

From IOCs to Group Profiles: On the Specificity of Threat Group Behaviors in CTI Knowledge Bases

Aakanksha Saha
TU Wien
aakanksha.saha@seclab.wien

Martina Lindorfer
TU Wien
martina@seclab.wien

Juan Caballero
IMDEA Software Institute
juan.caballero@imdea.org

Abstract—Indicators of Compromise (IOCs) such as IP addresses, file hashes, and domain names are commonly used for threat detection and attribution. However, IOCs tend to be short-lived as they are easy to change. As a result, the cybersecurity community is shifting focus towards more persistent behavioral profiles such as the Tactics, Techniques, and Procedures (TTPs) and the software used by a threat group. However, the distinctiveness and completeness of such behavioral profiles remain largely unexplored. In this work, we systematically analyze threat group profiles built from two open cyber threat intelligence (CTI) knowledge bases: MITRE ATT&CK and Malpedia. We first investigate what fraction of threat groups have group-specific behaviors, i.e., behaviors used exclusively by a single group. We find that only 34% of threat groups in ATT&CK have group-specific techniques. The software used by a threat group proves to be more distinctive, with 73% of ATT&CK groups using group-specific software. However, this percentage drops to 24% in the broader Malpedia dataset. Next, we evaluate how group profiles improve when data from both sources are combined. While coverage improves modestly, the proportion of groups with group-specific behaviors remains under 30%. We then enhance profiles by adding exploited vulnerabilities and additional techniques extracted from more threat reports. Despite the additional information, 64% of groups still lack any group-specific behavior. Our findings raise concerns on the belief that behavioral profiles can replace IOCs in threat group attribution.

1. Introduction

Indicators of Compromise (IOCs) – such as malicious IP addresses, file hashes, domain names, emails, and cryptocurrency addresses – are widely used for detecting and attributing threats but offer only a snapshot of an adversary’s activity. IOCs are often ephemeral and easily changed by threat actors, limiting their long-term effectiveness for threat detection and attribution. To address these shortcomings, prior work argues to instead focus on behavioral characteristics, such as the Tactics, Techniques, and Procedures (TTPs) used to gain access, move laterally, maintain persistence, and exfiltrate data [4], [6], [21], [38], [41], [44], [50].

Behavioral characteristics are thought to be more robust, remain more stable over time, have a larger cost for adversaries to change them, and be able to link seemingly unrelated attacks from the same threat group. Models like the Pyramid of Pain [5] place TTPs at the top of the pyramid in terms of cost for an adversary to change them. Apart from TTPs, other behavioral characteristics also exist, for example, the software tools used by a threat group may be distinctive, particularly if the malware is developed in-house. Similarly, exploited vulnerabilities can be characteristic, especially when a group targets uncommon software or uses custom-developed exploits. Even the textual content used in campaigns can be characteristic of a threat group with recent works leveraging phishing SMS contents [32], ransomware notes [48], and cross-file-type features [40] to identify attacks from the same campaign and threat group.

We refer to a threat group’s observed behaviors as its *group profile*. Accurate threat group profiles are fundamental for incident correlation, attribution, building behavioral detection rules, and proactive threat hunting. Group profiles can be built from the contents of the threat reports published by security vendors and analysts, which typically describe, in natural language, the analysis of specific attacks and their attribution to specific threat groups. Threat reports may come from a single source (e.g., a specific cybersecurity vendor) or be aggregated by cyber threat intelligence (CTI) knowledge bases [34], [45] and sharing platforms [7], [21].

While the potential of such behavioral profiles has been widely acknowledged, few studies have examined how distinctive the behavioral profiles of threat groups truly are, and how complete our understanding of those behaviors is, especially given the varying quality and scope of the data sources used to build these profiles. One concern is that many behaviors in these profiles can be *generic*, i.e., used by many threat groups, thus providing little information on the groups using them. These include common techniques (e.g., spearphishing, malware auto-start through registry keys), widely available software (e.g., abused penetration testing tools, open-source projects, malware kits sold in underground forums), and prevalent vulnerabilities (e.g., those affecting popular software with public exploits). Multiple threat groups often acquire such tools and exploits for reasons of convenience, reduced operational cost, or to obscure attribution.

This raises the question of which behaviors are truly *group-specific*, i.e., used by only a single threat group. Group-specific behaviors, when observed on a protected system, can serve as behavioral signatures that uniquely identify the responsible threat group. However, what fraction of threat groups exhibit group-specific behaviors remains an open question. Determining whether a behavior is truly group-specific requires not only analyzing the group that exhibits it, but also having comprehensive coverage of behaviors across other threat groups. Without this broader context, a behavior might appear unique when it is not. For example, as more threat reports become available, a behavior initially believed to be exclusive to group A may also be observed in group B, indicating it is not unique.

In this work, we perform a systematic analysis of group profiles in CTI knowledge bases, their complementary value, and how they can be extended. We focus on knowledge bases that are open (i.e., non-commercial), index many threat reports, are periodically updated, provide a taxonomy of threat groups, and organize information into group profiles. These profiles include basic group metadata (e.g., name, aliases, country), references to related threat reports, and descriptions of group behaviors mentioned in those threat reports. We find two knowledge bases satisfying those properties: MITRE’s Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) [27] and Malpedia [15]. We discard other open projects like the MISP threat actor galaxy [26], ThreatMiner [46], and APTnotes [3], as they collect threat reports and associate them with threat groups but do not extract or include behavioral information from those reports into their group profiles. Moreover, Malpedia incorporates data from the MISP threat actor galaxy, allowing us to cover that project indirectly. We also exclude commercial services that provide specialized threat reports to paying customers [6], as their knowledge bases are proprietary. Both ATT&CK and Malpedia provide their own taxonomies of threat groups and associated software tools. A key distinction is that ATT&CK also includes a taxonomy of TTPs, integrating attack techniques directly into the group profiles. Using the data in these two knowledge bases, we analyze the following three research questions:

RQ1: What fraction of the threat groups in ATT&CK and Malpedia have group-specific behaviors? We separately analyze the group profiles created using only the information available in ATT&CK and Malpedia. Identifying groups based on TTPs is challenging as only 52 (34.2%) groups in ATT&CK have group-specific techniques. It is somewhat easier to identify groups through the software they use with 111 (73.0%) groups in ATT&CK having group-specific software. However, this percentage is significantly lower in Malpedia, where only 192 (24.0%) have group-specific software. This discrepancy stems from threat reports disproportionately focusing on a subset of high-notoriety groups, leaving less-known groups with sparse reports to build their profiles. Combining techniques and software into joint profiles increases the groups with group-specific behaviors from 111 (73.0%) to 124 (81.6%).

RQ2: How complementary is the information in both datasets? How much do group profiles improve when combining data from both sources? Both datasets differ substantially in volume. Malpedia provides a broader coverage of the threat landscape than ATT&CK, comprising 5.2 times more threat groups (800 vs. 152) and 4.2 times more software (3,367 vs. 794). This expanded scope stems from Malpedia indexing 16.9 times more threat reports (15,699 report URLs vs. 930). To assess their overlap, we normalize group and software names across the datasets. Both datasets have little overlap with only 145 groups and 498 software entries in common. The corresponding Jaccard Index values are 17.7% for groups, 13.5% for software, and just 3.2% for report URLs. The low intersection indicates that each dataset captures a different view of the threat group landscape highlighting their complementary nature.

We create joint group profiles using the data from both datasets, identifying 236 (29.2%) groups with group-specific behaviors, compared to 124 groups using only ATT&CK and 192 using only Malpedia. Despite combining both datasets, over 70% of groups have no group-specific behavior.

RQ3: What additional information currently not in ATT&CK and Malpedia could make threat group profiles more complete? We examine how group profiles can be improved with additional information extracted from threat reports. First, we extract CVE identifiers to build vulnerability profiles for each threat group. The number of groups with at least one group-specific vulnerability is 48 (31.6%) in ATT&CK, 112 (14.0%) in Malpedia, and 119 (14.7%) when combining both datasets. Thus, exploited vulnerabilities tend to be less distinctive than the software used, but more than the techniques used. Next, we extend the group profiles with additional technique identifiers extracted from the threat reports. Since Malpedia does not provide TTPs, this step allows us to extend its group profiles with techniques. Incorporating these extracted techniques increases the number of groups with group-specific behaviors from 52 (6.4%) to 68 (8.4%). Finally, we combine all available behavioral indicators, including techniques, extracted techniques from reports, software, and vulnerabilities, into unified group profiles, identifying 291 (36.0%) groups with at least one group-specific behavior. Despite leveraging all available information, a majority of groups (64%) have no group-specific behaviors.

To better understand the limitations of current group profiles, we also discuss the impact of under-reporting, i.e., incomplete coverage of threat group behaviors. We observe that the number of technique identifiers extracted from the ATT&CK threat reports is larger than the number of techniques officially cataloged in ATT&CK from those same reports. This discrepancy likely arises from the manual nature of the report analysis process by ATT&CK contributors, emphasizing the need for automated approaches to extract TTPs from threat reports [2], [18], [37]. We also observe that only 46.3% of techniques and 64.1% of software entries in ATT&CK, and just 28.6% of software in Malpedia, are currently associated with at least one threat group. The re-

maining entries were likely added to the taxonomies because they were observed being used by adversaries in the wild. However, their lack of association with specific threat groups highlights the incomplete coverage of group profiles.

Artifacts. We will open-source our code and data at <https://anonymous.4open.science/r/ThreatGroupCTI-B746>.

2. Dataset Comparison

This section first details the information in ATT&CK [27] and Malpedia [15] and then sets the base for answering RQ2 by analyzing the extent of data overlap between the two datasets and assessing how their contents complement each other.

2.1. Datasets

ATT&CK. ATT&CK provides taxonomies of offensive and defensive techniques, software tools used by adversaries, and threat groups. The techniques taxonomy comprises three *domains*: Enterprise, Mobile, and ICS. Each domain defines a set of *tactics* that correspond to different steps in the kill chain, such as Reconnaissance (TA0043), Persistence (TA0003), and Lateral Movement (TA0008). Each tactic includes a set of *techniques*. For example, Active Scanning (T1595) and Phishing for Information (T1598) are techniques under the Reconnaissance tactic. Techniques can also contain *sub-techniques*. For example, T1204.002 corresponds to the Malicious File sub-technique under the User Execution (T1204) technique.

The threat group taxonomy covers nation-state actors, advanced persistent threats (APTs), and some large for-profit actors such as ransomware groups. The profile of a threat group contains a unique identifier, a name, a list of aliases (called *Associated Groups*), and the techniques and software used by the threat group.

Entries in the software taxonomy are categorized into Tools and Malware. Tools include commercial software (e.g., Cobalt Strike), open-source frameworks (e.g., Metasploit, Mimikatz), and built-in OS tools (e.g., PsExec, ipconfig). Malware includes families specific to a single threat group (e.g., Carbanak) as well as malware kits available in underground markets and used by multiple threat groups (e.g., PoisonIvy RAT). The focus is on software used by APTs listed in the group taxonomy; however, ATT&CK also catalogs non-APT malware such as the Conficker worm [29] and the SimBad Android malware [30]. Each taxonomy entry contains URLs to threat reports related to the entry, such as reports describing the techniques and software used by a threat group.

Since its public release in 2015, ATT&CK publishes a new version approximately every six months, with the latest version at the beginning of this work being 15.1, released in April 2024. Each version may add new taxonomy entries (e.g., groups, techniques, software), remove *revoked* entries, or mark entries as *deprecated* (i.e., to be revoked soon).

TABLE 1: Dataset summary. Malpedia does not have a techniques taxonomy. The low intersection and Jaccard Index show that both datasets have little overlap. We use the union of both datasets to build group profiles.

Data	ATT&CK	Malpedia	\cap	\cup	Jaccard
Groups	152	800	145	807	17.7%
Techniques	839	-	-	839	-
Software	794	3,367	498	3,663	13.5%
Report URLs	930	15,699	522	16,107	3.2%
Report FQDNs	218	2,002	194	2,026	9.6%
Reports	920	14,983	80	15,816	0.5%

Malpedia. Malpedia provides taxonomies of threat groups and software. It does not provide a taxonomy of techniques nor reference the techniques in ATT&CK. Similar to ATT&CK, the threat group taxonomy focuses on APTs and nation-state actors, whereas the software taxonomy aims to cover any malware family, regardless of whether it is used by APTs or other types of attackers (e.g., for-profit actors). The software taxonomy also includes a few security tools (e.g., Cobalt Strike) but does not differentiate between malware and tools. Each taxonomy entry for a threat group or software includes URLs of threat reports related to the entry. We collect information about groups and software through the Malpedia API and metadata about threat reports (e.g., URL, title, author, publication date) from the provided BibTex file. Malpedia is updated daily by adding new bibliographic references labeled with the associated threat groups and software. We obtained Malpedia data February 18, 2025.

Dataset comparison. Table 1 summarizes the contents of both datasets. ATT&CK v15.1 contains 152 groups, 358 techniques, 481 sub-techniques, 794 software entries, and 930 URLs of threat reports from which those associations are extracted. Of the 358 techniques, 121 (33.8%) have at least one sub-technique, while 237 (66.2%) do not have sub-techniques. Among the total 839 techniques and sub-techniques, 637 (75.9%) belong to the Enterprise domain (202 techniques and 435 sub-techniques), 119 (14.2%) to Mobile (73 techniques and 46 sub-techniques), and 83 (9.9%) to ICS (83 techniques and no sub-techniques). For simplicity, in the remainder of this paper, we use the term *techniques* to refer to the combined set of 839 techniques and sub-techniques.

In contrast, Malpedia does not include techniques; however, it is much larger, containing 800 groups (5.2 times more), 3,367 software (4.2x), and 15,699 (16.9x) report URLs. The report URLs in ATT&CK come from 218 domains, compared to 2,002 (9.2x) domains in the Malpedia URLs, showing that Malpedia draws from a significantly more diverse set of sources (e.g., cybersecurity vendors and analyst blogs). We download the content of each URL, filter errors, and identify reports by the SHA256 of the downloaded content (most often an HTML page or a PDF document). In total, we downloaded 920 unique reports from the 930 ATT&CK URLs and 14,983 unique reports from 15,699 Malpedia URLs. We use these downloaded reports to extract additional information for extending the group

profiles discussed in Section 4.

2.2. Dataset Intersection and Union

This section examines the overlap between the two datasets and evaluates the benefits of combining their data. A key challenge in this comparison is that the names of groups and software differ across the datasets. To address this, we first created a mapping to align them.

Each knowledge base provides a name and a list of aliases for each threat group. We first normalized all names and aliases by converting them to lowercase, removing common suffixes such as “group” or “framework,” replacing terms like “team” with a space, converting terms like “threat group” to “TG,” and removing prefixes such as “TEMP”. Then, we compute the intersection between the set of names and aliases for each group in each taxonomy. If a group in ATT&CK shares a name or alias with a group in Malpedia, we merge them by performing the union of their sets. After the merging, we select a unique name for each normalized group. The selected name is the one used in ATT&CK by default and the one used in Malpedia if the group is not in ATT&CK. The normalization process identified 145 groups common to both datasets, seven groups unique to ATT&CK, 655 groups found exclusively in Malpedia, and 807 groups in the union of both datasets. We perform a similar normalization for software. We first normalize all software names and aliases by removing common prefixes (e.g., trojan, win, apk, elf) and replacing special characters (e.g., _rat to rat). Then we merge software entries that share at least one normalized name. The normalization identifies 498 software that appear in both datasets, 2,869 only present in Malpedia, 296 only present in ATT&CK, and 3,663 in the union of both datasets. We will publicly release our mappings of group and software names.

Table 1 also presents the overlap and union of report URLs, their fully-qualified domain names (FQDNs), and the SHA-256 hashes of the downloaded reports. We find only 522 report URLs shared between ATT&CK and Malpedia, resulting in a low Jaccard Index (JI) of 3.2%, indicating minimal overlap in referenced sources. The overlap based on actual report content is even smaller, only 80 downloaded reports have identical SHA-256 hashes, yielding a JI of just 0.5%. This discrepancy arises because downloading the same URL multiple times, especially in the case of HTML pages, can produce different files due to non-deterministic content such as dynamic metadata or embedded advertisements. Overall, the overlap between the datasets is quite low, with Malpedia providing a much broader view of the threat landscape. This disparity may be partly due to ATT&CK accepting contributions only from selected entities, which restricts the number of threat reports included in its analysis. However, this selective approach contributes to under-reporting. To address this limitation, we build group profiles by combining group and software information from the union of both datasets. Note that technique-level information is only available from ATT&CK and thus cannot be supplemented from Malpedia.

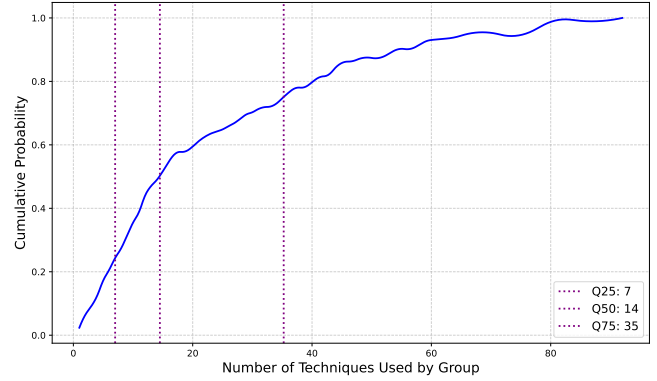


Figure 1: CDF of the number of techniques per group: 25% of groups use 7 or fewer techniques, 50% use 14 or fewer, and 75% use 35 or fewer.

Takeaway: Malpedia provides a larger coverage of the threat landscape, including 5.2 times more groups and 4.2 times more software than ATT&CK. While not a strict superset of ATT&CK, Malpedia covers 95.4% of ATT&CK’s groups and 62.7% of its software. Combining both datasets increases the overall coverage of the threat landscape.

3. Group Profiles in the Datasets

This section first addresses RQ1 by quantifying the proportion of threat groups in ATT&CK and Malpedia that have group-specific behaviors. It then addresses RQ2 by evaluating whether combining the datasets improves the group profiles.

3.1. Technique Profiles

In ATT&CK, the association of techniques to groups is provided as three separate group spreadsheets, one per domain. We combine the three group spreadsheets to obtain the set of techniques associated to each group, which we term the group’s *technique profile*.

We first measure the size of the technique profiles. Figure 1 shows the cumulative distribution function (CDF) of techniques per group. On average, each group uses 23.2 techniques. 38 (25%) have at most 7 techniques, 76 (50%) have between 7 and 36, and 38 (25%) have more than 35 techniques. The Lazarus Group (G0032) has the highest number of techniques, with 92 techniques. Four groups have no associated techniques and therefore cannot be identified through their TTPs. We next examine whether the remaining 148 groups contain group-specific techniques.

We build a mapping from each technique and sub-technique to the threat groups that use them. Among the 839 techniques cataloged in the ATT&CK framework, 388 (46.3%) have not been associated with any group, 147 (17.5%) are linked to a single group, 287 (34.2%) are associated to 2–37 groups, and 17 (2.0%) are used by at

TABLE 2: Top generic techniques, i.e., used by the largest number of groups.

ID	Technique Name	Groups
T1204.002	User Execution: Malicious File	79 (9.8%)
T1105	Ingress Tool Transfer	76 (9.4%)
T1566.001	Phishing: Spearphishing Attachment	72 (8.9%)
T1059.001	Command & Scripting: PowerShell	69 (8.5%)
T1588.002	Obtain Capabilities: Tool	66 (8.2%)
T1059.003	Command & Scripting: Win Command Shell	60 (7.4%)
T1036.005	Masquerading: Match Legitimate Name or Location	50 (6.2%)
T1547.001	Boot or Logon Autostart Execution: Registry Run Keys / Startup Folder	50 (6.2%)
T1071.001	Application Layer Protocol: Web Protocols	47 (5.8%)
T1082	System Information Discovery	46 (5.7%)

least one quarter (38) of all groups. The fact that 388 (46.3%) of all techniques in ATT&CK are not associated with any group raises concerns about coverage, as these techniques were presumably added to the taxonomy based on observed adversary behavior, yet remain unlinked to any known group.

Generic techniques. We call techniques used by many groups *generic*, as their presence in a protected environment offers limited value in distinguishing specific adversaries. Table 2 lists the top 10 techniques by number of groups. The most common technique is Malicious File (T1204.002) used by 79 groups where adversaries rely on users opening a malicious file, followed by Ingress Tool Transfer (T1105, 76 groups) where adversaries transfer tools or files from an external system into a compromised environment, and Spearphishing Attachment (T1566.001, 72 groups) where emails with a malicious attachment are used as a vector of initial compromise.

Additionally, we analyze which pairs of techniques tend to occur together. For this we compute the co-occurrence rate $\frac{|A \cap B|}{\max(|A|, |B|)}$ where A and B are the sets of groups using technique A and B , respectively. We find five pairs with a co-occurrence rate of at least 0.75. The highest rates are 0.951 between Malicious Link (T1204.001) and Spearphishing Link (T1566.002), followed by 0.886 for Malicious File (T1204.002) and Spearphishing Attachment (T1566.001). The four techniques in these two pairs are generic, each used by at least one quarter of the groups, and are also semantically related, where a spearphishing link is a type of malicious link, and the attachment in a spearphishing email is a malicious file that the user is encouraged to open.

Technique profile similarity. We compute the similarity of the technique profiles of each pair of groups using the Jaccard Index. The mean Jaccard Index is 0.07, the median is 0.06, and the maximum Jaccard Index is 0.55. Since most pairs have low similarity scores, it suggests that each group’s technique profile is quite unique. However, this may be due to a large number of possible techniques and limited visibility or incomplete data available in ATT&CK. We identified only 12 pairs of groups with a Jaccard Index larger than or equal to 0.4.

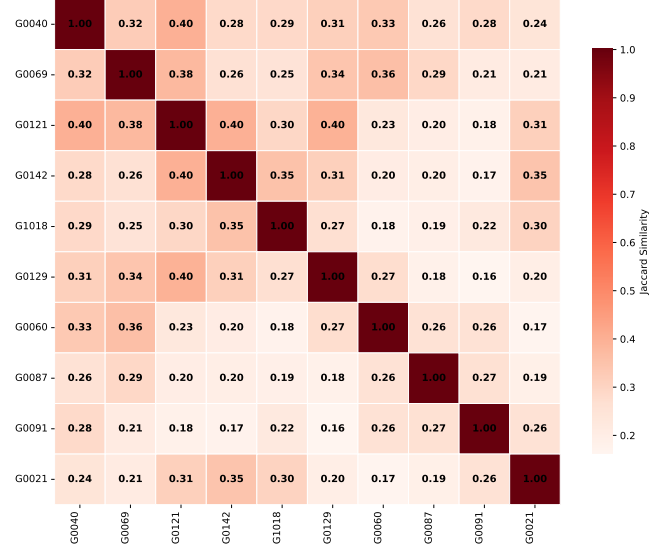


Figure 2: Jaccard Index between the 12 most similar groups (i.e., the ones with Jaccard Index ≥ 0.4). The mean Jaccard Index across all groups is 0.07, the median is 0.06, and the maximum Jaccard Index is 0.55.

Figure 2 shows a heatmap of these 12 group pairs. We observe that the moderate similarity between these pairs is often driven by generic techniques. For example, groups G0062 (TA459) and G0005 (APT12) each have five techniques in their profiles, sharing three, resulting in a Jaccard Index of 0.43. However, these three shared techniques are all generic techniques in Table 2: Exploitation for Client Execution (T1203), Spearphishing Attachment (T1566.001), and User Execution: Malicious File (T1204.002). This illustrates how generic techniques artificially increase group similarity.

Group-specific techniques. We call *group-specific* to the 147 (17.5%) techniques associated with a single group. Only 52 (34.2%) groups have group-specific techniques. The mean number of group-specific techniques is 0.99. While most of the techniques are commonly shared among groups, a few stand out by using distinct techniques, with the maximum number of group-specific techniques used by any group being 16 for Windshift (G0112).

A key question is whether these group-specific techniques appear unique because of limited coverage in ATT&CK, or if they truly represent capabilities developed or exclusively adopted by a single group. Table 3 provides examples of group-specific techniques. Some of these group-specific techniques appear indeed quite unique to their respective groups. For example, APT12 is the only group using DNS Calculation (T1568.003), where adversaries perform calculations on addresses returned in DNS results to determine which port and IP address to use for command and control. Conversely, some group-specific techniques may not be truly unique to their groups. For example, APT28 is the only group associated with Network Denial of Service (T1498), a fairly common attack technique likely

TABLE 3: Examples of group-specific techniques, some groups have multiple group-specific techniques.

Group Name	Technique ID	Technique Name
APT12	T1568.003	DNS Calculation
APT28	T1550.001	Application Access Token
	T1546.015	Component Object Model Hijacking
	T1001.001	Junk Data
	T1137.002	Office Test
	T1211	Exploitation for Defense Evasion
APT32	T1498	Network Denial of Service
	T1552.002	Credentials in Registry
APT37	T1564.004	NTFS File Attributes
	T1123	Audio Capture
APT38	T1562.003	Impair Command History Logging
	T1565.003	Runtime Data Manipulation
	T1565.001	Stored Data Manipulation
	T1565.002	Transmitted Data Manipulation
APT39	T1546.010	AppInit DLLs
	T1059.010	AutoHotKey & AutoIT
	T1056	Input Capture
APT41	T1596.005	Scan Databases
APT5	T1554	Compromise Host Software Binary
Axiom	T1563.002	RDP Hijacking
	T1001.002	Steganography
	T1553	Subvert Trust Controls
Chimera	T1110.004	Credential Stuffing
	T1556.001	Domain Controller Authentication
Cobalt Group	T1218.008	Odbcconf
DarkVishnya	T1200	Hardware Additions
Darkhotel	T1497	Virtualization/Sandbox Evasion

to be used by other groups at some point, suggesting this uniqueness may reflect limited coverage rather than actual exclusivity.

Takeaway: Only 52 groups (34.2%) have group-specific techniques. However, other groups may still be distinguishable by unique technique combinations, as seen by the low mean Jaccard Index of 0.06. Under-reporting remains a concern, as only 53.7% of ATT&CK techniques are observed in group profiles, and some seemingly group-specific techniques may not be truly unique.

3.2. Software Profiles

In this section, we explore how uniquely the software used by each group identifies it (see the classification of software in 2.1). For each group, we build three *software profiles* using the sets of normalized software names associated to the group in each dataset, and their union.

We first examine each dataset separately. Of the 794 software in ATT&CK, 509 (64.1%) are associated to at least one group. For Malpedia, the fraction is significantly smaller, where out of 3,367 software only, 963 (28.6%) are associated to at least one group. Software not associated to groups typically corresponds to non-APT malware. For example, the Conficker worm [29] and the Babuk ransomware [28] each appear in both ATT&CK and Malpedia and are not associated to groups in either dataset. The lower

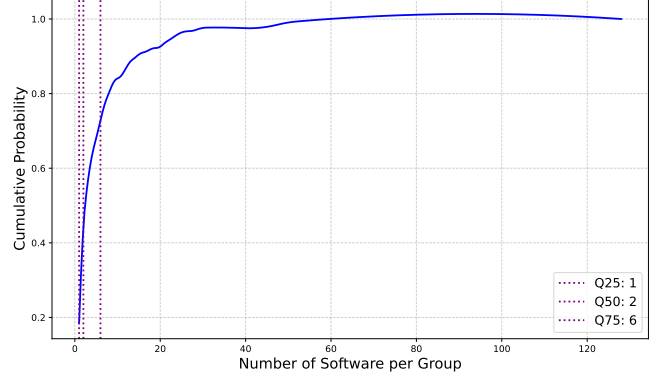


Figure 3: CDF of the number of software per group. 25% of groups used 1 or fewer software. 50% of groups used 2 or fewer software. Most groups used a relatively small number of software, with 75% using 6 or fewer.

TABLE 4: Top generic software by number of groups using them, their type in ATT&CK, whether they are in Malpedia, and the number and percentage of groups using them.

ID	Name	ATT&CK Type	Malpedia	Groups
S0002	Mimikatz	Tool	✓	46 (5.7%)
S0029	PsExec	Tool	✗	31 (3.8%)
S0039	Net	Tool	✗	30 (3.7%)
S0154	Cobalt Strike	Malware	✓	26 (3.2%)
S0013	PlugX	Malware	✓	25 (3.1%)
S0363	Empire	Tool	✗	15 (1.8%)
S0012	PoisonIvy	Malware	✓	14 (1.7%)
S0100	ipconfig	Tool	✗	13 (1.6%)
S0097	Ping	Tool	✗	13 (1.6%)
S0349	LaZagne	Tool	✓	12 (1.5%)

ratio of software associated to groups in Malpedia is likely due to Malpedia’s larger coverage of non-APT malware.

The fraction of groups with a non-empty software profile is also larger in ATT&CK where 138 out of 152 (90.8%) groups have associated software compared to 220 (27.5%) out of 800 groups in Malpedia (see Table 1 for the number of groups in each dataset). However, Malpedia contains 16.9 times more threat reports than ATT&CK, offering significantly more data for building software profiles. This difference arises because many smaller or lesser-known groups have few reports to support comprehensive profiling.

Next, we examine the unified software profiles. Out of the total 807 groups in both datasets, 264 (32.7%) groups have a non-empty software profile. Thus, over two-thirds of the groups cannot be identified by their associated software. Figure 3 shows the CDF of software per group for groups with at least one software. On average, each group uses 6.2 software but most groups used a relatively small number of software with 25% of groups have only one associated software and 75% using 6 or fewer. The APT38 group (G0082) has the highest count with 120 software.

Generic software. We call the software used by many groups *generic*, as their detection offers limited value in distinguishing specific adversaries. Table 4 lists the top 10

TABLE 5: Summary of group profiles across different data sources and combinations. The table reports the number of groups with non-empty profiles and the subset with at least one group-specific behavior. The top section presents results for profiles built using only techniques, only software, and a combination of both, as analyzed in Section 3. The middle section shows results from profiles enriched with extracted CVE identifiers and additional techniques from downloaded threat reports, discussed in Section 4. The bottom row shows the most comprehensive profiles, combining all available behavioral indicators.

Profile	ATT&CK		Malpedia		ATT&CK \cup Malpedia	
	Non-Empty	w/Group-Specific	Non-Empty	w/Group-Specific	Non-Empty	w/Group-Specific
Techniques	148 (97.4%)	52 (34.2%)	-	-	148 (18.3%)	52 (6.4%)
Software	138 (90.8%)	111 (73.0%)	220 (27.5%)	192 (24.0%)	264 (32.7%)	213 (26.3%)
Techniques \cup Software	151 (99.3%)	124 (81.6%)	220 (27.5%)	192 (24.0%)	331 (41.0%)	236 (29.2%)
Vulnerabilities	86 (56.6%)	48 (31.6%)	261 (32.6%)	112 (14.0%)	277 (34.3%)	119 (14.7%)
Techniques*	149 (98.0%)	60 (39.5%)	204 (25.5%)	69 (8.6%)	242 (30.0%)	68 (8.4%)
Tech* \cup Soft. \cup Vuln.	152 (100%)	128 (84.2%)	391 (48.9%)	265 (33.1%)	418 (51.7%)	291 (36.0%)

software by number of groups where the software appears in the group’s profile. All ten software appear in ATT&CK, while only five are present in Malpedia. The ones missing from Malpedia include four operating system tools (PsExec, Net, ipconfig, Ping) and the open-source Empire remote administration and post-exploitation framework [11]. Of these ten software entries, seven are classified as tools in ATT&CK and three as malware. Malpedia does not provide such classification. Among the three labeled as malware in ATT&CK, Cobalt Strike arguably should be categorized as a tool, as it is a commercial penetration testing package [14], while PlugX and PoisonIvy are remote administration tools (RATs) commonly available in underground markets. In summary, generic software typically refers to tools and malware kits that are either commercially sold or widely accessible.

We identify 88 software tools that are not marked as “tools” in ATT&CK but are used by multiple groups. These likely correspond to publicly available malware kits that are sold or shared in underground forums. In addition to Cobalt Strike, PlugX, and PoisonIvy (Table 4), other commonly reused malware include the gh0st RAT (used by 10 groups), China Chopper web shell (8), 8.t dropper (8), njRAT (7), and ShadowPad (7). Of these, gh0st is open-source [16], njRAT’s source code was leaked [13], China Chopper is publicly available [9], and both 8.t and ShadowPad have been reported to be privately shared among Chinese threat groups [25], [42]. In some cases, the groups using the same software may be related. For instance, Bistromath is reported by Malpedia to be used by both Lazarus Group and Silent Chollima, the latter being the subsidiary of Lazarus [24].

Group-specific software. We term software as group-specific if it has only been associated to one group and is not classified as a tool in ATT&CK. We exclude tools even if associated with a single group, because they can be easily adopted by others in the future, thus providing weak attribution. Of the 3,663 software across both datasets, 952 (26.0%) are associated with a single group, making them *group-specific* software. The detection of group-specific software may allow attributing the group behind an attack. Of the 807 total groups, 213 (26.3%) have at least one group-specific software associated with them. Among these 213 groups,

the mean and median number of group-specific software are 4.5 and 1.0, respectively, and 130 groups (16.1%) have only group-specific software. Some threat groups develop many custom tools for their attacks, for example, APT38 (G0082) has the highest number, with 99 group-specific software.

Takeaway: In ATT&CK, 111 groups (73.0%) have group-specific software, whereas in Malpedia, only 192 groups (24.0%) have that. Combining both datasets increases this number to 213 groups (26.3%), which still remains relatively low. Among the 264 groups with non-empty software profiles, the median number of software is 2, indicating that most threat groups operate with limited toolsets. However, some groups maintain extensive custom toolsets, for example, APT38 has 99 group-specific software.

3.3. Combining Group Profiles

Table 5 summarizes the number of groups with non-empty profiles and those with at least one group-specific entry across all the different profile-building methods examined in this work. The first three rows correspond to the profiles discussed in this section, based on techniques only, software only, and the combination of both.

To answer RQ1, the results show that identifying groups using techniques is challenging, as only 52 (34.2%) groups in ATT&CK have group-specific techniques. Identification is easier using software, with 111 (73.0%) of ATT&CK groups having group-specific software. However, this percentage is significantly lower in Malpedia, where only 192 (24.0%) of groups have group-specific software. The lower ratio in Malpedia likely reflects its focus on a subset of high-profile groups, while many smaller groups have too few reports to build robust profiles. Joint profiles combining both techniques and software improve identification in ATT&CK by 12.6 percentage points, increasing from 111 (73.0%) to 124 (81.6%) groups with group-specific behaviors.

To answer RQ2, the joint profiles built combining data from both datasets identify 236 (29.2%) groups with group-specific behaviors as compared to 124 groups using only ATT&CK and 192 using only Malpedia. Nonetheless, even

when combining techniques and software, over 70% of groups do not have any group-specific behavior.

4. Extending the Group Profiles

So far, the group profiles have included techniques and software from both ATT&CK and Malpedia. In this section, we address RQ3, i.e., whether we can extend the group profiles with additional data extracted from the downloaded threat reports. In Section 4.1 we build a *vulnerability profile* for each group with the vulnerabilities that the threat reports refer to as being used by the group in its attacks. Then, in Section 4.2, we discuss how we extend the technique profiles with additional technique identifiers extracted from the threat reports. This allows incorporating techniques mentioned from the threat reports in Malpedia and to examine how complete was the technique extraction in ATT&CK.

4.1. Vulnerability Profiles

The selection of which vulnerabilities to exploit is largely group-specific since it depends on the expected software used by the targets and the exploits the group has access to. The observation of specific vulnerabilities being exploited in a monitored system could potentially be used to attribute the group behind an attack. To build the vulnerability profiles, we first extract CVE vulnerability identifiers from the downloaded threat reports using a regular expression provided by the iocsearcher open-source tool [8], which can extract IOCs from text, HTML, PDF, and Word files. Then, we assign the CVEs to groups. ATT&CK reports are associated with a single threat group, whereas Malpedia reports may reference multiple groups. When a report mentions multiple groups, it is unclear which group the CVEs within the report should be assigned to. Therefore, we extract CVEs only from the 4,414 (29.4%) Malpedia reports referencing a single group, along with all 920 ATT&CK reports.

The left part of Table 6 summarizes the extraction of CVE identifiers from the downloaded threat reports. Among the 5,827 reports analyzed, 1,186 (20.3%) contain at least one CVE identifier for a total of 906 unique CVEs associated to 277 groups. Of the 807 groups, 277 (34.3%) groups have a non-empty vulnerability profile. The other 530 (65.7%) groups have no vulnerabilities that can be used to identify them. The mean CVEs per group is 8.9, and the maximum is 176 CVEs reported for APT28 (G0007).

Generic vulnerabilities. Overall, there are 368 vulnerabilities used by at least two groups, 114 used by more than five groups, and 28 used by more than 10 groups. Table 7 lists the top 10 CVEs by the number of groups using them. These generic vulnerabilities target popular software, with Microsoft Office being the most targeted with four vulnerabilities. Nine of the 10 vulnerabilities have publicly available proof of concept (PoC) exploits, either in the Exploit Database [12] or on GitHub. We did not find any PoC for CVE-2022-38028, which was a zero-day on the Windows Print Spooler used by Russian threat groups [17].

Group-specific vulnerabilities. Of the 906 CVEs identified in the reports, 538 (59.4%) are associated to a single group. We call these *group-specific vulnerabilities*. The *Vulnerability* row in Table 5 summarizes the generated profiles. Of the 807 groups, 119 (14.7%) have at least one group-specific CVE. Across these 119 groups, the mean and median vulnerabilities per group are 4.5 and 2.0, respectively. The maximum is for the Gorgon group (G0078), which uses 78 CVEs. Table 8 shows some example group-specific CVEs.

Takeaway: Only 119 (14.7%) groups have a group-specific vulnerability. The set of vulnerabilities exploited by a group is less unique than the set of software used (27.8% groups have group-specific software) likely because many groups focus on the same generic vulnerabilities affecting popular software, often with publicly available exploits. However, vulnerabilities tend to be more unique than techniques, as only 6.4% groups have group-specific techniques.

4.2. Technique Identifiers in Reports

In this section, we extend the technique profiles with explicit mentions of technique identifiers in the downloaded reports. It is important to note that threat reports may also include implicit references to techniques, such as stating that a rootkit was used without explicitly providing a technique identifier. We discuss the extraction of implicit references later in this section. To identify technique identifiers we use a regular expression provided by iocsearcher. We observe that the technique identifiers, if present, are typically provided in a table at the end of the threat report, although they may also appear throughout the text.

The right part of Table 6 summarizes the extraction of technique identifiers from the downloaded threat reports. From the 4,414 Malpedia reports uniquely assigned to one group, we find 626 unique technique identifiers associated to 204 groups appearing in 541 (12.2%) reports. Of these 626 techniques, 248 are not associated with any groups in ATT&CK, i.e., are only mentioned in the Malpedia reports. This shows that focusing on a small set of reports causes under-reporting and that technique profiles extracted from ATT&CK are likely to miss techniques used by a group.

Then, we apply iocsearcher to the 920 reports downloaded from ATT&CK reference URLs and identify 470 unique technique identifiers from 122 (13.2%) reports. These 122 reports are associated to 63 groups. ATT&CK has 451 techniques associated to 152 groups, while we find a larger technique set (470) mentioned for 63 groups in 13.2% of the same reports. This indicates that ATT&CK contributors may not be systematic in extracting all references of techniques in the report. Furthermore, we are only accounting for explicit references through technique identifiers. Threat reports may also include implicit references. However, extracting implicit references to techniques would reinforce the under-reporting trends we already observe.

The *Techniques** row in Table 5 captures the techniques profiles extended with the technique identifiers extracted

TABLE 6: Summary of CVE and technique identifiers extracted from the downloaded threat reports. For each dataset, it presents: (i) the total number of reports analyzed, (ii) the number of reports containing at least one CVE identifier, (iii) the number of unique CVEs extracted, and (iv) the number of threat groups associated with those CVEs. It then provides similar statistics for the extraction of technique identifiers.

Dataset	Reports	Reports w/CVEs	CVEs	Groups	Reports w/Tech	Tech.	Groups
ATT&CK	920	266 (29.0%)	325	86	122 (13.2%)	470	63
Malpedia	4,414	943 (21.3%)	853	261	541 (12.2%)	626	204
All	5,827	1,186 (20.3%)	906	277	650 (11.1%)	658	211

TABLE 7: Top generic CVEs, i.e., used by most groups, and whether a PoC exploit is publicly available.

Vulnerability	Affected Software	PoC	Groups
CVE-2017-11882	Microsoft Office	✓	46 (5.7%)
CVE-2012-0158	Microsoft Office	✓	41 (5.1%)
CVE-2017-0199	Microsoft Office	✓	34 (4.2%)
CVE-2021-44228	Apache Log4j	✓	28 (3.5%)
CVE-2022-30190	Microsoft Windows (MSDT)	✓	25 (3.1%)
CVE-2022-26134	Atlassian Confluence	✓	21 (2.6%)
CVE-2018-0802	Microsoft Office	✓	20 (2.5%)
CVE-2022-38028	Windows Print Spooler	✗	17 (2.1%)
CVE-2023-38831	RARLAB WinRAR	✓	17 (2.1%)
CVE-2024-37085	VMware ESXi	✓	16 (2.0%)

TABLE 8: Examples of group-specific vulnerabilities, some groups have multiple group-specific vulnerabilities.

Group Name	Vulnerability	Affected Software
APT37	CVE-2015-3636	Linux Kernel
	CVE-2016-0147	MSXML
Akira	CVE-2019-6693	Fortinet FortiOS
	CVE-2023-29336	Microsoft Windows
	CVE-2023-35078	Ivanti Endpoint Manager
Kimsuky	CVE-2012-4873	GNU Board
	CVE-2018-14745	Samsung Galaxy
	CVE-2018-2628	Oracle WebLogic Server
Gorgon Group	CVE-2015-7036	Apple iOS
	CVE-2019-8457	SQLite3
	CVE-2019-8598	iOS, macOS
Sidewinder	CVE-2018-4876	Adobe Experience Manager
	CVE-2018-7445	MicroTik RouterOS
	CVE-2019-2215	Google Android
Scattered Spider	CVE-2015-2291	Ethernet driver on Windows
	CVE-2021-35464	ForgeRock Access Management
	CVE-2022-0001	Intel Processors
Carbanak	CVE-2013-2463	Oracle JRE
	CVE-2015-2426	Microsoft Windows
	CVE-2016-1010	Adobe Flash Player

from the threat reports. It shows that we can identify 69 groups with group-specific techniques from the Malpedia threat reports, compensating for the lack of techniques in Malpedia. Notably, when we extract technique identifiers directly from the ATT&CK reports and combine them with the existing ATT&CK technique profiles, the number of groups with group-specific techniques increases from 52 (34.2%) to 60 (39.5%), despite analyzing the same set of threat reports. This highlights that the manual extraction of techniques by ATT&CK contributors is not always optimal.

Takeaway: The extraction of technique identifiers from threat reports increases the number of groups with group-specific techniques from 52 (6.4%) to 119 (14.7%) mostly due to additional techniques in the Malpedia reports. The comparison of technique identifiers extracted from ATT&CK reports with the techniques indexed in ATT&CK shows that the extraction process used by ATT&CK contributors may miss techniques, suggesting the need for an automated approach.

Implicit reference extraction. Previous work has proposed NLP techniques for recovering the ATT&CK technique identifiers implicitly mentioned in threat reports [2], [18], [37]. An alternative approach would be to use Large Language Models (LLMs), which have proved their flexibility in a number of security-related tasks involving natural language texts [10], [43]. We performed some preliminary experiments using large language models (LLMs) to extract techniques from the downloaded threat reports. Specifically, we used the commercial GPT-4 model, guided by a prompt shown in Figure 4 in the Appendix, which we experimentally identified as the most effective among other options. To ensure the model analyzed the report content, we removed any tables of techniques included at the end of the report.

Unfortunately, we obtained mixed results as the LLM frequently hallucinated techniques introducing false positives (FPs). An example is a 2022 report on the Lyceum group [23]. For this report, iocsearcher identified 8 techniques (without sub-techniques), all of them in a table at the end of the report. These are the same 8 techniques that ATT&CK associates to Lyceum from this report, suggesting that the contributor relied directly on the table for extraction. When we provided the same report after removing the table of identifiers, the LLM returned 11 technique identifiers (7 techniques and 4 sub-techniques). Of these, only two overlapped with the original table. We manually reviewed the remaining 9, finding that while two were valid, seven were false positives (FP). An example of a correctly extracted implicit reference is T1547.001 *Boot or Logon Autostart Execution: Startup Folder* (part of the *Persistence* tactic) that was extracted from the following text: “written into the Startup folder in order to maintain persistence”.

One example of a false positive is T1018, *Remote System Discovery*. When prompted to justify its inclusion, the model responded, “While the report does not specifically mention remote system discovery, many backdoors and malware en-

gage in network reconnaissance, which aligns with T1018”. In future work, we would like to compare LLM-based extraction with existing extraction tools [2], [18], [37].

5. Discussion

5.1. Implications of Results

Our work critically challenges the widespread notion that behavioral profiles for threat actor attribution can replace IOCs. We show that behavioral profiles are not as distinctive as expected with many groups employing generic techniques, software and vulnerabilities also used by other groups. Only a small fraction of threat groups have group-specific behaviors that uniquely identify them and thus could be used as behavioral signatures. Roughly two thirds (65.8%) of groups in ATT&CK have no group-specific techniques, challenging the distinctiveness of TTPs. The software used by a group is more distinctive with 73% groups in ATT&CK using group-specific programs. However, once we consider the larger number of groups in Malpedia the percentage drops to 26.3%. Despite leveraging information from ATT&CK and Malpedia and extending the profiles with the exploited vulnerabilities and additional techniques, the fraction of threat groups without unique behaviors remains at 64%.

As the number of profiled threat groups increases from 152 in ATT&CK to 800 in Malpedia, the fraction of groups with distinctive behaviors declines, as behaviors that initially looked unique may, with broader coverage, be observed across multiple groups. Consequently, even for the roughly one-third of groups exhibiting group-specific behaviors, a key question remains: are these behaviors genuinely unique, or do they appear so due to incomplete visibility into other groups? In fact, we observe cases where group-specific techniques are not inherently unique to a group (e.g., network DoS), but rather unassociated with other groups due to under-reporting. This raises concerns about the confidence of behavior-based attribution, which may impact critical attribution tasks such as those performed by legal or law enforcement, e.g., the attribution may not stand examination in a judicial process.

We observe that wide coverage is critical for constructing truly distinctive behavioral profiles. A key factor impacting coverage is the number of threat reports analyzed. Building profiles from small datasets of threat reports (e.g., those written by a single vendor or a select group of trusted sources) is tempting because those reports may be more uniform and less “noisy”, and it is easier to find behaviors that initially look unique. However, such behavioral profiles provide little confidence. Given the limited overlap we observe between ATT&CK and Malpedia, relying on a few sources may limit coverage too much. We argue that it is better to index more threat reports (as Malpedia does) rather than focusing narrowly on very selected sources (as ATT&CK seems to do) because it is hard and error-prone to predict which sources are the most accurate. Furthermore,

such filtering for trusted sources can always be performed a posteriori on the indexed data, as long as the mapping from behaviors to original sources is maintained (as the examined datasets do). Reports from known low-confidence sources can be excluded later in the pipeline. In contrast, assuming only a small set of sources is trustworthy may lead to overly constrained coverage. This is particularly problematic given that each vendor produces a limited number of reports annually and that reporting tends to concentrate on high-profile threat groups. Another important factor influencing coverage is the methodology used to extract behaviors from threat reports (e.g., manual or automated). We found that we could extract more explicit technique references from ATT&CK reports than those indexed by ATT&CK itself, even when analyzing the same reports. This suggests that the manual extraction process used by ATT&CK contributors may not always be exhaustive, further motivating the adoption of automated approaches [2], [18], [37].

Beyond techniques and software already indexed in the examined datasets, we have shown that behavioral profiles can be further extended with the exploited vulnerabilities. Additionally, other behavioral features present in the threat reports such as, payment services for adversaries to receive ransom, communication channels through which victims and adversaries interact, and the textual content that adversaries present to victims (e.g., emails, ransom notes).

Our analysis reveals a number of potential improvements to the datasets. First, we observe that the addition of OS-integrated tools (e.g., net, ping) into the ATT&CK software taxonomy provides little value, as those tools are already available in most target systems. Having a complete software taxonomy is not realistic, the focus should be on tools that adversaries deploy, which captures intent and provides more behavioral information. Second, the classification of software into tools and malware in ATT&CK is useful for analysts to quickly filter generic software but is missing in Malpedia. Furthermore, it is not clear where malicious kits should be placed, possibly indicating the need for a third category. We also observe some likely misclassified software, e.g., Cobalt Strike is arguably a tool rather than a malware. Finally, the split of techniques into domains used by ATT&CK seems quite arbitrary as techniques may apply to different domains, albeit with different implementations. For example, there is a Rootkit (T1014) technique in the Enterprise domain and another Rootkit (T0851) technique in the ICS domain. The latter includes in the description references to firmware rootkits and Stuxnet but having two equally named techniques is confusing and likely unnecessary. There is also overlap between techniques. For example, the Enterprise domain includes Pre-OS Boot: System Firmware (T1542.001) for capturing adversaries modifying system firmware to persist on systems, which seems the same as a firmware rootkit. Given an observation of a rootkit in a device, different security vendors and analysts may assign any combination of the above 3 techniques to the observation, which would complicate understanding which group may be behind the attack. This split complicates usage as three different technique taxonomies, one per domain,

need to be considered. This makes it tempting to focus on the Enterprise domain comprising 76% of all techniques and sub-techniques. A more unified taxonomy, with domains encoded as an attribute, could streamline analysis.

5.2. Threats to Validity

We discuss some potential threats to the validity of our results. First, our reliance on open-source threat intelligence (OSINT) introduces potential selection bias. Commercial CTI feeds could offer more detailed analysis of some threat groups. However, such commercial feeds are often limited to data from a single provider, significantly restricting their overall coverage. Second, a small number of threat reports could not be successfully downloaded, potentially leading to an underestimation of dataset coverage. Nonetheless, fewer than 5% of the URLs resulted in an error. To mitigate this, future work could incorporate archival sources such as the Wayback Machine [19] and AptNotes [3]. Third, our focus on group-specific behaviors as behavioral signatures may overlook groups characterized by unique combinations of non-exclusive behaviors. Unfortunately, the more behaviors needed to identify a group, the greater the risk that an attack goes unattributed. Finally, inconsistencies in naming conventions across knowledge bases pose a challenge for accurate data comparison. Although we applied a normalization strategy to align group and software names, there remains a risk that some mappings are incorrect.

6. Related Work

Our research relates to the following prior CTI research.

Knowledge bases. Previous work has presented the design of the two knowledge bases we use [34], [45]. Other works have analyzed the usage of ATT&CK by systematically reviewing literature on its applications [1], [20], [39]. Oftentimes, works use the knowledge bases simply as a source of threat reports from where IOCs can be extracted [8], [22]. Our work differs in measuring the utility of ATT&CK and Malpedia for the specific case of adversary profiling.

Application-oriented studies. Several studies have examined the use of ATT&CK across different cybersecurity contexts. Oosthoek et al. [33] employed ATT&CK to map sandbox evasion techniques across 951 Windows malware families, offering insight into both commonly used and increasingly adopted techniques in recent years. Virkud et al. [49] evaluate the ATT&CK framework in commercial endpoint detection products and assess its effectiveness as a security evaluation metric. They find that while these products typically cover between 48%–55% of ATT&CK techniques, much of this coverage consists of low-risk or less impactful rules. Their findings suggest that although ATT&CK is increasingly used to assess threat readiness, reported coverage frequently fails to reflect actual detection capabilities in real-world scenarios. In another line of work, Rahman et al. [36] investigate challenges in implementing security controls (e.g., strong password policies)

against ATT&CK techniques. In simultaneous and independent work, yet to be presented, Horst et al. [47] examine the role of low-level IOCs (e.g., domains) and high-level IOCs (e.g., TTPs) in ransomware attribution. They use a mixed-methods approach, combining interviews of 15 ransomware attribution experts and analyzing 27 incident reports from two sources. They show that experts leverage low-level IOCs for attribution more frequently than high-level IOCs, which they regard as too generic. Our results match theirs in raising concerns about using behavioral traits for attribution. But, our approaches are quite different. They examine 16 ransomware groups while we examine 807 threat groups covering different types of adversaries (e.g., APTs). We do not perform interviews but analyze over 15K threat reports from two popular knowledge bases. And, we measure for the first time the fraction of threat groups with unique behaviors.

Automated CTI extraction. Husari et al. [18] made early efforts to automate the extraction of TTPs from threat intelligence reports, using a context-aware, rule-based approach to identify and extract threat actions from both structured and unstructured CTI sources. The extracted TTPs are standardized using the STIX [31] format, with the tool achieving over 82% precision and recall on a proprietary dataset. Extending this work, Alam et al. [2] employed machine learning for automated extraction of attack patterns and IOCs. Their framework further mapped the extracted behaviors to the standardized ATT&CK framework and organized them in a knowledge graph to facilitate predictive analysis. Complementing these extraction-focused efforts, Rahman et al. [35] analyzed 667 CTI reports from the ATT&CK framework to study the prevalence and co-occurrence of TTPs used in APT campaigns, providing insights into adversary patterns. Our work builds upon these approaches by combining threat intelligence data from both the ATT&CK framework and Malpedia. We examine techniques and vulnerability usage across adversary groups, offering insights into building more comprehensive threat group profiles.

7. Conclusion

Our study critically evaluates the assumption that behavioral profiles can effectively replace Indicators of Compromise (IOCs) for threat actor attribution. By analyzing two open-source CTI knowledge bases, MITRE ATT&CK and Malpedia, we show significant limitations in the distinctiveness and completeness of group behavioral profiles. Specifically, only 34.2% of ATT&CK groups have group-specific techniques. Even after incorporating software and vulnerabilities from both ATT&CK and Malpedia, 64% of threat groups still lack unique behavioral signatures. As coverage expands from 152 groups in ATT&CK to 800 in Malpedia, the specificity of behaviors diminishes, with previously unique features proving to be more widespread. These findings highlight that group-specific behaviors are both rare and often overestimated.

Acknowledgements

This work was partially funded by the Spanish Government MCIN/AEI/10.13039/501100011033/ through grants TED2021-132464B-I00 (PRODIGY) and PID2022-142290OB-I00 (ESPADA). The above grants are co-funded by European Union ESF, EIE, and NextGeneration funds.

This work was further supported by the Vienna Science and Technology Fund (WWTF) and the City of Vienna [10.47379/ICT19056], the SecInt Doctoral College at TU Wien, and SBA Research (SBA-K1 NGC), a COMET Center within the COMET – Competence Centers for Excellent Technologies Programme and funded by BMIMI, BMWET, and the federal state of Vienna. The COMET Programme is managed by FFG.

References

- [1] Bader Al-Sada, Alireza Sadighian, and Gabriele Oligeri. MITRE ATT&CK: State of the Art and Way Forward. *ACM Computing Surveys*, 57(1), 2024.
- [2] Md Tanvirul Alam, Dipkamal Bhusal, Youngja Park, and Nidhi Rastogi. Looking Beyond IoCs: Automatically Extracting Attack Patterns from External CTI. In *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses (RAID)*, 2023.
- [3] APTnotes repository. <https://github.com/aptnotes/data/>, 2025.
- [4] Mohammed Asiri, Neetesh Saxena, Rigel Gjomemo, and Pete Burnap. Understanding Indicators of Compromise against cyber-attacks in Industrial Control Systems: A Security Perspective. *ACM Transactions on Cyber-Physical Systems*, 7(2), 2023.
- [5] David Bianco. The Pyramid of Pain. <https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html>, 2013.
- [6] Xander Bouwman, Harm Griffioen, Jelle Egbers, Christian Doerr, Bram Klievink, and Michel Van Eeten. A Different Cup of TI? The Added Value of Commercial Threat Intelligence. In *Proceedings of the 29th USENIX Security Symposium (USENIX Sec)*, 2020.
- [7] Xander Bouwman, Victor Le Pochat, Pawel Foremski, Tom Van Goethem, Carlos H. Ganan, Giovane C. M. Moura, Samaneh Tajalizadehkhoob, Wouter Joosen, and Michel van Eeten. Helping Hands: Measuring the Impact of a Large Threat Intelligence Sharing Community. In *Proceedings of the 31st USENIX Security Symposium (USENIX Sec)*, 2022.
- [8] Juan Caballero, Gibran Gomez, Srdjan Matic, Gustavo Sánchez, Silvia Sebastián, and Arturo Villacañas. The Rise of GoodFATR: A Novel Accuracy Comparison Methodology for Indicator Extraction Tools. *Future Generation Computer Systems*, 144, 2023.
- [9] Canadian Center for Cyber Security. Joint Report on Publicly Available Hacking Tools. https://www.cyber.gc.ca/sites/default/files/cyber/publications/joint_report_on_publicly_available_hacking_tools.pdf, 2018.
- [10] Hoang Cuong Nguyen, Shahroz Tariq, Mohan Baruwal Chhetri, and Bao Quoc Vo. Towards Effective Identification of Attack Techniques in Cyber Threat Intelligence Reports using Large Language Models. In *Companion Proceedings of the ACM on Web Conference (WWW)*, 2025.
- [11] Empire Post-exploitation Framework. <https://github.com/EmpirePrject/Empire>, 2025.
- [12] Exploit Database. <https://www.exploit-db.com/>, 2025.
- [13] Facundo Muñoz. njRAT: A Remote Access Trojan Widely Used by Various Cybercriminals. <https://www.welivesecurity.com/la-es/2021/09/29/que-es-njrat-troyano-acceso-remoto-utilizado-cibercriminales/>, 2021.
- [14] FORTRA. Cobalt Strike - Adversary Simulations and Red Team Operations. <https://www.cobaltstrike.com/>, 2025.
- [15] Malpedia. <https://malpedia.caad.fkie.fraunhofer.de/>, 2025.
- [16] gh0st RAT. <https://github.com/sin5678/gh0st>, 2025.
- [17] Dan Goodin. Windows Vulnerability Reported by the NSA Exploited to Install Russian Malware. <https://arstechnica.com/security/2024/04/kremlin-backed-hackers-exploit-critical-windows-vulnerability-reported-by-the-nsa/>, 2024.
- [18] Ghaith Husari, Ehab Al-Shaer, Mohiuddin Ahmed, Bill Chu, and Xi Niu. TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources. In *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC)*, 2017.
- [19] Internet Archive. Wayback Machine. <https://web.archive.org/>, 2025.
- [20] Yuning Jiang, Qiaoran Meng, Feiyang Shang, Nay Oo, Le Thi Hong Minh, Hoon Wei Lim, and Biplab Sikdar. MITRE ATT&CK Applications in Cybersecurity and The Way Forward. *arXiv preprint arXiv:2502.10825*, 2025.
- [21] Beomjin Jin, Eunsoo Kim, Hyunwoo Lee, Elisa Bertino, Doowon Kim, and Hyoungshick Kim. Sharing Cyber Threat Intelligence: Does it Really Help? In *Proceedings of the 31st Network and Distributed System Security Symposium (NDSS)*, 2024.
- [22] Robert J Joyce, Dev Amlani, Charles Nicholas, and Edward Raff. Motif: A Malware Reference Dataset with Ground Truth Family Labels. *Computers & Security*, 124, 2023.
- [23] lyceum. Lyceum .NET DNS Backdoor. <https://www.zscaler.com/blogs/security-research/lyceum-net-dns-backdoor>, 2022.
- [24] Silent Chollima. https://malpedia.caad.fkie.fraunhofer.de/actor/silent_chollima, 2025.
- [25] MalwareLab. On the Royal Road. https://blog.malwarelab.pl/posts/on_the_royal_road/, 2020.
- [26] MISP Threat Actors Galaxy. <https://github.com/MISP/misp-galaxy/blob/main/clusters/threat-actor.json>, 2025.
- [27] MITRE ATT&CK. <https://attack.mitre.org/>, 2025.
- [28] Babuk. <https://attack.mitre.org/software/S0638/>, 2025.
- [29] Conficker. <https://attack.mitre.org/software/S0608/>, 2025.
- [30] SimBad. <https://attack.mitre.org/software/S0419/>, 2025.
- [31] MITRE ATT&CK STIX Data. <https://github.com/mitre-attack/attack-stix-data>, 2024.
- [32] Aleksandr Nahapetyan, Sathvik Prasad, Kevin Childs, Adam Oest, Yeganeh Ladwig, Alexandros Kapravelos, and Brad Reaves. On SMS Phishing Tactics and Infrastructure. In *Proceedings of IEEE Security & Privacy Symposium (IEEE S&P)*, 2024.
- [33] Kris Oosthoek and Christian Doerr. SoK: ATT&CK Techniques and Trends in Windows Malware. In *International Conference on Security and Privacy in Communication Systems (SecureComm)*, 2019.
- [34] Daniel Plohmann, Martin Clauss, Steffen Enders, and Elmar Padilla. Malpedia: A Collaborative Effort to Inventorize the Malware Landscape. *The Journal on Cybercrime & Digital Investigations*, 3(1), 2018.
- [35] Md Rayhanur Rahman, Setu Kumar Basak, Rezvan Mahdavi Hezaveh, and Laurie Williams. Attackers Reveal their Arsenal: An Investigation of Adversarial Techniques in CTI Reports. *arXiv preprint arXiv:2401.01865*, 2024.
- [36] Md Rayhanur Rahman, Brandon Wroblewski, Mahzabin Tamanna, Imranur Rahman, Andrew Anufriyenko, and Laurie Williams. Towards a Taxonomy of Challenges in Security Control Implementation. In *Proceedings of the 40th Annual Computer Security Applications Conference (ACSAC)*, 2024.

- [37] Nanda Rani, Bikash Saha, Vikas Maurya, and Sandeep Kumar Shukla. TTPXHunter: Actionable Threat Intelligence Extraction as TTPs from Finished Cyber Threat Reports. *Digital Threats: Research and Practice*, 5(4), 2024.
- [38] Thomas Rid and Ben Buchanan. Attributing Cyber Attacks. *Journal of Strategic Studies*, 38(1-2), 2015.
- [39] Shanto Roy, Emmanouil Panaousis, Cameron Noakes, Aron Laszka, Sakshyam Panda, and George Loukas. SoK: The MITRE ATT&CK Framework in Research and Practice. *arXiv preprint arXiv:2304.07411*, 2023.
- [40] Aakanksha Saha, Jorge Blasco, Lorenzo Cavallaro, and Martina Lindorfer. ADAPT it! Automating APT Campaign and Group Attribution by Leveraging and Linking Heterogeneous Files. In *Proceedings of the 27th International Symposium on Research in Attacks, Intrusions and Defenses (RAID)*, 2024.
- [41] Aakanksha Saha, James Mattei, Jorge Blasco, Lorenzo Cavallaro, Daniel Votipka, and Martina Lindorfer. Expert Insights into Advanced Persistent Threats: Analysis, Attribution, and Challenges. In *Proceedings of the 34th USENIX Security Symposium (USENIX Sec)*, 2025.
- [42] Sentinel Labs. ShadowPad: A Masterpiece of Privately Sold Malware RAT. <https://assets.sentinelone.com/c/shadowpad?x=p42eqa>, 2021.
- [43] Sayuj Shah and Vijay K Madiseti. MAD-CTI: Cyber Threat Intelligence Analysis of the Dark Web Using a Multi-Agent Framework. *IEEE Access*, 13, 2025.
- [44] Florian Skopik and Timea Pahi. Under False Flag: Using Technical Artifacts for Cyber Attack Attribution. *Cybersecurity*, 3, 2020.
- [45] Lake E. Strom, Andy Applebaum, Doug P. Miller, Kathryn C. Nickels, Adam G. Pennington, and Cody B Thomas. MITRE ATT&CK: Design and Philosophy. https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf, 2018.
- [46] ThreatMiner: Data Mining for Threat Intelligence. <https://www.threatminer.org/>, 2025.
- [47] Max van der Horst, Ricky Kho, Olga Gadyatskaya, and Michel Mollema. High Stakes, Low Certainty: Evaluating the Efficacy of High-Level Indicators of Compromise in Ransomware Attribution. In *Proceedings of the 34th USENIX Security Symposium (USENIX Sec)*, 2025.
- [48] Kevin van Liebergen, Gibran Gomez, Srdjan Matic, and Juan Caballero. All your (data)base are belong to us: Characterizing Database Ransom(ware) Attacks. In *Proceedings of Network and Distributed Systems Security Symposium (NDSS)*, 2025.
- [49] Apurva Virkud, Muhammad Adil Inam, Andy Riddle, Jason Liu, Gang Wang, and Adam Bates. How does Endpoint Detection use the MITRE ATT&CK Framework? In *Proceedings of the 33rd USENIX Security Symposium (USENIX Sec)*, 2024.
- [50] Yiming Wu, Qianjun Liu, Xiaojing Liao, Shouling Ji, Peng Wang, Xiaofeng Wang, Chunming Wu, and Zhao Li. Price Tag: Towards Semi-Automatically Discovery Tactics, Techniques and Procedures of e-commerce Cyber Threat Intelligence. *IEEE Transactions on Dependable and Secure Computing*, 2021.

Appendix

Figure 4 shows an example of a prompt engineered to instruct LLM (GPT-4) to identify Tactics, Techniques, and Procedures (TTPs) from a natural language threat report and map them to their respective MITRE ATT&CK Technique IDs, formatted for JSON output.

I have a detailed threat report written in natural language. I need help identifying the TTPs (Tactics, Techniques, and Procedures) described in the report and mapping them to their corresponding MITRE ATT&CK Technique IDs. The output should include:

- A list of tactics (high-level strategic goals) based on the threat actor's behavior.
- A list of techniques (specific actions or behaviors), with their corresponding MITRE ATT&CK Technique IDs.
- A description of procedures (the exact implementation or variation of a technique as described in the report).

For each technique, provide both the name and the MITRE ATT&CK Technique ID. Please ensure all identified TTPs are clearly mapped to the most relevant MITRE ATT&CK entries.

Here's the threat report:
[Insert threat report text here]

The output should be in this JSON format:
Example Output:

```
{
  "tactics": ["Initial Access", "C2"],
  "techniques": [
    {"name": "Spear Phishing", "MITRE ID": "T1193"},
    {"name": "C2 Channel Over HTTPS", "MITRE ID": "T1071"}
  ],
  "procedures": ["Use of malicious scripts"]
}
```

Figure 4: LLM prompt.