mitigation efforts, and monitoring (Hettema, 2021). While the human factor of threat intelligence and attribution is critical; bias in any form can skew how an incident is handled.

Additionally, a major cause for the delay in awareness and response to breaches is that reporting often takes the form of human-generated documentation and is not automated or formatted to a standard of documents (Noor et al., 2019, p. 228). Due to this, one potential method to supplement human error in some form is to automate some steps fully or partially in the attribution process.

**Theoretical improvements for threat attribution**

Incident responders continuously strive to better cyber capabilities. Among the many options for improvement, a few are focused on in this section. Primarily, improvements to data sharing and normalization of threat information, legal self-defense policy for victims of cyber-attacks, and the improvement of artificial intelligence to compensate for human faults in data handling. There may be a grey area between current methods of attribution in the previous, near-future changes, and this theoretical improvement section. The main delimitation is the difference in acceptance and compatibility between various organizations.

The Trusted Automated eXchange of Intelligence Information (TAXII) standard framework was developed to assist in automation efforts for the dissemination of vital threat or attack information. This framework is designed to support the previously discussed STIX model. However, this new capability of automated data acquisition and reporting allows for much more complicated and proactive threat awareness. The API formatting STIX reports and being transferred and stored by TAXII opens the possibility for integration of additional capabilities. The ability to incorporate IDSs with forensics methodology in an automated format could provide a highly optimized model to deliver actionable evidence to incident responders.

Where IOCs are present and known by digital security teams; containment and mitigation efforts should always take effect. However, there are rare cases of organizations, often smaller and non-multinational, reaching out past their network to the adversary. This often is either to gather additional information from the assailant or actively defend the organization. Self-defense from a cyber-attack in the form of some offensive capability to degrade the perceived adversary to halt an attack is commonly known as *hack back*. The main reason that hack back is illegal in most countries is a threat actor may be spoofing its perceived source of the attack (Finlay & Payne, 2019, p. 205-206). However, there is a possibility that holding perceived sources partially responsible for an attack can promote transparency to benefit threat attribution efforts (p. 206). If in the future nations can agree on a standard of responsibility between a victim and traced source, possibly an internet service provider (ISP), there should be an increased willingness to provide information to the victim from the ISP. Even if the motivation for transparency from an ISP is only self-preservation, proper attribution of the true responsible party can be better exercised. This is a highly unstable concept and would take enormous restrictions and legal boundaries to ensure that all involved parties are justified in their actions and responsibility.

A large benefit to attribution can be found in removing human actors from steps of the information-gathering process at some variance. As we look to the future, we can expect repositories of TTPs and APT signatures to grow. With this growth, the human aspect of information retrieval, handling, and organization may begin to become more automated. We often hear artificial intelligence or machine learning (AI/ML) as buzzwords in nearly any industry. However, with such increasing pools of pertinent information, humanity may begin implementing even more dynamic processing approaches to the cyber threat attribution process.

**Conclusion**

The topic of threat attribution in cyber is an immensely complex one that is not easily accomplished properly after a cyber incident occurs. There are many methods of information retrieval available to an organization. Deciding what security tools to implement, what reporting tools to rely on, and if the security domain should be outsourced to a private vendor must be fully understood if any organization intends to survive and recover from an attack. Localized security is beneficial and should never be overlooked. However, if the source of an attack and the motivation behind the attack is not researched and documented, an organization is likely to fail to properly alter its approach to security to mitigate future attacks.

Among the many ways to attribute an attack to the actor, research of public documentation with a focus on information sharing between vendors appears to be the most prominent way to patch an entity's system. I find this lines up well with the definition of wisdom. Should an entity, unfortunately, experience a breach in security, they can share what they learned. In turn, other entities can work on a methodology to recognize and impede adversaries and share their approaches without having to suffer the same attack. However, this is only a reactive approach.

As automated reporting systems and IDSs improve, we can expect to see better tracking of TTPs and threat actor signatures to quickly discern and offer mitigation plans before an adversary completes an attack. If an incident response team can be notified of malicious behavior while an attack is just beginning, they may be able to stop it before the adversary can pivot within the network. To assist with this, we can expect models like STIX/TAXII to allow for abstracted threat reports to be disseminated quickly to parties involved or responsible for containment. If this can be paired with a machine learning module. These reports can be, in theory, read by the AI and proper mitigation steps can be presented or acted on long before the attack causes any real damage.

Even if these highly advanced systems are either completed or even initialized, it will always be the responsibility of security experts to recognize, track, monitor, and attribute attacks. Incident responders and data analysts must strive to stay ahead of the never-ending string of attacks on information systems and infrastructure daily. The only thing that keeps critical infrastructure running and recoverable is committed experts who focus on proper research, incident reporting, and technical skills for the systems they are responsible. The bad actors won’t give up and the good guys must be far more aggressive and determined to stop these persistent threats.

**Acronyms**

AI – Artificial Intelligence

API – Application Programming Interface

APT – Advanced Persistent Threat

CISA – Cybersecurity Information Sharing Act

CTA – Cyber Threat Actor

CTI – Cyber Threat Intelligence

DDoS – Distributed Denial of Service

ICS – Industrial Control Systems

IDS – Intrusion Detection System

IOC – Indicator of Compromise

IR – Incident Response †

IR – Information Retrieval †

ISP – Internet Service Provider

ML – Machine Learning

STIX – Structured Threat Information Expression

SCADA – Supervisory Control and Data Acquisition

† *Where observed in this document will be spelled out to avoid confliction & confusion*

TAXII - Trusted Automated eXchange of Intelligence Information

TTL – Time To Live

TTP – Tactics Techniques and Procedures

**References**

Finlay, L., & Payne, C. (2019). The Attribution Problem and Cyber Armed Attacks. *AJIL Unbound*, *113*, 202–206. https://www.proquest.com/docview/2245993625/fulltextPDF/975610009FFF4AC2PQ/1?accountid=12085

Hettema, H. (2021). Rationality constraints in cyber defense: Incident handling, attribution and cyber threat intelligence. *Computers & Security*, *109*, 1–13. <https://www.sciencedirect.com/science/article/pii/S0167404821002200?via%3Dihub>

Maglaras, L., Ferrag, M., Derhab, A., Mukherjee, M., Janicke, H., & Rallis, S. (2018). Threats, Countermeasures and Attribution of Cyber Attacks on Critical Infrastructures. *EAI Endorsed Transactions on Security and Safety*, *5*(16), 1–10. https://www.proquest.com/docview/2306299509?accountid=12085&pq-origsite=summon

Noor, U., Anwar, Z., Amjad, T., & Choo, K. K. R. (2019). A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise. *Future Generation Computer Systems*, *96*, 227–242. https://www.sciencedirect.com/science/article/pii/S0167739X18326141?via%3Dihub

Qamar, S., Anwar, Z., Rahman, M. A., Al-Shaer, E., & Chu, B. T. (2017). Data-driven analytics for cyber-threat intelligence and information sharing. *Computers & Security*, *67*, 35–58. <https://www.sciencedirect.com/science/article/pii/S0167404817300287?via%3Dihub>

U.S. Congress. (2015, October 28). *S.754 - 114th Congress (2015–2016): To improve cybersecurity in the United States through enhanced sharing of information about cybersecurity threats, and for other purposes.* Congress.Gov | Library of Congress. Retrieved November 25, 2021, from https://www.congress.gov/bill/114th-congress/senate-bill/754