Extracting Cyber Threat Intelligence from Port Scans: A Taxonomy-Based Approach

Jan Geisler, Robert Koch, Alexander Nußbaum and Gabi Dreo Rodosek
Universität der Bundeswehr, Neubiberg, Germany

Rober.Koch@unibw.de
Alexander.Nussbaum@unibw.de
Gabi.Dreo@unibw.de

Abstract: Port scans are a common preliminary step for a variety of cyberattacks, from simple hackers, attempted automated exploitation, to professional groups and state actors. They serve as a reconnaissance technique that facilitates the planning and execution of future attacks and are often conducted stealthily over extended periods to evade monitoring systems, making them challenging to identify and analyse. Despite this, effective detection and analysis of port scans can yield valuable cyber threat intelligence (CTI), enabling defenders to prioritize defensive measures, deploy and optimize protective infrastructure such as Intrusion Detection and Prevention Systems (IDS/IPS), and anticipate potential attacks by analysing the characteristics and frequency of scans. However, the huge amount of data generated by port scans and other network events hides the significant operations and complicates the extraction of actionable intelligence. We present a comprehensive taxonomy designed to classify and analyse port scans systematically. We focus on interpreting detected port scans rather than their detection, leveraging the wide availability of detection tools. Our taxonomy assesses key attributes of port scans, including the intent, origin, potential hostile gain, damage potential, available intelligence, and the necessity for responsive actions. We then propose an 8-step classification process to guide this analysis. It begins with a thorough technical analysis of the scan which can be provided by various detection frameworks. Based on that, the legitimacy of a detected scan is determined, distinguishing between malicious intent and benign activities like friendly analysis, general research, or internet background noise. Next, we generate a "fingerprint" of the scan and cross-reference it against a database of known scans, compiled from historical data, CTI repositories, and incident reports. The analysis further evaluates the scan's target, the information it may have revealed, and its success level. We also explore the broader intelligence that can be gleaned from the scan, enhancing situational awareness of our systems. Finally, we assess the technical response options, considering their feasibility and cost-effectiveness, and determine whether proactive measures are warranted. We show that our structured approach to port scan analysis improves the generation of actionable intelligence and supports informed decision-making for defensive strategies.

Keywords: Port scans, Cyber threat intelligence, Security operations, Intrusion detection, Network security

1. Introduction

Internet-facing gateways of any organization detect millions of firewall events each day, a majority of which are innocuous. While conventional detection methods can effectively thwart straightforward attacks through signature or anomaly-based techniques, it is significantly harder to do so for preparatory measures such as port scans (Lagraa, and François, 2017). Network gateways possess a certain degree of resilience against attacks and can provide significant contributions to threat intelligence through strategic sensor deployment in an ideal state. To prioritize the deployment of protective measures, reliable threat intelligence and threat analysis are essential (Conti, Dargahi, and Dehghantanha, 2018). These high-level recommendations are based log data from firewalls as well as Intrusion Detection Systems (IDS) or Intrusion Prevention Systems (IPS) which are generated at almost every gateway and are usually digested by a Security Information and Event Management (SIEM) system (Conti, Dargahi, and Dehghantanha, 2018). Based on this data and equipped with a database of Indicators of Compromise (IoCs), attacks can be detected somewhat reliably (Conti, Dargahi, and Dehghantanha, 2018). However, identifying attack preparations, such as port scans, proves to be significantly more challenging due to the constant and high background noise in the form of network scans for research purposes by universities, companies, and security-related authorities (Stallings, 2016).

2. Fundamentals of Port Scanning

Port scanning is a method used to identify open ports and services on a networked system. It is a valuable tool for network administrators to assess the security posture of their systems, but it is also frequently used by attackers to discover vulnerabilities. The process involves sending probes to each port on a target machine and analysing the responses to determine which ports are open and what services are running. This information can reveal more critical details about the system, such as which services are available, which users own these services, and whether anonymous logins are supported or certain network services require authentication (De Vivo et al, 1999).

2.1 Common Port Scanning Techniques

Various techniques are employed in port scanning to gather information about the target system. Each technique has its own special characteristics, advantages, and drawbacks, and knowing these can help both to detect scans and to perform them effectively. Roger (2001) as well as Bhuyan, Bhattacharyya, and Kalita (2011) consider the following techniques among the most important:

- Transmission Control Protocol (TCP) Connect Scan: This basic scanning technique uses the operating system's (OS) connect system call to open a connection to every port. It is very noisy and easily detectable, as it generates complete connections that are logged by the target system.
- TCP SYN (Half-Open) Scan: Often referred to as a half-open scan, this method sends a SYN packet to the target port and waits for a SYN-ACK response. If the port is open, the target responds with a SYN-ACK, and the scanner sends an RST to avoid completing the connection. This type of scan is harder to detect since it does not establish a full connection.
- TCP FIN Scan: In a FIN scan, the attacker sends FIN packets to the target ports. Closed ports respond with RST packets, while open ports ignore the FIN packets. This method can bypass certain firewalls and packet filters.
- TCP XMAS Scan: This scan manipulates the TCP header by setting the FIN, PSH, and URG flags. Closed ports respond with RST packets, while open ports do not respond. It is named XMAS scan because the set flags are analogous to lit-up Christmas tree lights.
- TCP NULL Scan: A NULL scan sends packets with no flags set. Closed ports respond with RST packets, while open ports do not respond. This method can evade some firewall and IDS/IPS.
- TCP ACK Scan: Used to determine whether a port is filtered, an ACK scan sends TCP packets with the ACK flag set. If the port is filtered, there is no response; if it is unfiltered, an RST packet is sent back.
- User Datagram Protocol (UDP) Scan: Unlike TCP, UDP is connectionless. A UDP scan sends a UDP packet to each port. If a port is closed, an Internet Control Message Protocol (ICMP) port unreachable message is returned. If the port is open, there may be no response or a UDP packet is returned. This scan is slower and less reliable due to the nature of the UDP protocol.

2.2 Detection of Port Scans

The detection of port scans is essential for identifying potential security threats. Several methods can be employed to detect port scanning activities:

- Signature-Based Detection: This method relies on predefined signatures of known port scanning tools and techniques. IDS and firewalls can use these signatures to recognize and block scanning attempts (Patel and Sonker, 2016). However, signature-based detection can be circumvented by novel or modified scanning methods (Patel and Sonker, 2016).
- Anomaly-Based Detection: This approach involves establishing a baseline of normal network activity and identifying deviations from this baseline (Ananin, Nikishova, and Kozhevnikova, 2017). Unusual patterns, such as repeated connection attempts to multiple ports, can indicate a port scan. This method is effective against new and unknown scanning techniques but may generate false positives (Ananin, Nikishova, and Kozhevnikova, 2017).
- Rate-Based Detection: This technique monitors the rate of connection attempts from a single source. A high rate of connection attempts in a brief period can indicate a port scan (Ring, Landes, and Hotho, 2018). This method is simple to implement but can be bypassed by slow scans that spread their attempts over a longer period (Ring, Landes, and Hotho, 2018).
- Behavioural Analysis: This method examines the characteristics of network traffic to identify patterns indicative of port scanning. For example, a single source Internet Protocol (IP) address attempting to connect to multiple destination ports or addresses in a short timeframe can be flagged as suspicious (Schäfer and Drozd, 2011). Behavioural analysis often involves advanced analytics and Machine Learning techniques (Schäfer and Drozd, 2011).
- Honeypots: Honeypots are decoy systems set up to attract and detect malicious activities. By monitoring interactions with these decoy systems, security teams can identify port scanning attempts and gather valuable information about the attacker's methods and objectives (Stallings, 2016).

Port scanning is a common technique used by both network administrators and attackers to identify open ports and services. Knowledge of the distinct types of port scans and employing effective detection methods are essential for maintaining network security and preventing potential attacks.

3. Taxonomy

Developing a comprehensive taxonomy for port scan detection and classification begins with a clear understanding of the types of activities involved. The goal is to create a systematic approach that can accurately identify and categorize various port scanning activities. To guide this process, we identified six critical questions, each addressing a specific aspect of port scanning behaviour and its implications for network security.

3.1 Six Critical Questions

The first question to consider is: *Is an activity malicious?* This distinction is crucial for identifying benign versus malicious port scans. Certain characteristics of a scan can help in this determination: Benign scans may come from known IP addresses associated with reputable organizations, occur at low and regular frequencies (Bhuyan, Bhattacharyya, and Kalita, 2011), target specific, consented networks (Hindy et al, 2021), use standardized methods (AlAhmadi and Martinovic, 2018), and/or involve clear communication of intent. In contrast, malicious scans are expected to originate from unknown or blacklisted IPs, exhibit unusual and irregular frequencies (Bhuyan, Bhattacharyya, and Kalita, 2011), target multiple hosts indiscriminately (Bhuyan, Bhattacharyya, and Kalita, 2011), employ advanced evasion techniques (AlAhmadi and Martinovic, 2018), and/or lack transparency. The intent behind a port scan is crucial information for all further response activities and an important filter for managing workload of responders.

The second question is: *Does an activity match the behaviour of known threat actors?* Matching the observed activity with known attack patterns can help attribute the scan to specific threat actors. This is achieved by comparing the activity with threat intelligence databases that document the tactics, techniques, and procedures (TTPs) of various attackers. Identifying the threat actor behind a scan provides context and helps anticipate future actions, making this step crucial for proactive defence (Conti, Dargahi, and Dehghantanha, 2018).

The third question is: *How much information could an attacker gain through this activity?* This involves assessing the potential information exposure resulting from the scan. Exposed information does not have any detectable impact but may make future attacks easier, similar to how a stolen password enables account theft. Information such as open ports, running services, and system vulnerabilities can be highly valuable to an attacker (Stallings, 2016). By evaluating what an attacker could learn from the scan, one can gauge the severity of future threats and prioritize responses accordingly (Kurose and Ross, 2017).

The fourth question is: *Are more activities necessary before damage is done?* This question helps us understand whether the detected scan is a standalone threat or part of a larger attack chain. Some port scans are simple wide-area reconnaissance activities without ulterior motives, while others may be the initial step in a multi-stage attack (Kurose and Ross, 2017). Identifying the role of the scan within the broader context of an attack strategy is essential for effective mitigation.

The fifth question is: *Can we generate threat intelligence from the observed behaviour?* Generating actionable threat intelligence from detected port scans enhances the overall security posture by informing proactive defences. This involves documenting Indicators of Compromise (IoCs) and sharing them with relevant stakeholders (Conti, Dargahi, and Dehghantanha, 2018). Effective threat intelligence generation helps create a collaborative defence environment.

The last question to consider is: *Should this activity be prevented?* This involves weighing the effort required to prevent the activity against its potential impact. It is essential to balance the need for security with the practicalities of network operations. Some scanning activities may have legitimate uses, and some activities identified as scans may be false positives of benign operations. In such cases, indiscriminate prevention could disrupt normal activities (Kurose and Ross, 2017). Therefore, careful consideration is necessary to ensure that preventive measures are effective and proportionate.

3.2 Eight-Step Classification Process

To answer these six questions in systemically and generalize the process into a taxonomy, we introduce an eight-step classification process, each step building upon the previous one both sequentially and logically to provide a comprehensive analysis of detected port scans. By systematically addressing these questions and following the eight-step process, the proposed taxonomy provides a robust framework for analysing and categorizing port scanning activities, ultimately enhancing the effectiveness of network security measures.

3.2.1 Step 1: Activity identification

The first step in our taxonomy development process is the identification of the type of activity. This step is crucial as it provides the foundational classification upon which further analysis will be based. Several types of port scanning activities can be identified, each with distinct characteristics and implications for network security.

Vertical port scans involve scanning a single host for multiple open ports. This type of scan is often used by attackers to identify vulnerable services on a target machine. Vulnerability scans are designed to identify known vulnerabilities within the systems being scanned. These scans can be legitimate, performed by network administrators to assess security, or malicious when conducted by attackers (Stallings, 2016). Sweep scans target the same port across multiple hosts, helping attackers to identify which machines on a network have a particular service running (Bhuyan, Bhattacharyya, and Kalita, 2011). ICMP scans use ICMP to discover hosts on a network. Slow scans, on the other hand, spread the scanning activity over a prolonged period to avoid detection by security systems (Ring, Landes, and Hotho, 2018). Each type of activity can be identified and classified based on specific criteria, such as the pattern of network traffic, the timing of scan attempts, and the range of ports or IP addresses targeted. Understanding these criteria is essential for accurate classification and subsequent analysis.

Port scanning activities can involve different protocols, each of which plays a distinct role in the scanning process. The most used protocols in port scans are TCP, UDP, and ICMP (Roger, 2001). TCP scans are prevalent due to the protocol's widespread use in Internet communication. Scanners might send SYN packets to initiate connections or use ACK, FIN, or NULL scans to probe for responses (Bhuyan, Bhattacharyya, and Kalita, 2011). Many types of TCP scans are easy to detect due to their connection-oriented nature (Bhuyan, Bhattacharyya, and Kalita, 2011). UDP scans, though less common, are used to detect services running on UDP ports. These scans are typically more challenging to detect and analyse due to the stateless nature of UDP (Kurose and Ross, 2017). ICMP scans are used primarily for network mapping and host discovery (Roger, 2001). These scans can identify which hosts are up and responsive to ICMP requests. Determining the protocol used in a scanning activity involves analysing the packet data and headers captured in firewall logs. Specific tools and techniques can parse these logs, highlighting the protocol and nature of the scan.

The identification of tools and techniques used in port scanning activities is another critical component of activity identification. Bhuyan, Bhattacharyya, and Kalita (2011) provide an overview of various tools that are commonly used by attackers and security researchers alike to perform port scans. Recognizing the signatures of these tools can help differentiate between legitimate and malicious activities. Nmap, for example, is one of the most widely used tools for port scanning. It can perform a variety of scans, including SYN scans, UDP scans, and OS detection (Nmap Project, 2024). Nessus and Greenbone are popular vulnerability scanners that can identify security weaknesses in a network. Each of these tools leaves distinctive signatures in log data. For instance, Nmap's SYN scan sends a series of SYN packets to different ports, which can be recognized by their characteristic patterns in firewall logs. Nessus and Greenbone scans typically generate more extensive and detailed logs due to the comprehensive nature of their vulnerability assessments. By understanding these signatures, we can accurately identify the tools and techniques used in scanning activities, providing further context for subsequent analysis steps.

3.2.2 Step 2: Intent assessment

Assessing the intent behind scanning activities is a critical step in classifying them as either legitimate or malicious. This distinction of intent is fundamental to enable appropriate responses and mitigate potential threats. Port scans can serve benign purposes, such as research conducted by universities or security firms, or they may be employed by malicious actors seeking vulnerabilities. However, it is important to recognize that the criteria used to differentiate between these intents are indicators rather than definitive proofs, as skilled attackers can often bypass such measures.

The source of a scan is a significant indicator of its intent. Benign scans often originate from known IP addresses linked to reputable institutions, such as universities, research organizations, or security firms (AlAhmadi and Martinovic, 2018). In contrast, malicious scans frequently arise from unknown, anonymized, or blacklisted IPs, particularly those associated with regions notorious for cybercrime (AlAhmadi and Martinovic, 2018). The frequency and volume of scans further differentiate legitimate from malicious activities. Low-frequency scans conducted at regular intervals typically align with research or security monitoring (Bhuyan,

Bhattacharyya, and Kalita, 2011). Conversely, high-frequency scans exhibiting irregular patterns are often indicative of attempts to quickly identify vulnerabilities (Hindy et al., 2021).

The selection of targets can also reveal the intent. Legitimate scans are usually focused on specific, consented networks or systems, often preceded by communication with the target (Hindy et al., 2021). Malicious scans, however, tend to have a broader scope, targeting multiple unrelated IP addresses in a generalized search for vulnerabilities (Ring, Landes, and Hotho, 2018). The patterns and techniques employed in scanning further illustrate this divide. Legitimate activities typically utilize standardized tools, such as Nmap, with well-documented methodologies (Roger, 2001). Malicious actors, by contrast, rely on advanced evasion techniques, unconventional scan types such as SYN, FIN, or Xmas scans, and less common or proprietary tools (Hindy et al., 2021).

Communication and transparency are key attributes of legitimate scanning. Researchers and security firms often notify network administrators about upcoming scans or publicly share their results to demonstrate accountability (Ring, Landes, and Hotho, 2018). Malicious actors, however, avoid communication, obfuscating their intentions and the origins of their scans (Ring, Landes, and Hotho, 2018). Timing and duration of scans also provide valuable context. Benign scans are often conducted during off-peak hours to minimize their impact, with limited durations (Hindy et al., 2021). In contrast, malicious scans may occur continuously over extended periods, often during peak hours to mask their activity amidst regular traffic (Hindy et al., 2021).

Legal and ethical compliance further distinguishes these activities. Legitimate scanners adhere to established guidelines, often securing prior permission to conduct their activities (Stallings, 2016). Malicious actors, on the other hand, disregard legal and ethical norms, performing unauthorized scans. Geolocation serves as another important criterion. Scans originating from regions associated with reputable organizations are more likely to be benign (Bhuyan, Bhattacharyya, and Kalita, 2011), whereas those from regions with lax cybersecurity regulations or high cybercriminal activity raise suspicion (Bhuyan, Bhattacharyya, and Kalita, 2011).

Finally, the context in which a scan occurs is crucial. Scans associated with recognized projects, academic research, or security initiatives are indicative of legitimate intent (Stallings, 2016). On the other hand, scans lacking context or linked to subsequent exploitation attempts suggest malicious behaviour (Ring, Landes, and Hotho, 2018). Assessing these various criteria enhances the accuracy of port scan classification, enabling more effective responses.

Security research and penetration testing often involve activities that may resemble malicious scans but are conducted for legitimate purposes. These activities are typically characterized by a clear scope, prior authorization, and a detailed follow-up report or action plan. Context is essential for distinguishing these scans from malicious ones. For example, scans originating from known security firm IP ranges are likely part of legitimate assessments. Additionally, the methodologies and patterns employed in these scans adhere to recognized standards, such as those outlined by the Open Worldwide Application Security Project (OWASP) or the National Institute of Standards and Technology (NIST).

Detecting scans with malicious intent requires recognizing patterns commonly associated with known attacks. These include high-frequency scanning, targeting multiple ports or IP addresses, and the use of tools favoured by attackers. Correlating scan patterns with threat intelligence databases, such as those provided by MISP, helps identify potential threats by linking scanning activity to known tactics, techniques, and procedures (TTPs). Historical data and records of prior incidents provide further context, offering valuable insights into the likelihood of malicious intent (Al-Haija, Saleh, and Alnabhan, 2021). By leveraging these criteria and contextual tools, it is possible to classify scanning activities more effectively and respond to threats with greater precision.

3.2.3 Step 3: Threat actor attribution

The third step in the classification process is attributing the detected activity to known threat actors. It should be noted, that any attempt at attribution is an estimate and can never be certain (Conti, Dargahi, and Dehghantanha, 2018). This involves comparing the activity's signature against databases of documented threat actor behaviours. Techniques for matching activities include analysing the TTPs associated with the scan and using threat intelligence platforms like MISP. These platforms provide detailed profiles of threat actors, including their preferred tools, techniques, and targets. By matching observed activities with known threat actor profiles, analysts can gain insights into the potential source and motivations behind the scan. This information is crucial for understanding the broader threat landscape and anticipating future actions. Classifying activities under specific threat actor profiles involves understanding the TTPs associated with different actors. For example, certain groups may be known for using particular scanning tools or targeting

specific industries or regions. Example profiles might include nation-state actors who conduct sophisticated and persistent scans to gather intelligence or financially motivated cybercriminals who perform broad sweeps to identify vulnerable targets for ransomware attacks. Understanding these profiles helps in tailoring defence strategies and prioritizing responses based on the actor's known behaviours and objectives.

3.2.4 Step 4: Information gain assessment

Assessing the potential information that an attacker could gain through a scanning activity is critical for understanding the threat's severity. Information exposed might include open ports, running services, and system vulnerabilities. Attackers could use such information to search for known vulnerabilities, discover or buy zero-day vulnerabilities, or even introduce new vulnerabilities via supply chain attacks. To assess the scope and value of exposed information, defenders should analyze the scan's results and cross-reference them with known vulnerabilities. Tools like vulnerability scanners can simulate the attacker's perspective, providing insights into what information is at risk. By knowing what the attacker knows and doing the same research an attacker would do to discover vulnerabilities, one can estimate likelihood and targets for follow up attacks. Ranking the potential severity of exposed information helps prioritize threat responses. The severity can be categorized as low, medium, or high based on the criticality of the exposed information and the potential impact on the network. Criteria for determining the level of exposure include the sensitivity of the information, the exploitability of identified vulnerabilities, and the potential damage that could result from an exploit. High-severity exposures typically involve critical systems or sensitive data, requiring immediate attention (Ananin, Nikishova, and Kozhevnikova, 2017).

3.2.5 Step 5: Pre-Damage activity analysis

The fifth step involves evaluating whether the detected activity alone could result in damage or if additional steps are required. This assessment helps in understanding the immediate threat posed by the activity. Activities that can cause damage on their own, such as direct exploitation attempts following a scan, require urgent responses. Conversely, activities that are part of a reconnaissance phase might be less immediately dangerous but still warrant close monitoring (Kurose and Ross, 2017). Determining if the activity is part of a multi-step attack involves identifying its role within a broader attack strategy. Techniques for this analysis include mapping the activity to known attack models and analysing the sequence of events. Recognizing the activity's place in an attack chain helps in predicting subsequent steps and implementing preventive measures to disrupt the attack progression.

3.2.6 Step 6: Threat intelligence generation

Generating threat intelligence from detected activities involves documenting IoCs associated with the activity. IoCs are data points that indicate the presence of a potential security threat, such as specific IP addresses, file hashes, or behavioural patterns. Methods for extracting and cataloguing IoCs include automated analysis tools and manual investigation. Sharing IoCs with relevant stakeholders enhances collaborative defence efforts and improves overall threat detection capabilities (Conti, Dargahi, and Dehghantanha, 2018). Documenting identified TTPs for inclusion in threat intelligence feeds helps in understanding and countering threats. TTPs provide a comprehensive view of how attackers operate, allowing for the development of targeted defences. The importance of TTPs lies in their ability to inform proactive security measures and improve the overall resilience of the network against similar future threats (Lagraa, and François, 2017).

3.2.7 Step 7: Response and prevention

Deciding if an activity warrants a response involves evaluating its intent, potential harm, and match with known threats. Criteria for determining the necessity of a response include the severity of the potential impact and the likelihood of the activity being part of a malicious campaign. Proposing strategies to prevent or mitigate the activity involves considering the balance between the effort required and the effectiveness of the measures. Effective prevention strategies might include updating firewall rules, implementing IDS, and conducting regular security assessments. Considering the potential harm to legitimate activities is also important. Strategies should aim to minimize disruptions to normal operations while enhancing security (Al-Haija, Saleh, and Alnabhan, 2021).

3.2.8 Step 8: Legitimate use consideration

Providing guidelines for distinguishing between legitimate and malicious use of similar techniques helps in managing future instances effectively. Developing policies and protocols that address both security and

operational needs ensures a balanced approach to threat management. Evaluating the cost and impact of mitigating actions against their benefits by assessing the resources required versus the potential security improvements ensures that mitigation efforts are proportionate and effective. Providing guidelines for distinguishing between legitimate and malicious use of similar techniques helps in managing future instances effectively. Developing policies and protocols that address both security and operational needs ensures a balanced approach to threat management.

4. Conclusion

The classification of port scanning activities plays a critical role in enhancing cybersecurity and fortifying defensive measures against potential threats. Port scanning, while a valuable tool for legitimate network security assessments, is frequently exploited as a reconnaissance method by malicious actors. By employing a taxonomy-based approach, the distinction between benign and malicious scans can be systematically assessed, enabling the generation of actionable CTI. Key criteria, such as the source, frequency, patterns, transparency, and historical data of scans, serve as indicators to determine their intent and potential threat levels. This comprehensive evaluation not only aids in identifying the immediate implications of port scans but also contributes to understanding their place within larger attack frameworks.

4.1 Vignette: Classifying and Responding to a Port Scan Attack

In a typical evening at a mid-sized tech company, Alice, a cybersecurity analyst, notices an unusual pattern of network activity. Her monitoring tools indicate a series of connection attempts to multiple ports on a single server hosting critical customer data.

Step 1: Activity Identification

Alice identifies this activity as a vertical port scan, characterized by the sequential probing of multiple ports on a single host. She verifies this using network traffic analysis tools, which confirm the signature of a vertical scan.

Step 2: Intent Assessment

The source IP address of the scan is traced back to a known cybercrime group. Given the targeted server's sensitive content, Alice assesses the intent as likely malicious rather than benign.

Step 3: Threat Actor Attribution

By comparing the scan's signature with their threat intelligence database, Alice tentatively attributes the activity to the "ShadowClaw" hacking group, known for sophisticated attacks on tech companies. She acknowledges that this attribution is just an estimate and not definitive.

Step 4: Information Gain Assessment

Using a vulnerability scanner, Alice simulates the scan to reveal exposed vulnerabilities, including an outdated service with known vulnerabilities. This information could be exploited by attackers to identify a potential entry point.

Step 5: Pre-Damage Activity Analysis

While the scan itself has not caused immediate damage, Alice recognizes it as a precursor to a potential attack. She understands that additional steps, such as exploiting vulnerabilities, would be required for actual system compromise.

Step 6: Threat Intelligence Generation

Alice documents the Indicators of Compromise (IoCs), including the source IP address, the specific port scanning pattern, and the targeted ports. She updates their threat intelligence platform with this new information to enhance future detection capabilities.

Step 7: Response and Prevention

Given the high-value target and the attribution to a known threat group, Alice initiates the incident response plan. She blocks the source IP and begins hardening the targeted server by updating vulnerable services and implementing additional security measures.

Step 8: Legitimate Use Consideration

Alice collaborates with the IT team to develop stricter firewall rules and policies for authorized port scanning activities. This ensures that legitimate security testing can continue while improving the overall security posture of the organization.

4.2 Evaluation of the Eight-Step Classification Process

The proposed eight-step classification process offers a structured methodology to analyse scans, assess their impact, and generate proactive defence strategies. It integrates the recognition of threat actor behaviours, assessment of information exposure, and consideration of preventive measures, while balancing operational requirements with security needs. Furthermore, distinguishing between legitimate activities, such as research or internal audits, and malicious scans is crucial to avoid disrupting essential network operations.

This structured approach to port scan analysis underscores the importance of integrating CTI, historical data, and collaborative defence mechanisms in modern cybersecurity strategies. By advancing the understanding and classification of port scans, organizations can not only enhance their situational awareness but also implement informed, efficient, and cost-effective defences against evolving cyber threats.

Acknowledgement

The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the programme of “Souverän. Digital. Vernetzt.”. Joint project 6G-Life, project identification number: 16KISK002

References

Abu Bakar, R. and Kijsirikul, B., 2023. Enhancing Network Visibility and Security with Advanced Port Scanning Techniques. *Sensors*, *23*(17), p.7541.

AlAhmadi, B.A. and Martinovic, I., 2018, May. MalClassifier: Malware family classification using network flow sequence behaviour. In *2018 APWG Symposium on Electronic Crime Research (eCrime)* (pp. 1-13). IEEE.

Al-Haija, Q.A., Saleh, E. and Alnabhan, M., 2021, December. Detecting port scan attacks using logistic regression. In *2021 4th International symposium on advanced electrical and communication technologies (ISAECT)* (pp. 1-5). IEEE.

Ananin, E.V., Nikishova, A.V. and Kozhevnikova, I.S., 2017. Port scanning detection based on anomalies. *2017 Dynamics of Systems, Mechanisms and Machines (Dynamics)*, pp.1-5.

Bhuyan, M.H., Bhattacharyya, D.K. and Kalita, J.K., 2011. Surveying port scans and their detection methodologies. *The Computer Journal*, *54*(10), pp.1565-1581.

Conti, M., Dargahi, T. and Dehghantanha, A., 2018. *Cyber threat intelligence: challenges and opportunities* (pp. 1-6). Springer International Publishing.

De Vivo, M., Carrasco, E., Isern, G. and De Vivo, G.O., 1999. A review of port scanning techniques. *ACM SIGCOMM Computer Communication Review*, *29*(2), pp.41-48.

Hindy, H., Atkinson, R., Tachtatzis, C., Bayne, E., Bures, M. and Bellekens, X., 2021. Utilising flow aggregation to classify benign imitating attacks. *Sensors*, *21*(5), p.1761.

Kurose, J. and Ross, K., 2017. Computer networking: A top-down approach, global edition.

Lagraa, S. and François, J., 2017, May. Knowledge discovery of port scans from darknet. In *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)* (pp. 935-940). IEEE.

Nmap Project (2024) “Nmap Reference Guide”, [online], Nmap Project, https://nmap.org/book/man.html.

Patel, S.K. and Sonker, A. 2016. Rule-based network intrusion detection system for port scanning with efficient port scan detection rules using snort. International Journal of Future Generation Communication and Networking. 9(6), pp.339--350.

Ring, M., Landes, D. and Hotho, A., 2018. Detection of slow port scans in flow-based network traffic. *PloS one*, *13*(9), p.e0204507.

Roger, C., 2001. Port Scanning Techniques and the Defense Against Them. *White Paper, SANS Institute*, pp.1-8.

Schäfer, J. and Drozd, M. 2011. Detecting network attacks using behavioural models. Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems. 2, pp.753--758.

Stallings, W., 2016. *Network security essentials: applications and standards*. Pearson.

u Nisa, M. and Kifayat, K., 2020, October. Detection of slow port scanning attacks. In *2020 International Conference on Cyber Warfare and Security (ICCWS)* (pp. 1-7). IEEE.