Detecting Lateral Movement: A Systematic Survey

Christos Smiliotopoulos[a,*], Georgios Kambourakis[a], Constantinos Kolias[b]

[a]*Department of Information and Communication Systems Engineering, University of the Aegean, Karlovasi 83200, Samos, Greece*
[b]*Department of Computer Science, University of Idaho, Idaho Falls, ID 83402, USA*

Abstract

Within either the cyber kill chain or MITRE ATT&CK frameworks, Lateral Movement (LM) allows adversaries to progressively move deeper into a system in seek of high-value assets. Although this timely subject has been studied in the cybersecurity literature to a significant degree, so far, no work provides a comprehensive survey regarding the identification of LM from mainly an intrusion detection system (IDS) viewpoint. To cover this noticeable gap, this work provides a systematic, holistic overview of the topic, paying also special attention to Internet of Things (IoT) ecosystems. The survey part, spanning a time window of eight years and 53 articles, is split into three domains, namely, Endpoint Detection and Response (EDR) schemes, machine learning oriented solutions, and graph-based strategies. On top of that, we reveal interrelations, mapping the progress in this field over time, and offer key observations that may propel LM research forward.

Keywords: Lateral Movement, Advanced Persistent Threat, Attacks, Network Security, IoT.

1. Introduction

Based on an initial point of compromise, typically through dropping malware or exploiting a vulnerability in a device or application, Lateral Movement (LM) involves moving deeper in terms of data or upwards in regard to access. Oftentimes, the attacker's goal is to remain in the system as an Advanced Persistent Threat (APT), attempting to gain as much loot as possible. What is more, in the Internet of Things (IoT) era, LM is gaining increased attention; attackers can exploit a plethora of IoT devices for achieving LM. That is, IoT devices represent a particularly alluring target for a variety of threat actors aiming to move laterally and create persistence within any network, especially enterprise ones. Namely, IoT seem ideal not only for obtaining initial foothold and persistent remote access, but also for gaining new permissions and user privileges in the breached environment; such devices are typically undersecured, they operate in a 24/7 basis, they are omnipresent and frequently in sensitive parts of the network, and, in many instances, not sufficiently updated or monitored, following an install-and-forget mentality [1]. In a typical scenario, a malicious binary can incorporate a scanner module to perform LM; in the presence of a vulnerability, the malware will propagate to, say, a local printer, Wi-Fi bulb, or smart humidifier with minimal built-in protections.

LM is a key tactic in the context of modern threat modeling and associated cyberattack methodologies; the MITRE ATT&CK matrix for enterprise classifies LM under ID TA0008, identifying nine basic techniques. However, even though a significance mass of contemporary research works address LM from either a defensive or offensive viewpoint, the literature lacks a comprehensive, contemporary survey of schemes proposed to detect LM, mainly in the context of an Intrusion Detection System (IDS). In fact, although some studies [2, 3, 4, 5] addressed LM as part of the general discussion on APT, they only did so in an abstract manner. In brief, Stojanovic et al. [2] focused on the description of the disposable feature selection and engineering techniques and the detailed presentation of the datasets leveraged by APT-dedicated IDS schemes. Tatam et al. [3] contributed a Systematic Literature Review (SLR) on the effectiveness of threat modelling

*Corresponding author
Email address: csmiliotopoulos@aegean.gr (Christos Smiliotopoulos)

frameworks regarding APT, while Talib et al. [4] presented a comparative study regarding the strengths and weaknesses of ML APT-related IDS schemes relied on the identification of APT's beaconing behavioral patterns. Furthermore, only Chen et al. [5] addressed the subject of APT models under the concept of IoT interconnected devices. Namely, they summarized the most impactful models on the area in terms of anomaly detection, signature-based, and hybrid ML identification of threats. All in all, the aforesaid studies although addressed the subject of APT identification and detection, they only touched upon IDS destined to LM.

Given the above, the work at hand seeks to address this noteworthy literature gap by offering a systematic, complete overview of this field of research. From a methodological standpoint, the survey spans eight years, from 2015 to 2023, covering all major bibliographic databases. Precisely, with reference to the LM detection methodology used, our analysis regarding the included literature works is split into three parts; Endpoint detection and response (EDR) log-based policy schemes, Machine Learning (ML)-powered schemes, and graph-based schemes. Additionally, given that IoT devices represent an enduring favorite among attackers attempting to move laterally, each of the abovementioned parts includes a separate section looking at schemes applied to the IoT ecosystem. Equally important, we provide extensive discussions for each category of schemes identifying key methodologies and techniques, benchmark datasets, challenges, and shortcomings as well as interdependencies among the various studies.

The remainder of this paper is organized as follows. The next section details the methodology. Sections 3 to 5 discuss the included works per category of schemes, namely EDR log-based policy-based, ML-based, and graph-based, respectively. Section 6 offers key observations on the surveyed works, including prevailing methodologies, inadequacies and challenges. Section 7 wraps up and points out the main directions for future research. For easier guidance throughout the manuscript, a list of abbreviations is included at the end of the article.

2. Methodology

As already mentioned in section 1, the current work contributes an SLR, discussing the various available techniques concentrating on the detection of LM incidents. Such techniques are basically used from and under the prism of small scale or corporate EDR policies, up to the more complicated advanced ML and graph-based algorithmic models. To this direction, equally important are the various testbeds, which most of the time rely on one or more benchmark datasets comprising either system logs or network traffic or both. Precisely, in most of the analyzed works in this survey, datasets are used to test the identification, evaluation and classification standards of pertinent threats for each LM detection model under the simulation of real-life scenarios.

As already outlined in section 1, the scope of the SLR at hand can be outlined as follows:

- To identify the various EDR, ML algorithms, graph-based techniques and testbeds that have been incorporated in IDS schemes towards the identification of LM adversary incidents.

- To identify and classify the applicability and effectiveness of the various embodied IDS schemes per utilized identification technique, namely EDR policies and log-based concepts, supervised or unsupervised ML algorithms and graph-based models.

- To point out the types of the utilized testbeds, listing briefly the hardware and software tools, the equipment, and finally the various metrics through which the evaluation of the model's effectiveness has been conducted.

- To identify shortfalls, unattended issues and future potentials upon the field of LM incidents identification.

The SLR conforms to the methodology presented in [6], outlining the basic steps that should be used towards the design and execution of an SLR. Specifically, the steps that were followed at first stage under the present work have as follows:

- A variety of the most popular literature databases, namely Scopus, ACM, IEEE Xplore, Science Direct, and Springer Nature, were scrutinized in search of the related to LM works.

- The aforementioned databases were queried based on the combination of the following dedicated to the studied subject keywords: "lateral movement" **AND** ("endpoint detection and response policy" **OR** "EDR policy" **OR** "machine learning" **OR** "supervised machine learning" **OR** "unsupervised machine learning" **OR** "graph-based analysis" **OR** "graph based analysis") **AND** "security" **AND** ("intrusion detection systems" **OR** "IDS" **OR** "anomaly detection"). An additional query was used to scan the literature for relevant works focused on the IoT and IIoT ecosystems: "lateral movement" **AND** "IoT" **AND** "IoT vulnerabilities" **OR** "IoT-dedicated datasets" **OR** "IoT IDS applicable schemes".

- The examined literature spans a period of nine years, i.e. from 2015 to 2023.

- The SLR took seven months to complete, i.e. from April to October 2023.

Second, as presented in Table 1, the selection process was aggregated into the list of the finally selected papers through a set of core inclusion / exclusion criteria.

Table 1: List of inclusion and exclusion criteria.

	Inclusion-Exclusion	Description
Inclusion	Endpoint Detection and Response Policies (EDR policies)	The EDR policies need to focus on LM detection. Works referencing pivoting and general APT identification are also included, as the criteria addressing them could find applicability to LM incidents too.
	Supervised ML models	Consider works that only describe the implementation of ML algorithms towards the exclusive identification of LM adversary events.
	Unsupervised ML models	Due to the limited variety of studies upon the specific category, studies related to LM events will be examined in conjunction with the corresponding ones affiliated to APT and pivoting incidents.
	Graph-based models	Only the dedicated to LM identification literature are considered.
	LM-oriented Testbeds	Any case-study scenario that involves a testbed along with the utilized dataset(s), if any, is examined and analyzed.
	Published papers written in English language.	-
Exclusion	Genre of literature	Book reviews - chapters, conference abstract - information chapters, mini blog reviews, editorials and online discussions, blog discussions, news.
	Not exclusively related to LM	Any other paper that could be categorized to the general category of generic defensive schemes or generalized IDS concepts has been excluded.

3. Endpoint detection and response log-based policy schemes

3.1. EDR policy schemes

The current section briefly reviews the key pertinent literature on the subject of endpoint and log-based detection of LM events. The concentration is on the methodology of each relevant work regarding event-driven identification of the most impactful LM techniques. Particularly, we focus on the collection of system and network related logs to the specified examined attacks, the utilized rule-based policy, and the impact upon the successful incidence response. To facilitate the parsing of the relevant literature, Table 2 recaps the relevant key characteristics per work included in this section. The

various studies are chronologically arranged in ascending order. Finally, Table 3 recapitulates all the public benchmark datasets that were leveraged as proof-of-concept per case-study scenario.

Ki et al. [7] considered the detection of malware Application Programming Interface (API) call patterns under the concept of dynamic analysis and identification of anomaly behavioral signatures that may indicate LM activity. They deployed sequence alignment algorithms in an effort to extract common API call sequence patterns for malevolent functions generated from diverse categories of malware. According to the authors, what made ideal sequence alignment algorithms for malware's behavioral identification was their convenience in extracting similarities and patterns from different incrementally executed sequences, such as malware's API calls. To experiment upon their proposed methodology, the authors created a virtual MS Windows-based environment of hosts, while for hooking the various executables during runtime and monitor API calls, they relied on *Detour* hooking library. As it concerns sequence alignment algorithms, the ClustalX freeware library was imported to trace a malware dataset comprising more than 23K samples. Approximately 2.7K different API calls were identified during the experiment; these were grouped under 26 common categories based on the Microsoft Development Network (MSDN) list. To contribute with the extraction of the common API call sequence patterns among the executed malware, they introduced a custom "Longest Common Subsequences (LCSs)" formula, through which a signature based on the longest API call is created and attributed to each piece of malware. Despite the effectiveness of the proposed scheme, the formula was able to identify and categorize only user-level APIs and not the kernel-level ones.

The work of JPCERT/CC [8, 9], which is the first leading Computer Security Incident Response Team (CSIRT) established in Japan, was also one of the first security organizations around the world that conducted a full-fledged survey upon the categories of logs produced during the execution of LM techniques. Such techniques aim to initiate reconnaissance and identification of potential vulnerabilities upon the targeted host, they continue with the infection of the target to end up to the exploitation of critical information for malevolent purposes. JPCERT/CC first identified the standard schemes that adversaries follow in most of the cases during their targeting of network facilities. The first scheme is used to infiltrate fundamental network information, as those were collected from the infected network infrastructure using tools as "ipconfig", "systeminfo" or MS Windows core applications such as "Windows Event Viewer" and "Sysmon". The second focused on the identification of valuable information regarding the network hosts, namely, OS version, account-domain characteristics, open ports, and many others, with a variety of tools like "net". Finally, the most vulnerable host was targeted by means of a credential exploitation procedure with tools like "Mimikatz" and "pwdump", towards the expansion of the infection to all network users. Such adversary patterns leave a far from negligible variety of diverse log files, which can be used by audit teams towards the identification of the existence of LM. To evaluate the aforesaid log files, the most prominent at that time LM techniques, namely exploitation of remote services, password exfiltration, privilege escalation, capturing windows active directory database and more, were applied on both compromised servers and targeted clients, related to each attack method. Logging information was collected and categorized via Sysmon [10] and MS Windows audit-policy. The work concluded with the proposition of an MS Windows endpoint detection and response audit policy, allowing the optimal collection of useful information related to potential LM methods. The final report was published in June 2017 and updated as v2 [9] in Dec. of the same year.

Mavroeidis et al. [11] experimented on the creation of a data-driven threat classification methodology, that is relied on the continuous aggregation and analysis of voluminous log files collected from Microsoft's Sysmon security information and event management (SIEM) tool. As a first step, the authors highlighted the importance of the notions of "threat intelligence" and "threat information sharing", as part of a security ontology. Precisely, that ontology forms a collective cognitive library on which all the related to cyberthreat detection identified aspects and elements are imported, based on well-defined semantic concepts and their relationships. Recall that the creation of an ontology, allows the aggregation of information derived from various multipurpose sources in a singled knowledge database. According to the authors,

that database may evolve over time periods for supporting as reasoning evidence the logical sequence of identifying and recording malevolent inconsistencies and events. Overall, the authors' work is twofold. First, the authors concentrated on the deployment of their proposed Cyber Threat Intelligence Ontology (CTIO), which was based on the Cyber Threat Intelligence (CTI) model, also published by them in [12]. CTIO may act as an ensemble of information from numerous multifarious sources, from low-level technical log-based events to more advance high-level observations and threat actors. The aforementioned composite information grid aims to support the decision-making procedure under the concept of a security policy. To evaluate the robustness of the proposed ontology, CTIO was incorporated alongside a threat assessment system towards the classification of event-logs generated by Sysmon in four distinct threat categories, namely high, medium, low, and unknown.

Berady et al. [13] relied on the analysis of the distinct elements that contribute to the success or failure during the adversaries "threat hunting" EDR operation. Namely, the early identification of malevolent actor's areas of applicability and the qualification of their level of effectiveness in the shortest possible time are included among the most impactful factors for a successful EDR policy. To this direction, the authors presented an element-based threat hunting model based on the common acceptance of the conclusion that the attacker and defender should mutually understand each other. The analysis was conducted from both an offensive and defensive viewpoint, allowing the collation of the two contradictory perceptions regarding the same attack vector. The proposed model contributes in a twofold way to both adversaries and defending EDR "Blue teams". Regarding the former, the malevolent entities become more aware of the traces left during the execution of their attacking techniques. As it concerns the latter, "threat hunting" is improved though the identification and elimination of False-Positives (FP) and the implementation of a realistic and calibrated log-based policy oriented around forensic and cybersecurity investigations. To evaluate their model, they conducted experiments based on offensive and defensive perspectives. At first, offensive experiments were done upon the APT29 dataset, that is included in the Mordor Project of pre-recorded event-related logs of malicious activities [14]. Note that the Mordor Project relies on the real-world APT29 threat, and was conducted based on a scenario designed and presented by MITRE's ATT&CK evaluation tactics list. The two-staged attack emulation scenario incorporated initially the reconnaissance and the compromise of the target via an injected toolkit, followed by the extraction of sensitive information from the targeted host. As it concerns "Blue teams", they used Sysmon to monitor the targeted host's network activity and produce a dataset comprising two-days of log recording. The paper concludes with the proposition of an "Indicators of Compromise" list of elements. This contributes to the enhancement of defender's base knowledge of the adversaries best practices, enhancing seemingly the preventive identification of threats. We argue that, although this work examined only APT attacks against the dedicated APT29 subset, the underlying ideas are generic and can be applied to LM techniques as well.

Matsuda et al. [15] proposed a Dynamic Link Library (DLL)-oriented methodology for detecting malicious files upon logsets collected via Microsoft's Sysmon tool. To create a rich dataset, numerous DLL files were collected and analyzed through four different tools namely, China Chopper, Mimikatz, PowerShell Empire, and HUC Packet Transmitter. The systematic observation of the aforesaid tools' special characteristics and traffic, revealed the existence of significant differences of the specific files among various versions of the MS Windows Operating System (OS). The proposed methodology concluded with the extraction of the "common DLL list", which comprises a collection of the most commonly loaded per malevolent tool DLL files, not dependent on the MS Windows OS version. Moreover, as an extension to that list of DLL files, a DLL logs-oriented detection method based on the open-source ELK Stack SIEM was introduced. According to the authors, the use of the well-known ElasticSearch engine, allows real-time monitoring of the DLL information loaded on different hosts, leveraging the detection of potentially malicious files through their comparison with the aforesaid common list. The detection accuracy of the proposed methodology was evaluated by means of the same four tools. The generated logs were filtered via ELK Stack in a real-time manner, revealing a promising detection rate. Overall, it can be said that

the enhancement of the DLL detection policy, particularly focused on each LM method's Sysmon event logs, expands the detection accuracy and diminishes false positives.

The ELK Stack was also used by the authors in [16, 17] for the analysis of massive log records and the identification of malicious behavior. Specifically, Jaim et al. [16] deployed various LM techniques on a sandbox Windows-based networking virtual lab, producing massive log-based traffic related to LM events. The latter were generated from the execution of 24 LM case studies, namely PSExec, Mimikatz (Golden/Silver ticket, Pass the Hash, Pass the Ticket, etc.), PWDump, vssadmin, etc. The traffic was captured and filtered via ELK stack towards the final identification and detection of the behavioral characteristics of the tools exploited by LM techniques. The authors concluded with a presentation of a proof-of-concept of the most prominent EventID characteristics, as those were derived from each of the 24 executed LM tools and captured via ELK stack. On the other hand, Rajesh et al. [17] deployed an ELK stack-based massive data processing pipeline to collect, analyze, and identify anomalies on voluminous log-based data structures. The massive log-sets were collected through "logstash", while ElasticSearch was used for the querying and filtering of malicious traffic among the collected samples. The authors' experiments include the examination of local outlier factor for MySQL queries, numerical, and Boolean values, classification and regression detection. We argue that although that paper incorporated ML notions that should be categorized in the cases of Section 4.1, the vague and abstract representation of the reasoning upon the results render it a good fit for the EDR category.

El Hadidi et al. [18] dealt with the detection of LM and APTs, as those are executed through any version of the well-known "Mimikatz" tool, and via the identification of Mutex objects and loading DLL files through the numerous generated logs. According to the authors, during their LM, APT groups follow five basic stages, namely infection, compromise, reconnaissance, credential theft, and LM. The authors' focus was on the fourth stage of the credential heist, as during its execution, the adversaries exploit the various dedicated for that purpose tools, such as Mimikatz, leading to the production of Mutex objects and DLL as evidences. To evaluate the robustness of their proposed identification scheme, they executed four different versions of Mimikatz on three different MS Windows Server editions, namely, 2008 R2, 2012 R2, and 2016 Standard Edition. Their findings, through the persistent analysis of the collected logs, revealed the existence of dedicated to Mimikatz Mutex objects, such as "conhost.exe", that permitted increased accuracy levels towards the identification of LM.

Agarwal et al. [19] presented a custom-built LM and pivoting EDR tool dedicated to the Linux platform. Their tool expands on the "Osquery" and "Elastic" open-sourced ones, being capable to aggregate endpoint logs at a pre-configured common server. In this way, the log-oriented correlation between events stemming from various endpoint devices permits the investigation and detection of LM and pivoting incidents. According to the authors, the idea for developing their Linux EDR tool derived by the limited variety of tools and policies available to enable monitoring and threat detection on Linux endpoint devices and servers. In general, the collection of logs in Linux environment may be accomplished through the use of "system logs" or "audit logs", respectively. Both the aforementioned methods suffer from various disadvantages. The former lacks the capability to identify user information towards the identification of the threat actor who caused the malicious event; as in case of privilege escalation, the attacker will always be captured as the root user. On the other hand, although the audit logs allow the identification of the perpetrator and provide relevant alerting, it still falls short in recognizing the endpoint device through which the attack was executed. In this respect, the major contributions of the EDR tool assemble to continuous monitoring and information gathering and analysis of the collected information in a historical archiving manner towards knowledge gaining and proactive detection of malicious actors through logs iteration. Nevertheless, the proposed tool is incapable to provide real-time incremental information regarding the events captured during the monitoring and capturing procedure. To evaluate the vigor of their contribution, the authors executed CVE-2019-2725 vulnerability of Oracle Weblogic server; through it, an adversary may gain control of a targeted host via the

execution of arbitrary OS commands.

Niakanlahiji et al. [20] presented *ShadowMove*, a corroboration strategy for APT and LM techniques execution and host compromising. Recall that APT attacks leverage numerous techniques and strategies towards compromising an enterprise networking environment through LM. Although such techniques tend to be omnipresent and hard to be encountered, due to the consecutive evolution of malevolent tools and the variety of threat actors, they are limited to incorporate a set of legacy unique features related to their core functionality. Precisely, the execution of APT's LM has as prerequisites the creation of multiple new connections, the performing of authentications, or the requirement for process injections. Via such procedures, the adversary may stealthily move laterally within a networking facility, incrementally through privilege escalation. According to the authors, *ShadowMove* is designed to work stealthily, overcoming the aforementioned APT's shortcomings. Namely, the proposed methodology neglects the legacy features of APT's functioning in favor of a novel approach related to "socket duplication", through which a malicious process silently leverages TCP connections of benign equivalents. Moreover, the presented LM strategy, permits APT threat actors to move stealthily among hosts in enterprise networks without being identified or detected by host-based, network level, or IDS systems. *ShadowMove* does not inject in any manner arbitrary malevolent code or commands on the benign processes. Instead, it passively monitors the targeted networking host's traffic to identify established connections. By this, the attacker avoids the creation of new connections, and therefore the need for new authentication. On top of that, *ShadowMove* is not restricted to the established connections only, but also to application protocols such as "WinRM" and "FTP", both of which allow the injection of remote commands on a compromised remote server. The effectiveness of *ShadowMove* was successfully evaluated on MS Windows 10 and Ubuntu 18.0.4 with the majority of the stealthiness being presented on Windows systems instead of Linux. They finally deployed five top-notch antivirus products (McAfee, Norton, Webroot, Bitdefender, and Windows Defender) and two EDR systems (CrowdStrike Falcon Prevent and Cisco AMP) to confirm that the proposed methodology evade detection. Although this work is dedicated to implement an offensive methodology for compromising enterprise networks via APTs and LM techniques, we opt to include such schemes as new offensive techniques do propel the corresponding defensive systems.

Smiliotopoulos et al. [21] presented a novel initialization rule-based EDR policy for the Sysmon tool dedicated to the identification and detection of LM events within the MS Windows ecosystem. They elaborated on ordinal identified patterns related to LM events through the execution of the most frequently encountered LM techniques, as those are presented on MITRE's ATT&CK database of adversaries, tactics, and techniques [22]. Specifically, regardless the complexity of any LM method, the consecutive experimentation with LM tactics in [22] revealed that, during the occurrence of such conventional LM incidents, the attacker tends to repeatedly manipulate a limited number of penetration tools. Such tools strive for the identification and enumeration of the targeted host and the exfiltration of crucial information related to OS characteristics; typical tools to this end are ipconfig, systeminfo, Mimikatz, etc. Moreover, the adversary aims the acquisition of credential information related to the clearance level of the targeted host with tools such as Mimikatz, pwdump, LazagneProject, or even its infection with dedicated to credential theft malware. To identify the fundamental technical particles of the most common LM attacks, the authors exploited a testbed with nine LM techniques of diverse targeting subject, including four variant of the "Exploitation of Remote Services" (T1210, T1021), "Pass the Hash" (T1550.002), "Pass the Ticket" (T1550.003), "Golden Ticket" (T1558.001), "Silver Ticket" (T1558.002), and post exploitation on stored passwords with Lazagne Project (T1555, T1003, T1552). Experimentation over the aforementioned LM techniques revealed numerous interrelated features that were incorporated as custom rules in Sysmon's config.xml file. Through the proposed policy oriented initialization, Sysmon acts more than a dedicated to LM SIEM, than just a log monitoring tool. Above that, the authors contributed the LM-oriented log-based LMD-2022 [23] corpus, comprising more than 870K Sysmon event logs. To evaluate the proposed rule-based policy, an extensible Python .evtx file analyzer,

dubbed PeX, was developed and assessed in terms of TP and FP rates over the LMD-2022 logset. The PeX tool can be used towards automatizing the parsing of voluminous log files in .evtx format.

Noor et al. [24] addressed the subject of massive sets of data manipulation as part of raw logs auditing during forensic evidence content approximation and defense of information computing facilities. The authors expanded the already conducted work on the subject and contributed to the extraction of sufficiently approximated audit logs through their experimentation upon a variety of threat models. They provided a renewed pool of metrics that acts as a catalyst towards the quantification of audit log forensic evidence validity. These metrics can be used to measure the usefulness of logs under different test-case scenarios. In addition, they identified independently a preliminary entry point in the approximation design space techniques related to the elimination of, typically to system, activity related logs. Precisely, any log-oriented event that is related to typical system activity may be forced to approximation, while events identified as malicious are preserved with lossless fidelity for further analysis. This entry point approach was incorporated in "LogApprox" technique, which aims at the preservation of attack-related forensic evidence through logs. This is achieved through the creation of an extensive record comprising "process-to-process" and "process-to-network" dependencies, through which the forensic causative analysis process is aggregated. The efficiency of the proposed scheme was evaluated over the DARPA Transparent Computing Engagement 5 Data Release corpus [25], revealing a promising equilibrium among the reduction of redundant logs and the preservation of attack-oriented information. Moreover, "LogApprox" applicability to identify and reconstruct patterns of adversary components was tested against the "Webmin" exploit APT case-study; the latter comprises a web-based Linux configuration tool that can be leveraged for LM over networking facilities.

Guri et al. [26] addressed the wider thematic area of LM techniques through the study of "USBCulprit", which forms an usb-oriented APT, exclusively designed to breach corporate or governmental air-gapped networks. Namely, "USBCulprit" is classified in the malware category that incorporates LM, spreading, and data exfiltration characteristics, and acts through the exploitation of USB thumb drives. The authors experimentally evaluated the performance of the malware over various case study scenarios. Moreover, the numerous interrelated features of the APT were identified and isolated through reverse engineering techniques. Specifically, the source code for the malware's data collection and the air-gap exfiltration mechanisms were extracted; namely, encrypted payloads, mutex files, registry records, API calls, DLL interrelated files, etc. The authors concluded with the presentation of a number of proposed countermeasures, including policy, software, and hardware mitigation measures.

Mundt et al. [27] focused on the ever rising impact of network based attacks through infamous ransomware, such as WannaCry and NotPetya, including their elevated double extortion versions that in addition exfiltrate valuable data prior to being encrypted. The authors contribute to the subject in a twofold way. At first, they present an automated ransomware mitigation concept that acts in parallel as a first line of defense towards the protection of private and corporate networking facilities, and as a measure to cope with the constantly evolving techniques of adversaries to leverage new victims. Moreover, they implement in Python and evaluate a simulation tool that can dissimulate the most impactful ransomware attacks in favor of knowledge and experience acquisition in advance of the actual occurrence of an incident. To achieve their goal and identify the most current practices regarding data exfiltration, the authors relied on the MITRE ATT&CK [28] adversary tactics knowledge database. The finally presented automated mitigation concept was combined with an Extended Detection and Response (XDR) scheme and continuous Security Orchestrated Automation and Response (SOAR) practices, towards the automation and improvement of protective measures.

Mahmoud et al. [29] presented *APTHunter*, a detection system dedicated to the identification of APTs during the initial stages of a system's breach. Recall that APT is a generalized notion representing numerous sophisticated cyberattacks, among which LM techniques are also incorporated. APTs typically target impactful corporate and governmental targets of high value. The difference between *APTHunter* and other data analysis approaches of common provenance relies

on the fact that the former focuses on identifying the most characteristic to each APT threat indicators and inbound correlations on the very early stages of their existence. The proposed methodology allows the identification of adversaries with precision and sensitivity on a real-time basis. The aforesaid characteristics seem rather effective compared to the already existing works, which considered APTs as short period threats that should be examined as a whole after all the enclosed stages are completed. *APTHunter* implements "kernel" audit logs as a reliable source of information related to system activities that could reveal signs of adversarial activity. Based on this information, the proposed method creates a flowchart of causal relationships, through which the already identified indicators of each APT is used to identify abnormal activity. *APTHunter* was evaluated via "DARPA Transparent Computing" [25] comprising the most important types of cyberthreats, particularly APTs. The methodology was applied over different OS platforms, presenting in all cases consistent and reliable detection rates in the early stages of an APT's appearance.

Park et al. [30] also recaps the subject of the timely identification and detection of APTs during the first stages of their manifestation. According to the authors, the traditional signature-based endpoint detection methods lack the adequate level of efficiency to adapt to the ever-changing environment of APT-oriented attacks. In this context, the authors focus on the introduction of a rule-based methodology that leverages well-known open-source tools to succeed the identification and detection of an APT adversary at the early stages of its appearance, all these under the concept of an EDR-oriented methodology. Specifically, the proposed EDR scheme incorporates Google Rapid Response (GRR) and Auditbeat of Elasticsearch, as two open-source live incident response framework and logging tools, respectively. A dedicated to the dissimulation of an APT attacking environment simulator was also implemented for the analysis of the attack's stages based on the established EDR rulesets. Their EDR scheme was evaluated against the well-known APT29 testbed [14], comprising real-life APT scenarios.

Bajpai et al. [31] readdress the subject of LM endpoint detection through the scrutinization of the behavior of ransomware. To this end, they presented a ransomware detection and response framework that can be leveraged by organizations as a first line of defense against ransomware. Through the examination of numerous ransomware test-cases, the authors mined the most crucial for each malware's functioning characteristics, and incorporated them under a procedural basis to the policy of their proposed framework. Precisely, 25 notorious ransomware variants were examined through both static and dynamic behavioral analysis, including *REvil*, *LockBit*, *IDAFree*, *Cutter*, and *Binary Ninja*. The aforesaid 25 samples were initially mapped through CISA's Decider tool for Time-Triggered Protocol (TTP) calls. That is, the samples were unpacked through static behavioral reverse engineering to reveal the source code of their payload and the pertinent intricacies of the different enclosed classes. On the downside, the finally proposed framework was presented theoretically without revealing any results of its efficiency based on a real-life scenario.

3.2. EDR schemes applied to the IoT ecosystem

Marquez et al. [32] presented a modern detection scheme for pivoting attacks, namely *APIVADS*. The latter analyzes APT pivoting tactics based on the flow of traffic in Small Office Home Office (SOHO) and corporate facilities, where IoT interconnected devices find great applicability. The presented identification scheme aims to overcome the literature gaps that were permitting adversaries to infiltrate the targeted traffic, breach and gain access through parallel obfuscation of their pivoting traces. To succeed that, the authors focused on the flow-based statistical analysis of the generated traffic during pivoting incidents, in an effort to identify anomaly indicators produced by malicious events. *APIVDAS's* core functions can be considered as a crossbreed of decentralized host-based pivoting detection with centralized practices towards the ensemble consideration of the extracted results. The efficiency and scalability of the proposed scheme was evaluated through empirical experimentation, presenting a promising rate of 98.54% successful separation of the examined traffic as normal or malicious, including the identification of TLS, HTTPS, DNS, and P2P events.

Xiao et al. [33] presented *SoK*, an EDR scheme oriented to secure emerging technological assets related to cloud/edge computing and IoT interconnected devices. Above that, the introduced policy converges three of the most contemporary pillars of security practices, namely Zero Trust (ZT) systems, context-based, and risk-based oriented access control. The authors examined the three aforementioned security practices for potential overlapping sectors and collaboration during their applicability. The results of their examination revealed the existence of an extended list of interrelated elements that can be used as incorporated features within a security EDR policy. However, the actual development and presentation of such a policy was left for future work. This also applies to any evaluation based on a real-life attacking scenario.

Sarfaraz at al. [34] focused on the detection of LM APT malicious incidents over interconnected IoT devices via Linux web servers. Towards this goal, the authors amended four out of six layers of the so-called "List of Pain" [35], namely Tactics - Techniques - Procedures (TTPs), tools, network, and host artifacts. Precisely, the aforesaid list forms a relational presentation of the indicators through which malicious activities can be identified. The list is presented in the form of a pyramid-like diagram, each level of which represents a number of delimiters dedicated to the identification of a specific gender of adversaries tactics. Above that, each level denotes the effort that needs to be paid ("pain") by an adversary to overcome the aforementioned delimiters. Towards the quantization of the aforementioned policy of "pain", the authors leveraged the popular Unix-based OS instrumentation framework, called *osquery*. They, implemented the framework on a Linux OS machine to develop a dedicated OS heuristic-based log-query script that seeks for anomalies based on the aforementioned criteria, as those are related to initial connections, escalation of privileges, and above all, LM malicious events.

The work of Weisman et al. [36] focused on the improvement of Security Operation Center's (SOC) performance towards the identification of potentially malevolent incidents in SOHO and corporate networking environments that incorporate IoT interconnected devices through ad-hoc connectivity. Specifically, they revisited the concept of IoT integration in modern attack vectors (namely hardware, APT, LM, etc.), through the classification of the most contemporary adversary tactics that target IoT devices and relevant protocols. The specified characteristics of each attack were thoroughly recorded along with the related in each case-study scenario countermeasures. The work ends with the presentation of an EDR framework targeted to the enhancement of the automation security mechanisms of monitoring and detection.

Ricardo dos Santos et al. [37] addressed the subject of the ever-evolving cybersecurity concerns in Smart Building (SB) facilities. According to the authors, the integration of smart technology with building's management operations, such as automated power distribution and energy saving systems, namely Building Automation Systems (BAS), despite its great potential, enlarges the attack surface, leaving each resident exposed to adversaries. In more detail, the article concentrates on BAS interconnected smart subsystems, including smart surveillance, lightning, and automated heating actuators. Above that, the authors examine the most substantial factors that make building smart devices an impactful attack vector for adversaries. The characteristics of several known and zero-day attacks were examined and analyzed within the paper under the concept of vulnerabilities exploitation of IoT specified operational protocols. The results were aggregated and provided as proof-of-concept with the presentation of an IoT devices-oriented malware. The malware is deliberately tuned to remain inactive for as long as the attacker wishes within the IoT network, and being activated as zero-day to compromise targeted devices.

Süren et al. [38] contributed a four-part vulnerability assessment methodology, namely *PatrIoT*, dedicated to the examination of IoT smart devices. The proposed methodology expands over four pillars of security evaluation: (a) the logical partitioning of the attack surface, (b) the aggregation and leveraging of the most distinct characteristics of the top 100 Common Vulnerabilities Exposures (CVEs) and Common Weakness Enumeration (CWEs) metrics for risk assessment, (c) an overall risk scoring system, and (d) a complete guideline for the execution of penetration testing over IoT devices and related software. The presented methodology underwent a two-year evaluation via the penetration testing of over 30

IoT devices. The authors demonstrated the penetration assessment results for seven of these devices, utilizing top-notch IoT software, such as Xiaomi Mi Home Security Camera, AI robot, and others. The acquired results revealed that the proposed evaluation finds applicability in a great range of interconnected smart devices, with quite promising results for security experts in the field.

Table 2: Summary of the most important aspects of the works included in Section 3. The works are presented in chronological ascending order. The literature titles denoted with the double-star delimiter ** are exclusively related to EDR schemes oriented in the IoT ecosystem.

Summary of the LM EDR-related identified literature			
Title	**Year**	**Testset**	**Summary**
A novel approach to detect malware based on API call sequence analysis [7]	2015	Malicia [39]	Detection of malware API call patterns towards the identification of anomaly behavioral signatures and through the implementation of sequence alignment algorithms and Malicia-Project malware dataset.
Detecting lateral movement through tracking event logs (v1 & 2) [8, 9]	2017	-	Execution of LM through well-known penetration testing tools. Collection of logs with Sysmon and MS Windows Audit Policy. Categorization of logs per attack and proposition of optimal MS Windows infiltration settings for effective identification of LM activity.
Data-driven threat hunting using Sysmon [11]	2018	-	Data-driven threat classification methodology based on aggregated logs created by Sysmon. Introduced a threat analysis system based on the developed by the authors' CTI ontology.
Lateral movement detection using ELK stack [16]	2018	-	A way to generate event-logs with Sysmon, related to the execution of various LM attacks and the associated malicious tools on MS Windows environments. Implementation of ELK stack towards the analysis of possible abnormalities in the collected logs due to existence of LM.
Real-time detection system against malicious tools by monitoring DLL on client computers [15]	2019	-	Proposed a DLL-oriented method for malicious files detection through logs collected by Sysmon [10]. Common DLL list per malicious tool, independent of the underlying MS Windows OS version.
Detecting Mimikatz in lateral movements using Mutex [18]	2020	-	APT detection of Mimikatz, while utilized in LM. Implementation of Mutex memory objects in conjunction with DLL files analysis. Mimikatz's misidentification while deliberately obfuscated.

cont'd on next page

Table 2 – cont'd from previous page

Title	Year	Testbed	Summary
ShadowMove: A stealthy lateral movement strategy [20]	2020	-	A strategy dedicated to the stealth execution of APT and LM techniques towards Windows-based and Linux host compromising over enterprise networks. The presented approach is related to "socket duplication" through which a malicious process silently leverages TCP connections of benign equivalents. The strategy permits APT threat actors to move "stealthily" among hosts in enterprise networks without being identified or detected by host-based, network level, or IDS systems.
On the Forensic Validity of Approximated Audit Logs [24]	2020	DARPA Transparent Computing Engagement 5 [25]	Creation of a renewed pool of metrics towards the quantification of audit log forensic evidence validity and measure of the usefulness of logs under different test-case scenarios. Presentation of the "LogApprox" technique that is deliberately dedicated to the preservation of attack-related forensic evidence through logs. The proposed scheme was assessed by means of the DARPA Transparent Computing Engagement 5 Data Release corpus [25].
Integrating IoT monitoring for security operation center [36] **	2020	-	Presentation of an EDR framework aiming at the enhancement of the automation security mechanisms of monitoring and detection under the concept of SOK operation.
From TTP to IoC: Advanced persistent graphs for threat hunting [13]	2021	APT29 [14]	A threat hunting model dedicated to the evaluation of Sysmon's logs from both an offender's and defender's perspective. Based on indicators of compromise, proactive threat detection can be possibly enhanced. A high rate of FPs due to the absence of rule-based Sysmon's configuration.
Network forensics investigation in virtual data centers using ELK [17]	2021	-	Generation of log records with *Logstash*. Network forensic analysis via the implementation of ElasticSearch and towards the identification of RDP LM-related attacks, ransomware, data exfiltration, etc., as part of a criminal investigation.
USBCulprit: USB-borne air-gap malware [26]	2021	-	Contributed to the wider thematic area of LM techniques through the study of "USBCulprit", which forms an usb-oriented APT exclusively designed to breach corporate or governmental air-gapped networks. Interrelated features of the APT were identified and isolated through extensively persistent reverse engineering techniques. A number of proposed countermeasures including policy, software, and hardware mitigation measures is also given.

cont'd on next page

Table 2 – cont'd from previous page

Title	Year	Testbed	Summary
Leveraging operational technology and the Internet of things to attack smart buildings [37] **	2021	-	Focused on BAS interconnected smart subsystems, including smart surveillance, lightning, and automated heating actuators. An IoT-oriented malware was presented as proof-of-concept that acts as a zero-day, being able to leverage smart devices vulnerabilities.
Threat detection and response in Linux endpoints [19]	2022	-	Custom-built LM and pivoting EDR tool dedicated to Linux OS platforms. The tool expands on the "Osquery" and "Elastic" open-sourced tools and is capable to aggregate endpoint logs at a pre-configured common server. Thereafter, through log-oriented correlation between events from various endpoint devices, the investigation and detection of LM and pivoting incidents is permitted.
Threat-based simulation of data exfiltration towards mitigating multiple ransomware extortions [27]	2022	-	Automated ransomware mitigation concept that acts in parallel as a first line of defense towards the protection of private and corporate networking facilities. It also acts as a measure to deal with the constantly evolving techniques of adversaries to leverage new victims. Implementation of a Python simulation tool of the most impactful ransomware attacks in favor of knowledge and experience acquisition in advance of the actual occurrence of an incident.
Revisiting the detection of LM through Sysmon [21]	2022	LMD-2022 [23]	Introduced a rule-based EDR policy for Sysmon tool dedicated to the identification and detection of LM events within the MS Windows ecosystem. Elaboration on ordinal identified patterns related to LM events through the execution of frequently encountered LM techniques, based on their impact, and as those are presented in MITRE's ATT&CK database [22]. Exploitation of a properly designed testbed with nine LM techniques of diverse targeting subject that revealed a great number of interrelated features that were incorporated as custom rules in Sysmon's config.xml file. Creation of LMD-2022 LM-oriented log-based corpus [23], comprising more than 870K Sysmon event logs. Evaluation of the efficiency of the proposed methodology through the development of an extensible Python .evtx file analyzer, dubbed PeX, which was assessed in terms of TP and FP rates over the LMD-2022 logset.

cont'd on next page

Table 2 – cont'd from previous page

Title	Year	Testbed	Summary
APIVDAS: A Novel privacy-preserving pivot attack detection scheme based on statistical pattern recognition [32] **	2022	-	A detection scheme for pivoting attacks through the implementation of a hybrid methodology that crossbreeds decentralized host-based detection with centralized practices towards the ensemble consideration of the extracted results.
SoK: context and risk aware access control for zero trust systems [33] **	2022	-	Presented *SoK*, an EDR scheme oriented to secure emerging technological assets related to cloud / edge computing and IoT interconnected devices. They examined the three aforementioned security architectures for potential sectors of overlap and collaboration during their applicability.
Real-time heuristic-based detection of attacks performed on a Linux machine using Osquery [34] **	2022	-	Focuses on the detection of LM APT malicious incidents over interconnected IoT devices via Linux web servers. They implemented the framework on a Linux machine to develop a dedicated OS heuristic-based log-query script that seeks for anomalies based on criteria extracted by [35], as those are related to initial connections, escalation of privileges, and above all LM malicious events.
APTHunter: Detecting advanced persistent threats in early stages [29]	2023	DARPA Transparent Computing [25]	Introduced *APTHunter*, a detection system dedicated to the identification of APTs during the initial stages of a system's breach; this is done on a real-time basis and achieved through the implementation of "kernel" audit logs as a reliable source of information related to system activities that could reveal signs of adversarial activity.
Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks [30]	2023	APT29 [14]	Introduction of a rule-based methodology that leverages well-known open-source tools to succeed the identification and detection of an APT adversary at the early stages of its appearance; this is done under the concept of an EDR focused methodology. The EDR scheme incorporates Google Rapid Response (GRR), Auditbeat of Elasticsearch, and an open-source simulator, and it was tested over the APT29 testbed [14].
Know Thy ransomware response: A detailed framework for devising effective ransomware response strategies [31]	2023	-	Ransomware detection and response framework that can be leveraged by organizations as a first line of defense against ransomware attacks. The framework incorporates the most impactful malware's characteristics, as those were derived from the static and dynamic source code examination of 25 real-life ransomware samples. Time-Triggered Protocol (TTP) calls were also mapped through the CISA's Decider tool.

cont'd on next page

Table 2 – cont'd from previous page

Title	Year	Testbed	Summary
PatrIoT: practical and agile threat research for IoT [38] **	2023	-	Presentation of a four-parted vulnerability assessment methodology dedicated to the examination of IoT smart devices. The methodology underwent a two-year evaluation via the penetration testing of over 30 IoT devices, revealing a great applicability on a large range of interconnected smart devices.

Table 3: Identified testbeds incorporated in LM identification and detection EDR schemes. The works highlighted with the star (*) delimiter utilize "APT29" and "LMD-2022" datasets, which comprise real-life examples, as those are presented in Mitre's ATT&CK database [22].

Year	Organization	Dataset	Description
2015 [7]	IMDEA Software Institute	Malicia [39]	Malware-oriented, comprising both malicious and normal traffic as well as API calls.
2020 [24]	U.S. Department of Defense - Defense Advanced Research Projects Agency (DARPA)	DARPA Transparent Computing Engagement 5 [25]	APTs related malicious traffic combined with normal logs.
2021 [13]*, 2023 [31]*	Mordor Intelligence	APT29 [14]	APT threats' evaluation log-based dataset, created based on Mitre's ATT&CK real-life examples.
2022 [21]*	University of the Aegean	LMD-2022 [23]	LM dedicated Sysmon log-based dataset.

4. Machine learning IDS schemes

The current section provides a condensed summary of the key pertinent literature on the subject of LM-specific ML IDS concepts. The concentration is on the methodology of each relevant work regarding the implementation of ML models, either via supervised (SML) or unsupervised (UML) data manipulation techniques. Particularly, we focus on the identification of the core ML algorithm's function in terms of shallow and DNN applicability, the implementation of feature selection processes, and the leverage of publicly available benchmark datasets, if any, towards the evaluation of the effectiveness of each proposed IDS. To facilitate the parsing of the relevant literature, Table 4 recaps the relevant key characteristics of every work included in this section. The various works are chronologically arranged in ascending order. Following the same chronological structure, Table 6 summarizes the public datasets leveraged as proof-of-concept per case-study scenario.

4.1. Supervised learning based schemes

The work of Kaiafas et al. [40] is considered cutting-edge compared to their related counterparts dedicated to the field of log-based identification and anomaly detection. Precisely, the authors focused on the creation of an adaptable and efficient anomaly IDS methodology that leverages the combination of 10 generic log-based features and other eight custom-made, respectively. Both these set of features were extracted from the popular, publicly available Los Alamos

National Laboratory (LANL) log-based corpus that was gathered between 1996 and 2005 [41]. The collected subsets were manipulated via various sampling techniques towards the facilitation of pre-processing and computational power issues with such voluminous data volumes. The authors followed a binary classification scheme exploiting popular SML techniques, namely Random Forest (RF), LogitBoost (LB) and Logistic Regression (LoR). The results were evaluated under False Positive (FPR) and False Negative (FNR) rates, while the malicious classified predictions were re-fed to the ensemble Majority Voting uniform weighted algorithm and re-evaluated.

Bian et al. [42] presented a hybrid anomaly detection methodology, dedicated to the identification and detection of LM-related vulnerabilities. Specifically, their scheme aims to protect targeted networking hosts during the early stages of their exposure to the threat. The authors leveraged the popular open access LANL dataset [41] towards the extraction of 29 custom composite features and other six more flow-based related to authentication and flow event traffic, respectively. Both set of features were leveraged and aggregated into a composite graphical representation of the log authentication traffic. SML techniques, namely Decision Tree (DT), RF, Linear Regression (LiR), Gaussian Naive Bayes (GNB) and Label Binarizer (LaBi), were implemented for the evaluation of the 35 extracted features as part of the proposed anomaly detection scheme. The aforesaid ML techniques were enhanced with under and oversampling techniques to improve their applicability due to the highly imbalanced nature of the LANL logs.

Interestingly enough, during the same year (2019), Bian et al. revisited [42] with the work presented in [43]. This was done under a realistic case study scenario that was focused on RDP-based LM techniques. Keeping the same methodology that was presented in [42], the authors extracted two MS Windows host-based RDP-related event logs subsets from the LANL testbed [41], namely, *Comprehensive* and *Unified*, respectively. The two subsets were evaluated under supervised classification algorithms. The concept of identification and detection of LM malicious incidents in authentication logs via SML techniques was also revisited by Bian et al. in [44]. Moreover, the two above-mentioned RDP-based subsets were also leveraged under the same scheme presented in [42, 43]. The model's classification efficiency was evaluated against the two aforesaid LANL-originated subsets in terms of LM detection and overhead. Additionally, the attack frequency of the evaluated LM patterns was tampered by importing artificially generated noise and traffic variations towards the model's evaluation against adversarial LM scenarios.

Chen et al. [45] introduced a log-based anomaly identification and detection IDS scheme, dedicated to the examination of Sysmon event logs via supervised shallow classification (SSC) and Deep Neural Networks (DNN) supervised techniques. In more detail, three popular algorithms were implemented upon experimentally collected Sysmon logs, namely Support Vector Machines (SVM), Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN), for each of the two experimental categories, respectively. On top of that, a generic set of custom features were presented based on the transformation of Sysmon *EventIDs* and evaluated in terms of TP and TN rates.

Narouei et al. [46] presented *DLLMiner*, that is, a heuristic DLL-oriented malware detection methodology, which is also applicable to the general area of the identification of APT and LM activity. This scheme was constructed upon the results of the static analysis of portable executable's dynamic-link library features. As evident from the LM EDR policy presented in [21], *DLLMiner* incorporates significant characteristics regarding potentially malevolent behavior, without even executing the files.

Juwono et al. [47] examined the effectiveness of various ML models under the concept of intrusion detection over voluminous log files related to real-life malware infections. The authors created two sandbox environments, namely *Cuckoo* and *Anubis*, within which the well-known *Weka* ML framework was exploited for assessing four popular shallow classification algorithms, namely Support Vector Machine (SVM), DT, Nearest Neighbour (NNeighbour), and RF, over the legacy Malheur dataset [48].

Smiliotopoulos et al. [49], contributed a detailed methodology specifically dedicated to the identification of LM via

the implementation of SML algorithmic models. They detailed the significant importance that human-driven feature selection may impose to the effectiveness of ML models, especially when those are combined with elevated pre-processing and feature importance processes. Several SML techniques were evaluated over an exclusively created for that purpose LM log-based dataset, namely *LMD-2023* [23]. The latter is unbalanced, comprising more than 1.8M (both normal and infected traffic) log samples collected through multiple, virtual and physical, MS Windows terminals by means of Sysmon SIEM. A variety of 10 base shallow estimators, one ensemble meta-estimator, and five DNN models were evaluated upon the multiclass classification of *LMD-2023's* traffic, yielding an F1-score of 99.41%. The same work additionally contributed an open-source tool called *ETCExp* for converting *EVTX* monitor log files to a *CSV* equivalent format.

He et al. [50] dealt with the presentation of a multidimensional detection methodology, specifically designed to identify the LM behavioral stages of APT malware threats targeting the MS Windows Server Message Block (SMB) sharing protocol. The authors extracted and analyzed the most impactful "honeytoken" nodes related to the LM techniques executed after a malware's initial access to a targeted system in terms of an SMB-dedicated honeypot. Further, feature generation and engineering techniques were applied towards the preparation of a dataset to be implemented as input to the core NN supervised classification scheme. Specifically, the NN model enclosed a multi-layered combination of hidden and convolution NN, namely TextCNN, LSTM, and FastText layers. The models' evaluation was conducted on a manually collected benchmark dataset, comprised by more than 10K publicly available malicious malware APT files.

4.2. Unsupervised learning based schemes

So far, only a limited number of works considered unsupervised ML towards the classification of a vast diversity of collected logs exclusively related to LM. In general, the incorporated features were either included as generic to the originally analyzed log-based testbeds or aggregated from numerous interrelated nodes and edges included in the network topology in a custom-made manner.

Bohara et al. [51] examined the composition of an anomaly detection scheme that performs on top of ensemble UML techniques towards the identification of LM event traces on infected hosts. The LANL testbed [41] was amended though the graphical representation of the various communication nodes of the targeted hosts and via a graph-based model. This was done to conclude to the extraction of the related to the classification experiments features. The authors continued with the implementation of ensemble of UML models, namely Principal Component Analysis (PCA), k-means clustering, and Median Absolute Deviation-based outlier (MADO). The efficiency of the proposed scheme was evaluated under a trace-related simulation case study.

Le et al. [52] manipulated four UML methods, namely Autoencoder (AE), Isolation Forest (IF), Lightweight On-line detection of anomalies (LODA), and Local Outlier Factor (LOF), for creating an IDS scheme that targets the identification of LM insider attacks. The data were manipulated via pre-processing techniques to fit with Deep Learning (DL) models and contribute to reveal anomaly behavior. Various UML ensembles were created and evaluated against several state-of-the-art works exploiting popular benchmark testbeds, namely CERT [53], LANL [41], and TWOS [54].

Inspired by the hybrid approaches presented in [51] and [52], Chen et al. [55] aggregated the existed theories of a network's nodes and vectors graphical representation mapping via network embedding with feature manipulation and pre-processing techniques towards the creation of composite features. The finally selected features were evaluated under a semi-supervised classification model via Denoising autoencoder algorithm. To this direction, the authors leveraged a balanced subset of the LANL dataset [41], namely *"The Comprehensive, Multi-Source, Cyber-Security Event"*. The experimental results were evaluated via FPR, TPR, accuracy, and precision metrics.

4.3. ML schemes applied on the IoT ecosystem

Noor et al. [56] introduced an IDS framework, that leverages shallow and DL techniques combined with semantic networking representation practices to build a cyber-threat identification scheme, applicable to APT targeted devices including IoT ones. Their concept relies on the contemporary notion of sharing Cyber Threat Incident Reports (CTIR) towards en effective proactive countermeasure against the ever-evolving adversary tactics. Precisely, the proposed framework leverages the popular Mitre's ATT&CK taxonomy to search an entire network for the existence of malicious TTPs, which are then semantically interrelated through a TTP-Detection diagram. The recognized attack patterns are imported as features and evaluated though various ML models, namely DT, RF, DL, SVM, and Bayesian Belief Networks (BBNs). The achieved accuracy of the aggregated models reaches an average of 92%, presenting a low FPR at the same time.

We observed that the relevant literature contains several works similar to [56]. All of them suggested a shallow or DNN classification scheme, originally proposed by others, with the goal to make them more efficient in terms of the identification and detection of malevolent APT or LM incidents. For instance, the work of Powel [57] presented an unsupervised learning model of LM detection, based on the role-based approach of clustering the system connections to remote hosts into distinct roles. That is, the successful identification of unusual process sequences or generic connections to remote hosts may reveal the existence of an adversary. Moreover, Imran et al. [58] evaluated the performance of multiple SSC and DNN models, namely RF, SVM, AdaBoost (AD), Stochastic-Gradient Descent (SGDC), Gradient-Boosting (GBC), MLP and LSTM, to discerning APT and LM activity. For balancing their dataset, the authors relied on the Synthetic Minority Oversampling (SMOTE) balancing technique. A similar approach is presented by González-Manzano [59], in which numerous APT-related pieces of malware were analyzed towards feature extraction and evaluated under a SSC scheme towards effectively distinguishing from regular malware. Although not specifically defined, the aforementioned works are generally evaluated upon logs extracted from IoT or IIoT devices. However, for reasons of completeness, we opt to include such schemes in Table 4.

Arifeen et al. [60] contributed an automated micro-segmentation ML model destined to IIoT. Specifically, the model is dedicated to the identification and detection of LM events derived from a malevolent actor or malware over IIoT devices. This is achieved through the generation of micro-segments via the fragmentation of the various network blocks of traffic into normal or malicious. The performance of the presented model was evaluated over two IoT-oriented testbeds, namely UNSW-ND15 [61] and IoTD20 [62], revealing a promising effectiveness to curb malicious actions due to LM or malware.

Koroniotis et al. [63] contributed a DNN forensic framework, dubbed as *Intelligent Satellite Deep Learning Network Forensic (INSAT-DLNF)*. Their scheme aims at the timely and efficient identification and detection of malevolent LM incidents targeting satellite smart networks. This is achieved through the consecutive training of a hybrid Neural Network (NN) consisted from a Long Short-term Memory Recurrent Neural Network (LSTM-RNN) and a Gated Recurrent Unit (GRU) models. The efficiency of the presented network was evaluated through the implementation of three popular IoT-oriented benchmark testbeds, i.e., NSL-KDD [64], UNSW-ND15 [61], and Bot-IoT [65]. The derived evaluation results were compared with three SML models (ANN, NB, Association Rule Mining (ARM)) and two unsupervised algorithms (k-Means, Expectation-maximization).

The focus of Altunay et al. [66] was on the creation of an advanced IDS scheme devoted to the protection and security of IIoT interconnected devices. Specifically, the authors assessed three distinct IDS DNN models, namely CNN, LSTM, and the hybrid combination of both CNN+LSTM. To evaluate the results of each model, they relied on two well-known testbeds, namely UNSW-NB15 [61] and X-IIoTID [67], defining successfully normal and malicious records with a rate of 93.21% and 92.9% regarding binary and multiclass classification, respectively. The effectiveness of the conducted experiments via each of the three aforementioned models were compared both with each other and with other relevant studies in the field.

A similar approach with [66] was presented by Sarhan et al. [68]. The authors centered on the design of an IDS scheme over heterogeneous network data samples that targets the identification and detection of LM and APT. In detail, they aggregated cyberthreat intelligence practices from various independent organizations to build a collaborative federated learning IDS scheme, namely Threat Intelligence Sharing Scheme (TISC), towards the design and effective training of DNN ML models. Their proposal was evaluated over two popular datasets, namely UNSW-NB15 [61] and Bot-IoT [65], via three different models, namely federated, centralized, and localized. It should be noted that the federated model outperformed the localized in terms of F1-score and Accuracy, however still remained behind the centralized one.

Jayalaxmi et al. [69] also examined the heterogeneous nature of IIoT interconnected device's traffic, as a key cause that affects the effectiveness of IDS schemes against zero-days. That is, they proposed a comprehensive DNN zero-day IDS framework, called *PINGUS*, which combines existing feature mapping techniques with cascading models. Optimal features selection was conducted through the Denoising Autoencoder (AE) algorithm, while the classification and attack detection itself was accomplished via Cascade Forward Back Propagation Neural Network (CFBPNN). The proposed scheme was evaluated through five open access datasets, namely the Natural Gas Pipeline [70], WA Water Tank [71], NSL-KDD [64], UNSW-NB15 [61], and Bot-IoT [65]. The derived results were compared to the related literature on the subject, exceeding by 25% on average similar models.

Table 4: Summary of the key aspects of the works included in Section 4. The works are presented in chronological ascending order. The paper's titles marked with the ^{*1} or ^{*2} delimiters represent SML or UML techniques, respectively. The literature titles denoted with the double-star delimiter ** are exclusively related to ML schemes destined to the IoT or IIoT ecosystem.

Summary of the LM ML-related identified literature

Title	Year	Dataset	Method	Summary
DLLMiner: structural mining for malware detection [46] ^{*1}	2015	Malicia [39]	SSC (RF, NB [WEKA])	Heuristic DLL-oriented malware detection methodology, which is also applicable to the general area of the identification of APT and LM.
A comparative study of behavior analysis sandboxes in malware detection [47] ^{*1}	2015	Malheur [48]	SSC (SVM, DT, RF, NNeighbour [WEKA])	ML IDS scheme integrated with *Cuckoo* and *Anubis* sandbox environments, towards the identification of anomalies.
An unsupervised multi-detector approach for identifying malicious lateral movement [51] ^{*2}	2017	LANL [41]	UML (PCA, k-means, MADO, Ensemble ML)	Automated unsupervised anomaly detection ensemble method for identifying LM traces on infected hosts. The LANL dataset was used for the log-based graphical representation and the extraction of features. The extracted data were evaluated via two independent anomaly detection methods, which incorporate PCA, k-means, and MADO techniques. The results of these two methods were combined and re-evaluated under a parameter-based ensemble learning method.

Cont'd on next page

Table 4 – Cont'd from previous page

Title	Year	Dataset	Method	Summary
A novel approach for identifying lateral movement attacks based on network embedding [55] *2	2018	LANL [41]	UML (Network Embedding, Denoising Autoencoders)	Semi-supervised classification through Denoising autoencoder of features stemming from network embedding and feature aggregation techniques.
Detecting malicious authentication events trustfully [40] *1	2018	LANL [41]	SSC (RF, LB, LoR, MV)	A log-based anomaly detection approach applied on generic and custom-made/engineered (artificial/synthetic) log features, which were extracted from the LANL dataset. The application of sampling techniques precedes the samples' classification with SSC ML techniques in normal or malicious, while the final results were evaluated under the FPR and FNR metrics.
Host in Danger? Detecting network intrusions from authentication logs [42] *1	2019	-	SSC (DT, RF, LiR, GNB, LaBi)	Hybrid anomaly detection perspective regarding the identification of LM techniques on hosts during the early stages of their threat exposure. They extracted 35 composite log-based features from the LANL dataset, which were classified under SSC ML algorithms.
A machine learning approach for RDP-based lateral movement detection [43] *1	2019	LANL [41]	SSC (DT, RF, LiR, GNB, LaBi)	Re-examination of the work presented in [42] under the prism of RDP-based LM.
A machine learning framework for investigating data breaches based on semantic analysis of adversary's attack patterns in threat intelligence repositories [56] **	2019	-	SSC (DT, RF, SVM, BBN)	An IDS framework that leverages both shallow and deep learning techniques combined with semantic networking representation practices to build a cyberthreat identification scheme applicable to APT targeted devices including IoT ones.
Analyzing system log based on machine learning model [45] *1	2020	-	SSC (SVM) - NN	Shallow and DNN ML classification analysis on Sysmon log samples.
Uncovering lateral movement using authentication logs [44] *1	2021	LANL [41]	SSC (DT, RF, LiR, GNB, LaBi)	The same authors of [42] and [43] revisit the subject of LM detection through the classification of LM-related authentication logs with shallow ML techniques.
Anomaly detection for insider threats using unsupervised ensembles [52] *2	2021	CERT [53], LANL [41], TWOS [54]	UML (AE, IF, LODA, LOF, unsupervised ensembles)	Unsupervised ML detection of anomalies on user behavioral habits. Creation of unsupervised ML ensembles to evaluate the performance of the proposed anomaly detection scheme under various algorithmic combinations.

Cont'd on next page

Table 4 – Cont'd from previous page

Title	Year	Dataset	Method	Summary
Automated microsegmentation for lateral movement prevention in industrial internet of things (IIoT) [60] **	2021	UNSW-ND15 [61], IoTD20 [62]	SSC (DT)	Automated micro-segmentation ML model destined to IoT devices integrated on industrial operational networks. Generation of micro-segments via the fragmentation of the various network blocks of traffic into normal or malicious.
A new intelligent satellite deep learning network forensic framework for smart satellite networks [63] **	2022	NSL-KDD, UNSW-ND15 [61], BoT-IoT	DNN (LSTM, RNN, GRU), SSC (ANN, NB, Association Rule Mining (ARM), BBN), UML (k-Means, Expectation-maximization)	A DNN forensic framework centered on the detection of malevolent cyberattack incidents targeting satellite smart networks. This is done through the consecutive training of a hybrid NN consisted from LSTM-RNN and GRU models.
Role-based lateral movement detection with unsupervised learning [57] **	2022	-	UML (SVC)	UML model based on the role-based approach of clustering the system connections to remote hosts into distinct roles.
On the detection of lateral movement through supervised machine learning and an open-source tool to create turnkey datasets from Sysmon logs [49] *1 *2	2023	LMD-2023 [23]	SSC, DNN (MLP, CNN, LSTM, RNN, AE)	An SML-based methodology to detect LM. Human-driven feature selection combined with elevated pre-processing and feature importance processes. Unbalanced dataset, comprised from more than 1.8M (both normal and infected traffic) Sysmon log samples. Contributed the publicly available ETCExp tool to easily transform *EVTX* monitor log files to a CSV equivalent format.
A performance overview of machine learning-based defense strategies for advanced persistent threats in industrial control systems [58] **	2023	-	SSC (RF, SVM, AD, SGDC, GBC), DDN (MLP, LSTM)	Performance evaluation of multiple SSC and DNN models to identify APT and LM infections, via the SMOTE balancing technique upon log-based datasets.
A technical characterization of APTs by leveraging public resources [59] **	2023	-	SSC (RF, KNN), DDN (MLP)	Analysis of numerous APT-related pieces of malware towards feature extraction and evaluation under a SSC scheme. The goal is to tell apart legacy from APT-related malware.

Cont'd on next page

Table 4 – Cont'd from previous page

Title	Year	Dataset	Method	Summary
A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks [66] **	2023	UNSW-ND15 [61], X-IIoTD [67]	DNN (CNN, LSTM, CNN+LSTM)	An IDS scheme dedicated to the protection and security of IIoT interconnected devices. Presentation of three distinct IDS DNN models, namely CNN, LSTM, and a hybrid combination of both CNN+LSTM.
Cyber threat intelligence sharing scheme based on federated learning for network intrusion detection [68] **	2023	UNSW-NB15 [61], Bot-IoT [65]	DNN, LSTM, Federated Learning	A collaborative federated learning IDS scheme, aiming at the design and effective training of DNN ML models. The presented framework is applicable to the identification and early detection of LM and APT.
PIGNUS: A deep learning model for IDS in industrial internet-of-things [69] **	2023	NSL-KDD [64], UNSW-NB15 [61] and Bot-IoT [65]	DNN (AE, CFBPNN)	A comprehensive DNN zero-day focused IDS framework, which combines existing feature mapping techniques with cascading models.
A comprehensive detection method for the lateral movement stage of APT attacks [50]	2023	-	DNN (TextCNN, LSTM, FastText NN)	Supervised DNN LM multidimensional framework targeting the identification of LM behavioral stages of APT malware related to the MS Windows SMB protocol.

5. Graph-based schemes

Although the literature is full of works addressing graph-based IDS schemes, only a limited number of them are dedicated to the identification and detection of LM. Table 5 summarizes the key elements of the most prominent published works, while Tables 6 and **??** recap the publicly available datasets leveraged as proof-of-concept per case-study scenario.

The work of Purvine et al. [72], described a LM detection methodology dubbed "dynamic graph-based reachability model (DGBR)". The scheme builds upon the definition of an impact-oriented metric on graph-based techniques. In more detail, the authors developed a custom algorithm towards defining the evolution of the multiple paths that an adversary could follow while moving laterally around the various network nodes following the exploitation of a critical vulnerability. This model is the core of a network-level impact score, which is quantified based on the value and reachability score assigned to each network node that could be compromised by adversaries. The effectiveness of the presented DGBR model was tested along with a case study scenario related to "Pass-the-Hash" MS Windows credentials vulnerability over the LANL dataset [41]. It should be noted that the impact metric model was implemented via a C++ source code script that was, however, not made publicly available.

A similar approach was presented by Liu et al. [73] that describes a graph-based IDS scheme, called *Latte*, which subserves two purposes. First, it handles the multi-layered nature of voluminous data samples stemming from LM incidents, and secondly, it addresses the lack of knowledge regarding the tactics that an adversary might use. Moreover, *Latte* contributes in a twofold way to the identification and detection of LM attacks. That is, initially, the presented methodology identifies and marks host and user accounts and their various interconnections as nodes and edges, respectively.

Once an infected node is identified, it leads through the proposed forensic algorithmic analysis to any other compromised element(s). The identification and detection process continues with an algorithmic approach that leverages a remote file execution detector towards recognizing suspicious path anomalies caused by unknown LM attempts.

The work of Liu et al. [73] is regarded a milestone for researchers in the domain of LM identification, inspiring two more similar approaches on the same subject. Precisely, Ho et al. [74] presented *Hopper*, an LM identification tool that is fed by real-life generated log-based traffic. Hopper, tracks user's login activities and outlines their correlations among hosts on a graph-based algorithmic representation. The effectiveness of *Hopper* was tested upon a 15-month custom dataset that was specifically injected with LM events. The results revealed that Hopper contributes to the detection of anomalies among multiple logins, related to LM attacks.

Also inspired by the work in [73], Fang et al. [75] examined the literature gaps regarding the efficiency of the existing IDS LM identification models. In this endeavor, they presented *LMTracker*, a LM detection scheme that leverages two custom graph-based algorithmic models, supported by advanced graph neural networks theory. The two algorithms are dedicated to the graphical representation of the LM-related paths and the unsupervised anomaly path detection based on a predefined threshold, respectively. The robustness of the presented NN graph model is supported via the implementation of features of various elements included in the captured log-based traffic, including users, computers, processes, etc. The features were pre-processed as nodes for the construction of heterogeneous graphs that depict the various relationships among them. The *LMTracker's* effectiveness was evaluated upon two popular benchmark testbeds, namely LANL [41] and CERT 6.2 [53].

Chen et al. [76] engaged with the fundamentals of how a "Blue" defensive team may design, implement, and execute ML algorithmic models towards the perpetual hunting of APTs. The presented methodology is a step-by-step tutorial of how security models should be built so that security teams avoid the challenges of threat's low signature footprint, the imbalanced nature of the tested datasets, and the lack of knowledge when a zero-day emerges. These guidelines were presented through two case study models, namely *Fuchikoma* and *APTEmu*. The former comprises an example of how autonomous threat hunting via Natural Language Process (NLP) NN and graph-based algorithms could be accomplished, while the latter is an emulator for APT3, as those are presented in Mitre's ATT&CK vulnerability list [22]. The presented models were evaluated and discussed over APT3 malware dataset, that was originally created by Haddadpajouh et al. [77].

5.1. Graph-based algorithmic schemes applied on the IoT ecosystem

Agmon et al. [78] dealt with the vectors that make IoT interrelated devices prone to LM techniques. These include the low-end nature of the device, the great diversity of vendors, the low security standards on which the device's firmware is developed, its location within the network, and its communication capabilities and supported protocols. In more detail, they presented a graph-based network-level LM risk quantification methodology, which can be evaluated via an incrementally constructed attack graph model comprising nodes and vectors derived from aspects related to the location and key communication features of the IoT device. Precisely, they proposed a depth-first branch and bound (DFBnB) heuristic search algorithm, aiming at the optimization of the structure of interconnected IoT devices through the mitigation of risk regarding full deployment and maximum utility.

Yang et al. [79] contributed a hybrid IoT-IDS model targeting at the protection of the numerous data transferred frequently via IoT interconnected devices. Emerged on top of the inabilities of the already presented solutions in terms of robustness and real-time detection, the authors' work concentrated on the combination of data fusion practices on feature semantics level with the CNN bidirectional LSTM (BiLSTM) DNN algorithm towards the effective identification of anomalies within device's requests. In particular, the IoT extracted traffic is statistically analyzed to produce informational-enriched features related to semantic relationships. Moreover, multi-view feature fusion and alignment practices transform the extracted features into word vectors that in turn are injected into the CNN-BiLSTM DNN model

to be classified. The effectiveness of the presented model was evaluated via the NSL-KDD dataset [64] over 43 extracted features derived from four distinct attack categories, namely *Intrinsic*, *Content*, *Time-based*, and *Host-based*.

An alternative forensic investigation and traceability approach destined to LM and APT was presented by Wang et al. [80]. That is, the authors presented a graph-based reconstruction methodology of APT malicious incidents in large networks, including IoT and mobile ones. They composed an APT alert correlation model targeting the elimination of FP indicators as those are generated by known and unknown APT scenarios identified by open-source or proprietary IDS tools. Precisely, for the needs of the alert generation process, the authors incorporated the *Zeek* IDS and firewall, injected with various customized configuration scripts such as those available by Mitre ATT&CK framework [81]. The presented model mined the aforementioned voluminous log-set history and leveraged the Monte Carlo Tree Search (MCTS) heuristic algorithm to reveal the existence of advanced APT existence via incident's reconstruction. The efficiency of the model was evaluated over the CSE-CIC-IDS2018 dataset [82], revealing promising results related to the elimination of FP alerts.

A similar approach to [80] was presented by Javed et al. [83], as part of their behavioral analysis effort to identify complex hidden APT scenarios over IIoT interconnected devices. According to the authors, while traditional ML methods presented significantly promising results, it struggled to recognize in a real-time manner APT adversaries efforts to exploit cyber physical IIoT systems. The proposed methodology leverages the robustness of novel graph-based schemes, more specifically, the Graph Attention Neural Networks (GAN NN), that comprises a feature analysis multidimensional algorithm. Put simply, the GAN algorithm is combined with Convolutional NN (CNN) in an effort to improve the early detection of APT. The final model was evaluated against two publicly available datasets, DAPT2020 [84] and Edge-IIoT [85].

Sharadqh et al. [86] presented, *HybridChain-IDS*, a promising IDS model that aggregates the powerful features of the bi-level optimization theory, the data privacy and security of Blockchain technology, graph-based attack's features reconstruction, and the classification robustness of adversary paths via the Enhanced KNN (eKNN) algorithm, under a unified hybrid IDS concept. The presented model was assessed in terms of its capacity to identify three contemporary attacks, namely brute force, SYN flood, and phishing, over an IoT network comprised of 5% of malevolent nodes. Although promising, the results still need to be appraised against voluminous corporate sets of data to prove the model's robustness.

Kumar et al. [87] also focused on the presentation of an IIoT IDS framework, named *RAPTOR*. The latter is focused on the early stages identification and detection of APT adversarial campaigns, including LM, over IIoT corporate networks. RAPTOR correlates data from multiple open-source origins for creating a multi-level APT campaign graph. This graph is followed by feature's vector selection and pre-processing processes for concluding to the final classification of the analyzed traffic via SML schemes. Specifically, for the classification task, RAPTOR exploits two ML models, namely SVM and RF.

Table 5: Summary of the key aspects of the works included in subsection 5. The works are presented in chronological ascending order. The literature titles denoted with the double-star delimiter ** are exclusively related to GB schemes destined to the IoT or IIoT ecosystem.

Summary of the LM GB-related identified literature				
Title	**Year**	**Dataset**	**Method**	**Summary**
A Graph-Based impact metric for mitigating lateral movement cyber attacks [72]	2016	LANL [41]	DGBR model	Introduction of a DGBR model that keeps track of the various adversarial paths that may be followed during the exploitation of vulnerabilities with LM techniques. Through the DGBR model, the authors calculate a network-level impact score and conduct a PtH case study on the LANL dataset.
Latte: Large-scale lateral movement Detection [73]	2018	-	GB model	Identification of an infected host as an anchor point to reveal other compromised hosts through forensic graph-based algorithms. Their scheme reveals anomalies on rare paths through the detection of remote file execution.
Deployment optimization of IoT devices through attack graph Analysis [78] **	2019	-	DFBnB GB model	A graph-based, network-level LM risk quantification methodology. Through an augmented attack graph model, the proposed methodology can benchmark the location and communication characteristics of IoT devices.
Hopper: Modeling and detecting lateral movement [74]	2021	-	GB model	A graph representation system for LM attack detection based on real-life generated logs. Login activity is tracked and outlined through a graph of interrelated logins among the implicated hosts that is dedicated to track and outline login activity towards the detection of anomalies among logins referring to LM.
LMTracker: Lateral movement path detection based on heterogeneous graph embedding [75]	2022	LANL [41], CERT [53]	GB model - NN	Presentation of the *LMTracker* custom LM identification algorithm. *LMTracker* is a mixture of LM paths representation via heterogeneous graphs construction and anomaly detection through graph-based NN algorithmic theory.
Building machine learning-based threat hunting system from scratch [76]	2022	APT3 [77]	NLP NN - GB models	A step-by-step methodology, in kind of tutorial, regarding how security models should be built so that security teams avoid the challenges of APT threat's referring to low signature footprint, the imbalanced nature of the tested datasets, and the lack of knowledge when a zero-day is exploited. Two case study scenarios, namely *Fuchikoma* and *APTEmu*.

Cont'd

Table 5 – Cont'd from previous page

Title	Year	Dataset	Method	Summary
An enhanced intrusion detection system for IoT networks based on deep learning and knowledge graph [79] **	2022	NSL-KDD [64]	Hybrid (GB model - CNN-BiLSTM)	A hybrid IoT-IDS model combining data fusion practices on feature semantics level and the CNN bidirectional LSTM (BiLSTM) DNN algorithm towards the effective identification of anomalies within requests produced by network devices.
An end-to-end method for advanced persistent threats reconstruction in large-scale networks based on alert and log correlation [80] **	2022	CSE-CIC-IDS2018 [82]	GB model - Monte Carlo Tree Search (MCTS) heuristic algorithm	A graph-based reconstruction methodology of APT malicious incidents in large-scale networks, including IoT and mobile ones. The APT alerts' correlation model targets the elimination of FP indicators, as those are generated by known and unknown APT scenarios. The latter are identified by open source or proprietary IDS tools and via the implementation of the MCTS heuristic algorithm.
APT adversarial defense mechanism for industrial IoT enabled cyber-physical system [83] **	2023	DAPT2020 [84], Edge-IIoT [85]	GB (GAN) model - CNN	A GAN NN APT behavioral analysis graph-based model comprising a feature analysis multidimensional algorithm. GAN algorithm is combined with Convolutional NN (CNN) towards an improved detection model.
Hybrid Chain: Blockchain enabled framework for bi-level intrusion detection and graph-based mitigation for security provisioning in edge assisted IoT environment [86] **	2023	-	GB model, eKNN, Blockchain authentication	A unified hybrid IDS model that aggregates the powerful features of the bi-level optimization theory, the data privacy and security of Blockchain technology, graph-based attack's features reconstruction, and the classification robustness of adversary paths via the eKNN algorithm.
RAPTOR: Advanced persistent threat detection in industrial IoT via attack stage correlation [87] **	2023	-	GB model, SML (SVM, RF)	An IIoT graph-based and SML IDS framework dedicated to the identification and detection of APT malevolent campaigns, including LM techniques, on the early stages of their existence.

6. Analysis and Discussion

With reference to Sections 3- 5, Tables 2- 6, and Figure 1 it becomes apparent that the literature works regarding EDR policies, ML, and GB IDS frameworks focusing on the presentation of IDS solutions towards the identification, detection and elimination of LM malevolent incidents follows an increasing trend from 2015 onwards. Naturally, this reflects the ever-going increase in the complexity and frequency of cyberattacks, that is, APT, LM, pivoting, malware, etc., along with the offensive nature of their existence. Amongst others, since 2020, the drift brings progressively the experimentation

Table 6: Identified testbeds incorporated in LM intrusion detection ML and Graph-based schemes.

Year	Organization	Dataset	Description
2015 [46]	IMDEA Software Institute	Malicia [39]	Malware-oriented. It includes both malicious and normal traffic as well as the respective/matching API calls.
2015 [47]	University of Mannheim, Germany	Malheur [48]	Legacy dataset that contains the recorded behavior of malware. It can be used for classifying and clustering malware behavior.
2016 [72],2018 [40, 55],2019 [42, 43],2021 [44, 52],2022 [75]	U.S. Los Alamos National Lab (LANL)	LANL [41]	Publicly available log-based set of data comprising numerous collected normal and malicious MS Windows event viewer logs.
2021 [52], 2022 [75]	Carnegie Mellon University - Community Emergency Response Team (CERT)	CERT [53]	Synthetic insider threat testbed that includes both background and malicious actor's synthetic data.
2021 [52]	Harilal et al.	TWOS [54]	Publicly available dataset including both normal and malicious user interactions with each other. The testbed was created during a double-role simulation gamified competition specifically conducted to obtain normal and adversaries instances of insider threats.
2021 [60], 2022 [63], 2023 [66, 68, 69]	University of South Wales Sydney	UNSW-ND15 [61]	Collection of 100 GB raw source files in pcap, BRO, Argus, and CSV format, including the relevant reports per file. The testbed was created by the IXIA PerfectStorm tool at UNSW, aimed to provide a stable source of normal and malicious log files for security experts and researchers.
2021 [60]	Ullaf et al.	IoTD20 [62]	IoT botnet testbed comprising recorded raw normal and malicious traffic over IoT devices. It is provided as a reference point for the creation and benchmarking of IoT IDS related schemes.
2022 [76]	Haddadpajouh et al. - Cyber Security Lab	APT3 [77]	It contains more than 12K samples of APT malware threats, as those were derived from five well-known APT groups, namely APT1, APT3, APT28, APT33, and APT37.
2022 [80]	University of New Brunswick - [Communications Security Establishment (CSE) & the Canadian Institute for Cybersecurity (CIC)]	CSE-CIC-IDS2018 [82]	Multipurpose IDS benchmarking dataset generated through the collection of numerous events and behaviors representations captured during the user profile's creation processes. It contains seven attack scenarios, namely Brute-force, Heartbleed, Botnet, DoS, DDoS, Web attacks, and insider's network infiltration. The testbed upon which the dataset was created implemented 50 red-team machines against 420 corporate targeted terminals and 30 servers. The network and log traffic is arranged in 80 extracted features.
2022 [63, 79], 2023 [69]	University of New Brunswick - Canadian Institute of Cyberdefense	NSL-KDD [64]	Popular benchmark dataset in the field of CIS and IoT security. The testbed was produced as an upgraded version of KDD'99, however, it suffers from the lack of public related events.
2022 [63], 2023 [68, 69]	University of South Wales Sydney	Bot-IoT [65]	Benchmarking testbed comprising a combination of normal and botnet traffic logs. The captured files in pcap format are 69.3 GB in size, with more than 72M records. The extracted flow traffic in CSV format is 16.7 GB. The collected traffic incorporates logs from DDoS, DoS, OS and services scan, keylogging, and data exfiltration attacks.
2023 [66]	Muna Al-Hawawreh et al.	X-IIoTD [67]	The X-IIoTID dataset represents a voluminous collection of adversary's tactics, techniques and procedures, and various activities related to IIoT devices. The list of the included malicious traffic events is rather extended and includes generic scanning, vulnerability scanning, webSocket fuzzing, discovering Constrained Application Protocol (CoAP) resources, brute force attacks, reverse shell, man-in-the-middle, Message Queuing Telemetry Transport (MQTT) protocol cloud broker-subscription, data exfiltration, ransom distributed DoS, etc.
2023 [83]	Myneni & Chowdhary et al.	DAPT2020 [84]	Benchmark dataset which incorporates normal and APT-oriented malicious traffic. The dataset was created via the recording of five days of continuous simulated network traffic; according to its creators, this is equivalent to three months of real-life traffic.
2023 [83]	Ferrag et al.	Edge-IIoT [85]	Contemporary cybersecurity testing and benchmarking dataset, which encloses IoT and IIoT application traffic derived from 10 IoT physical devices. It is designed to serve ML experimental concepts, both centralized and federated. It incorporates seven distinguished layers, namely cloud computing, network functions virtualization, Blockchain network, fog computing, software-defined networking, edge computing, and IoT and IIoT perception. Nearly 1.2K features were identified in total, however 61 of them were finally proposed; these were derived from diverse sources, including alerts, system resources, logs, and network traffic.

with benchmark corpora derived from IoT and IIoT interrelated devices, as those pose a progressively alluring attack vector for future threat actors.

A brief explanation for this ascendant trajectory stems from the extensive observation of how the works presenting LM-oriented detection policies [21] may be implemented as base knowledge to others dedicated to the presentation of LM IDS schemes [49]. As a characteristic example, the dedicated to LM events identification through Sysmon policies presented as Appendix "A" in [21] was implemented as the basic criteria criterion for labelling the LMD-2023 dataset, and also as fundamental during the data manipulation processes and setting of the SML models (Shallow and DNN) in [49]. Nevertheless, as already pointed out, the effectiveness of LM ML or GB IDS schemes is interrelated to the existence of robust log-based EDR procedures that will guarantee the in-time identification and alerting during potentially malevolent events in terms of EDR teams. The identified elements will be forwarded in turn as features to "Blue teams" for the second stage of the ML or GB algorithmic analysis. Especially for ML techniques, three different pools of available data manipulation techniques, namely supervised, unsupervised, and semi-supervised, together with the numerous algorithmic models available enlarge dramatically the response vector's possibilities during the construction of IDS schemes, either shallow, DNN, graph-based, or some combination of them. To this end, the availability as public of IDS-destined security testbeds is key to identifying the robustness and efficiency of the aforesaid detection schemes, in terms of a sandbox environment, without jeopardizing the continuity of, say, critical infrastructures during a real-time evaluation.

The following three subsections offer a deeper discussion regarding the identified EDR, ML (SML or UML), and GB frameworks and models given in Sections 3 to 5, respectively. The reader can also refer to Figure 1 for obtaining an aggregated view of each category along with the possible interrelations among the works included in each category.

6.1. EDR log-based policy schemes

As observed from the first and second columns of Table 2, a great variety of 23 multipurpose EDR-related works were identified; all of them proposed some log-based and policy specific technique against the first stages of the existence of LM malevolent incidents. Specifically, all 23 works centered on the presentation of an EDR threat hunting framework, either dedicated exclusively to examine LM incidents or to APT threats, which among others enclose malware LM, pivoting, and other relevant techniques. No less important, five out of 23 works are dedicated exclusively to the endpoint analysis of threats applied on IoT or IIoT realms.

Precisely, with reference to the fourth column of Table 2, six different subcategories of studies were identified among the general theme of EDR threat hunting: data-driven identification, vulnerability assessment via characteristics found in DLL directories of malware executable, APT's API calls, APT malware, APT and LM threats, and threat identification in the IoT ecosystem. That is, a significant number of works from the presented literature emphasized the analysis of general APT malware [18, 13, 30, 31] or LM equivalent techniques [20, 26, 21, 29]. What is more, several researchers concentrated on the identification of APT and LM malevolent events through data-driven procedures and policies towards incident forensic analysis [8, 9, 11, 16, 24, 17, 19, 27].

On the other hand, a DLL-oriented LM identification method destined to the detection of infected pieces of executable code, as those are evidenced in log-traffic collected through Sysmon, was presented in [15]. The subject of LM identification through the analysis of DLL files as part of executable snippets of code is currently limited in terms of dedicated works. To this direction, the work in [21] expanded [15] by proposing a richer set of DLL features towards the presentation of an improved LM log-based detection Sysmon policy. This improves substantially the framework's effectiveness and robustness in terms of FPR and TPR rates. Above that, API-calls patterns were identified and analyzed in the methodology presented in [7]. With reference to the aforementioned 17 papers, it should be noted that only four of them centered on the presentation of a complete EDR framework; these are *ShadowMove* [20], *USBCulprint* [26], *PeX* [21] and *APTHunter* [29].

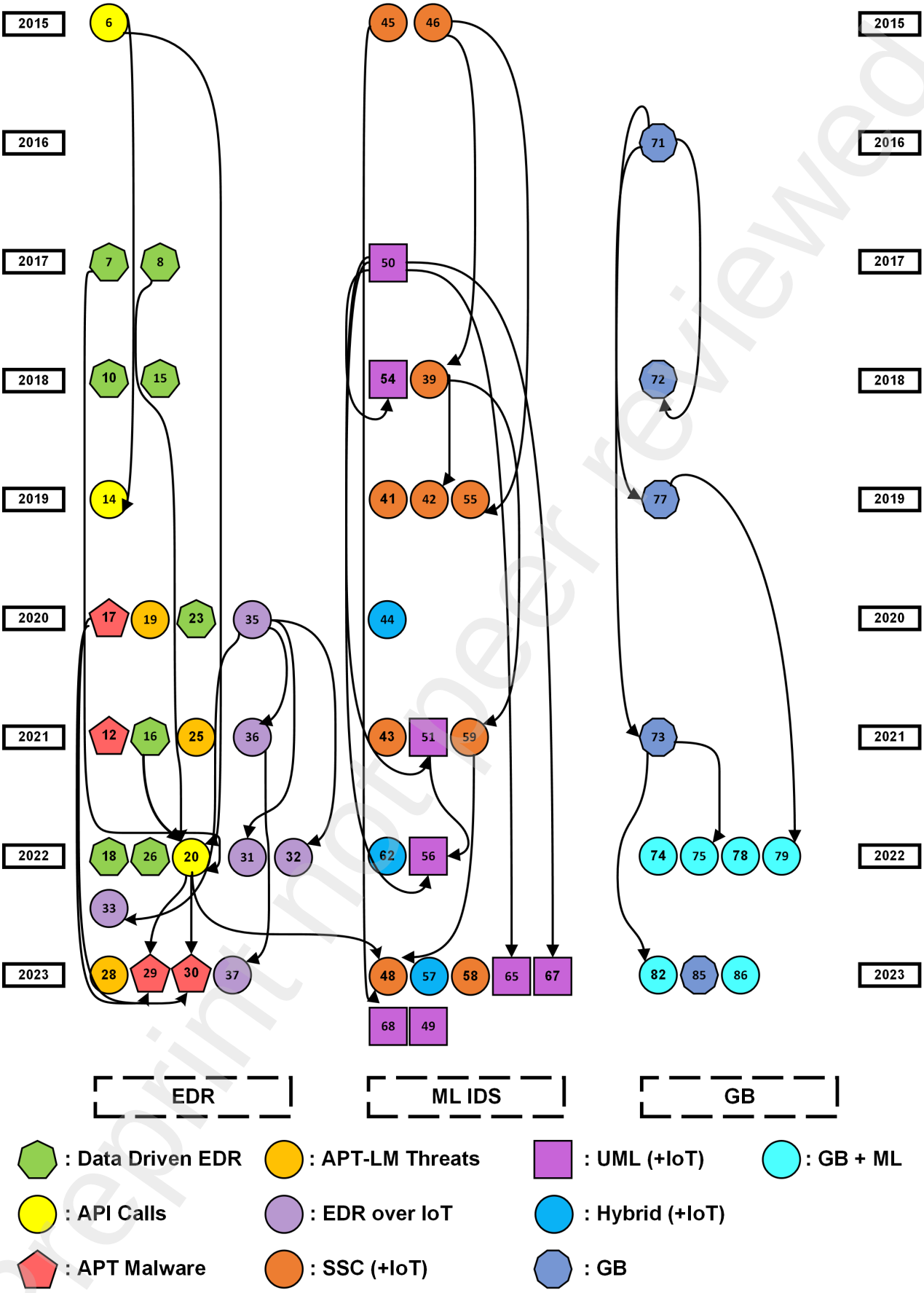

Figure 1: Major EDR, ML (SML, UML) and GB approaches per category in chronological order. Arrows indicate other methods that possibly influenced each detection model. The IoT category is also enclosed as depicted in the index at the bottom of the Figure.

The 13 rest were dedicated to the presentation of general purpose APT and LM characteristics that could be implemented as key feature elements in audit-log SIEM tools and policies.

The six last papers [36, 37, 32, 33, 34] included in the last six bottom lines of Table 2 are related either to the detection of intrusion attempts or the general endpoint security of IoT and IIoT interrelated devices implemented as core to SOHO or corporate networking environments. As opposed to the five aforementioned EDR categories, in the case of IoT technology, more than half of these works presented an autonomous and complete concept that could be leveraged as a general purpose framework for IoT device's security. These are *APIVDAS* [32], *SoK* [33], and *PatrIoT* [38]. Finally, a characteristic common to most works is that they neither construct their own set of data logs and samples, nor provide adequate reference to regularization techniques and hyperparameter optimization steps in case they were evaluated through some ML model.

The contributions in [7, 24, 13, 21, 29, 30] are an exception to the aforesaid rule related to the lack of datasets and ML parameters explanations, creating or using an existing benchmark dataset for evaluation purposes as presented in Table 3. Precisely, four datasets were identified; in chronological order, these are: *Malicia* [39], DARPA Engagement 5 [25], APT29 [14], and LMD-2022 [23]. Regarding these four datasets, the following observations can be made. Malicia is nowadays discontinued, the *DARPA* project, although still quite popular, is considered largely outdated. *APT29* on the other hand is regarded as a descent benchmark dataset for the sandbox evaluation of EDR concepts. Finally, as its name preludes, the *LMD-2022* collection is the first to our knowledge contemporary dataset that was created as a whole from LM Windows logs collected via the Sysmon SIEM tool. In this respect, LMD-2022 is specifically destined to LM scenarios. Interestingly, LMD-2022 is publicly available in three different versions (in terms of the number of data samples) and is delivered in both .evtx and .csv format. Last but not least, all the works included in Table 2 but two have been published from 2018 onwards.

6.2. Machine learning IDS schemes

6.2.1. General IDS ML models

With reference to Section 4 and Tables 4 and 6, half of the presented works (10 out of 20) introduced a general purpose LM IDS model focused on the identification and classification of LM malevolent events via contemporary ML algorithms. The remaining 10 works contributed ML-based schemes destined to the IoT and IIoT ecosystem [56, 60, 63, 57, 58, 59, 66, 68, 69]. As it concerns the former general LM IDS category, six contributions exploited for detection purposes SSC classifiers [46, 47, 40, 42, 43, 45, 44], whereas the rest three utilized UML techniques (either NN, DNN, or a combination of them) [51, 55, 52]. Still in the ML category, as it was also observed in subsection 6.1, most of the presented in Table 4 studies provide very little knowledge related to the data samples upon which their ML models have been tested; naturally, this does not work in favor of repeatability. As detailed next, the majority of these works utilized legacy datasets (such as LANL, KDD, Bot-IoT etc.) or even created their own custom one for evaluation purposes. No adequate references are also provided regarding the regularization and hyperparameter optimization of the aforesaid ML models.

As already briefly mentioned, the great majority of the literature works enclosed in Table 4 up to 2021, utilized either their own custom solutions related to logs collected via the MS Windows Event Viewer tool or their public equivalent *LANL* [41] dataset of multi-source cybersecurity events. As it concerns logs gathered through a custom testbed, all the contributions except [49] did not deposit them in a public repository. However, the *LANL* [41] dataset is nowadays considered an almost outdated corpus due to the many years, since 2015, that it has to be updated with contemporary attacks, including novel LM techniques. More precisely, although exploited so far in six works [55, 40, 43, 44, 52, 51] for evaluation purposes of the presented LM identification ML schemes, the small proportion of the malicious traffic enclosed, legally compels the majority of the authors to conduct artificial reproduction of malevolent samples for balancing the

dataset. However, this manipulation of the dataset samples through over- or under-sampling methods, creates significant concerns regarding the validity of the conducted experiments. We argue that those techniques, say, SMOTE and others, should be treated with great concern. That is, although the initially presented results may be rather promising, in the majority of the cases, the system logs or network traffic under analysis correspond to artificial, unrealistic, samples. This in turn may yield false results regarding the benchmarking process and prediction rates, especially those related to FP classified events. This concern is rather obvious in the works [42, 43, 44, 45], in which the authors, although presented an alternative dataset based on *LANL* [41], they disregarded to release it as public, impeding reproducibility. Another outdated legacy dataset that have been introduced in two different versions, namely Malicia [39] and Malheur [48] was utilized in [46, 47]; however, from 2017 onwards no other reference to it has been detected. As depicted in Table 4, since 2021, the ML experimentation's evaluation practices seem to be ameliorated, with the implementation of datasets enriched with modern APT and LM techniques. Among the aforesaid, most of the works [60, 63, 66, 68, 69] exploited the UNSW-ND15 [61] dataset, while a few implemented the CERT and TWOS corpora, respectively. There is also the case of [49], in which the authors presented a dedicated to LM techniques enhanced version of the LMD-2022 dataset, coined as LMD-2023 [23].

Another noteworthy aspect is that most of the presented works in Table 4, either SSC or DNN, except those in [51, 40, 42, 45, 44, 49], omitted to properly justify their ML model's contribution via the presentation of their selected for the SSC or UML classification experiments features or the distribution of any repository with, say, *Python* or *R* scripts. Another drawback, which also does not aid reproducibility, is that the majority of the authors neglect to mention the hyperparameters on which their ML models were built.

6.2.2. IDS ML models for the IoT or IIoT ecosystem

As already pointed out in subsection 6.2.1, nine out of 19 collected ML LM-related works contributed or evaluated ML-powered IDS schemes targeting the IoT and IIoT realm [56, 60, 63, 57, 58, 59, 66, 68, 69]. Nearly one-fourth of them (2 out of 9) presented SSC models [56, 60], four more [63, 59, 58, 66] utilized DNN algorithms for the binary or multiclass classification of the analyzed logs or traffic, while the rest three [57, 68, 69] presented hybrid multiplexed solutions of both SSC and DNN algorithmic models. From their thorough review in Section 4 and Table 4, almost half of them (4 out of 9) omitted to provide significant details regarding their feature selection or ML model construction processes or any hyperparameter tuning or optimization being applied. This is despite the fact that these proposals did exploit for the evaluation of their models prominent but still not LM-focused benchmark datasets, including UNSW-ND15 [61] and Bot-IoT [65]. On the other hand, the works in [63, 58, 66] presented in a rather descriptive and documented way both their selected features and the algorithmic notion behind the hyperparameter's implementation in their ML models.

Moreover, two more papers categorized as IoT-dedicated [68, 69], presented in full details their DNN model's hyperparameters initialization, omitting however any reference regarding feature selection. Overall, from the extended review of the nine aforesaid IoT or IIoT related studies, it must be emphasized that the improvement in terms of the selection of ML models (namely the majority of the papers implemented novel DNN classification algorithms over multiclass schemes), features selection, presentation of model's initialization is rather noticeable. It can be argued that this situation reflects how important the scientific community considers the implementation of ML IDS knowledge into schemes which are intended for the security of smart devices in the IoT ecosystem.

Finally yet importantly, regarding any datasets created for evaluation purposes of the presented ML schemes, almost half of the IoT-focused studies [60, 63, 66, 68, 69] (5 out of 9) employed contemporary corpora as those are presented in Table 6. In more detail, among the most recent IoT-oriented datasets stand out the IoTD20 [62], the Bot-IoT [65], the X-IIoTD [67], and the Edge-IIoT [85], all created from 2019 onwards. A noteworthy observation is that the work of Jayalaxmi et al. [69] was the first in which the evaluation of the proposed IoT LM IDS scheme was conducted against

five datasets, two of which were derived from corporate IIoT devices implemented in gas pipeline [70] and water tank [71] facilities.

6.2.3. Overall Remarks

Given the above discussion regarding the various ML-powered LM IDS studies reviewed, it can be argued that their majority has been designed and evaluated on datasets that do not conform to a number of key aspects. Specifically, all the datasets up to 2018 (except those mentioned in the last part of subsection 6.2.2 related to the evaluation of IoT or IIoT LM IDS frameworks) do not meet important criteria related to (a) contemporary LM or general purpose APT and malware schemes, (b) public disposal of the selected for the experimentation process features, (c) the regularization and hyperparameter optimization check the st of the ML algorithms, (d) multiclass labeling of the implemented samples instead of the straightforward binary one. As already point out, this situation deprives any chance for reproducibility in favor of potential researchers in the same field.

It should also be mentioned that only three ([45, 21, 49]) out of the total 20 reviewed works in this category were conducted on the basis of Sysmon SIEM's collected log-based traffic, as a means to take advantage of its descriptive header's and event-oriented structure. The remaining ones either omitted to do so or were bounded to the limited legacy equivalent of MS Windows event viewer, limiting significantly the quantity and quality of the collected information. We consider that this shortcoming is mainly due to the lack (at least up to 2022) of a publicly available, open-source converter able to transform the extracted from Sysmon .evtx files to a .csv unlabeled equivalent and compatible for ML algorithmic experimentation. This justifies to an extent why most of the works up to 2018 relied on custom manual identification of LM events, as those were collected via the MS Windows Event Viewer. Therefore, most studies resorted to the pre-processed in comma-separated format legacy *LANL* [41] dataset. Towards this extent, the work in [21] was the first to introduce an EDR policy solution dedicated to the endpoint identification and detection of LM techniques upon Sysmon raw log traffic on an EDR team level. Precisely, they evaluated the proposed EDR policy through the presentation of the *PeX* EDR tool [88], which parses raw .evtx Sysmon files and iterates over them based on the criteria imposed by its incorporated LM policy rules. Further, the work in [21] transitioned from EDR to an ML IDS equivalent [49] through the presentation of a dataset creation tool dubbed *"ETCExp"* [49]. Briefly, *ETCExp* serves researchers in the field to transform voluminous .evtx log files into compatible with ML algorithms unlabeled datasets in .csv format. For more details about the ETCExp tool, the reader is referred to [89].

6.3. Graph-based algorithmic schemes

With reference to subsection 5 and Tables 5 and 6 it can be argued that although the literature abounds of works dedicated to graph-based algorithms under general IDS schemes, only a limited number of them is devoted to the identification of LM techniques, not to mention that those presenting hybrid models for LM incident detection over IoT interrelated incidents outnumber the former. Particularly, 12 works in total have been recognized as the most related to the subject of graph-based model's presentation that classify incoming traffic based on the exploitation of model multidimensional feature analysis techniques over the construction of heterogeneous algorithmic graphs. Almost half of them [75, 76, 72, 73] presented general LM identification concepts, while the other half [83, 86, 87, 78, 79, 80] dedicated to the hunting of LM malevolent incidents over IoT or IIoT interrelated devices.

What is more, 6 out of 12 works, namely [75, 76, 79, 83, 86, 87], presented a hybrid framework, which leverages the powerful features of graph-based algorithmic techniques with the robustness of modern DNN models. This is rather obvious in the work that presented *LMTracker* [75], a custom LM and APTs identification and detection algorithm. Specifically, LMTracker's functionality hinges on the graphical representation of features as nodes and vectors through the construction of heterogeneous graphs, finally classifying the graphically represented features in a multidimensional

way via advanced multilayered Autoencoders NN. Finally, it is worth to be mentioned that the work in [86] was the only that mixed four pillars of contemporary technology paradigms, namely bi-level optimization theory, Blockchain, feature reconstruction into nodes and vectors via graph-based algorithms, and traffic anomaly detection and classification via the eKNN model. In terms of the identified datasets, the works in [72, 75] exploited again the legacy LANL dataset [41] towards the evaluation of their general graph-based IDS schemes. Above that, only the authors in [76] employed the enhanced contemporary APT3 [77] corpus, comprising more than 12k of APT samples. Once again, since all the LM-related IoT or IIoT specified works were created from 2022 forward (except [78] in 2019) utilize contemporary datasets for their model's evaluation. A prominent example of this situation is the multipurpose CSE-CIC-IDS2018 dataset [82], comprising traffic related to more than seven attack scenarios over 420 infected corporate terminal, 30 servers and all these organized ready for ML manipulation into 80 unbalanced features.

7. Conclusion

In the era of Internet of Everything (IoV), LM has evolved as a game-changing tactic in the quiver of cybercriminals, especially when it comes to APT. In this volatile ecosystem, the present article offers the first to our knowledge full-fledged, systematic review of literature works about schemes designed to timely identify and possibly counteract LM in the network perimeter or deeper in the network, mainly in the form of an IDS. We differentiate among three kinds of such defensive solutions, namely, those that hinge on either graph algorithmic schemes, ML classification models, or EDR log-based policy strategies. Interestingly, for each category of solutions, an additional distinction is made between schemes proposed for general network domains, especially enterprise intranets, and IoT or IIoT ones. Just as important, an extensive and thorough array of essential observations and discussions is given, highlighting the utilized methodologies and datasets, as well as certain deficiencies and challenges.

With reference to Section 6, several issues exist that are addressed only partly or not at all, leading to the following key takeaways. First, all the published works up to 2022, although built on prominent EDR, ML, or GB solutions, largely neglected to present adequate information regarding the technical characteristics of each model's evaluation, the feature engineering and selection techniques followed, and the dataset upon which the proposed framework's effectiveness was tested. Second, most of the published studies relied on legacy dataset solutions, such as LANL [41] and UNSW-NB15 [61], which however are not LM-oriented. This observation also largely applies to contributions from 2022 onwards, which leaned on contemporary, yet not-so-relevant datasets, namely Bot-IoT [65] and Edge-IIoT [85]. An exception to this rule are the works in [49, 21]. Third, an interesting, yet disregarded, topic in LM identification is the creation of unsupervised and federated learning models. That is, UML detection methods will provide to current studies an adequate level of applicability in real-life IDs scenarios, where labelled datasets are scarce. On the other hand, federated models are expected to decentralize LM detection, also increasing privacy in terms of data transfer between the participating clients and ML server. All in all, we anticipate that this work will provide fundamental insight into this rapidly changing and interesting research branch, and fulfil the needs of a solid reference point for the interested readers.

Abbreviations List - The following abbreviations are used in this manuscript:

AD	AdaBoost
AE	Autoencoder
API	Application Programming Interface
APTs	Advanced Persistent Threats
BAS	Building Automation Systems

BBN	Bayesian Belief Networks
CB	CatBoost Classifier
CERT	Computer Emergency Response Team
CFBPNN	Cascade Forward Back Propagation Neural Network
CIC	Canadian Institute for Cybersecurity
CNB	Categorical Naive Bayes
CNN	Convolutional Neural Networks
CoAP	Constrained Application Protocol
CSE	Communications Security Establishment
CSIRT	Computer Security Incident Response Team
CSV	Comma-Separated Values
CTIO	Cyber Threat Intelligence Ontology
CTI	Cyber Threat Intelligence
CTIR	Cyber Threat Incident Reports
CVE	Common Vulnerabilities Exposures
CWE	Common Weakness Enumeration
DARPA	U.S. Department of Defense - Defense Advanced Research Projects Agency
DFBnB	Depth-first branch and bound heuristic search algorithm
DGBR	Dynamic Graph-based Reachability Model
DL	Deep Learning Algorithms
DLL	Dynamic Link Library
DNN	Deep Neural Networks
DT	Decision Tree Algorithm
EDRPolicy	End-point Detection and Response Policy
eKNN	Enhanced K-Nearest Neighbor Algorithm
EoHT	Exploitation of Hashing LM Techniques (PtH, PtT, GT, ST via Mimikatz)
EoRS	Exploitation of Remote Services LM Techniques
ERS	Exploitation of Remote Services
ETCExp	evtx_To_CSV_Export Tool
ET	Extra-Trees
EVTX	Windows XML EventLog
FF	Feed-Forward
FNR	False Negative Rates
FPR	False Positive Rates
GAN	Graph Attention Networks
GB	Graph-based Model
GBC	Gradient-Boosting Classifier
GNB	Gaussian Naive Bayes Algorithm
GRR	Google Rapid Response
GRUs	Gated Recurrent Units
GT	Golden Ticket attack
IDS	Intrusion Detection System (IDS)

IF	Isolation Forest
IIoT	Industrial Internet of Things
IoT	Internet of Things
IoV	Internet of Everything
JD	Jackard Distance
KNN	K-Nearest Neighbors Algorithm
LaBi	Label Binarizer
LANL	Los Alamos National Laboratory Dataset
LB	LogitBoost Algorithm
L-based	Linear-based Algorithms
LCSs	Longest Common Subsequences
LGBM	Light Gradient Boosting Model Algorithm
LiR	Linear Regression
LM	Lateral Movement Techniques
LODA	Lightweight On-Line Anomaly Detection
LOF	Local Outlier Factor
LoR	Logistic Regression Algorithm
LSTM	Long Short-Term Memory
MADO	Median Absolute Deviation-based outlier detection
MCTS	Monte Carlo Tree Search
Min-Max	Min-Max Scaler
ML	Machine Learning
MLP	Multilayer Perceptron
MQTT	Message Queuing Telemetry Transport
MSDN	Microsoft Development Network
MV	Majority Voting Algorithm
NB	Naive Bayes Algorithm
NLP	Natural Language Process
NNeighbour	Nearest Neighbour
NN	Neural Networks
NT	Network Traffic
OHE	One-Hot Encoding
OS	Operating System
PeX - v1	Python_Evtx_Analyzer (PeX - v1)
PCA	Principal Component Analysis
PtH	Pass the Hash attack
PtT	Pass the Ticket attack
RNN	Recurrent Neural Networks
RF	Random Forest Algorithm
SB	Smart Buildings
SGDC	Stochastic Gradient Descent Classification Algorithm
SIEM	Security Information and Event Management logging tools

SLR	Systematic Literature Review
SML	Supervised Machine Learning
SMOTE	Synthetic Minority Oversampling Technique
SOAR	Security Orchestrated Automation and Response
SOC	Security Operation Center's
SOHO	Small Office Home Office
Softmax	Softmax activation function
SSC	Supervised shallow Classification
ST	Silver Ticket attack
ST-based	Stochastic - Based algorithms
STD	Standard Deviation
SVC	Support Vector Classification
SVM	Support Vector Machines
TISC	Threat Intelligence Sharing Scheme
TTPs	Tactics - Techniques - Procedures
TXT	Standard Text Document (Contains Plain Text
UML	Unsupervised Machine Learning
WEV	Windows Event Viewer Application
XDR	Extended Detection and Response schema
XML	Extensible Markup Language
ZT	Zero Trust

References

[1] G. Kambourakis, C. Kolias, A. Stavrou, The mirai botnet and the iot zombie armies, in: 2017 IEEE Military Communications Conference, MILCOM 2017, Baltimore, MD, USA, October 23-25, 2017, IEEE, 2017, pp. 267–272. doi:10.1109/MILCOM.2017.8170867.

[2] B. Stojanović, K. Hofer-Schmitz, U. Kleb, Apt datasets and attack modeling for automated detection methods: A review, Computers & Security 92 (2020) 101734. doi:https://doi.org/10.1016/j.cose.2020.101734.
URL https://www.sciencedirect.com/science/article/pii/S0167404820300213

[3] M. Tatam, B. Shanmugam, S. Azam, K. Kannoorpatti, A review of threat modelling approaches for apt-style attacks, Heliyon 7 (1) (2021) e05969. doi:https://doi.org/10.1016/j.heliyon.2021.e05969.
URL https://www.sciencedirect.com/science/article/pii/S2405844021000748

[4] M. Abu Talib, Q. Nasir, A. Bou Nassif, T. Mokhamed, N. Ahmed, B. Mahfood, Apt beaconing detection: A systematic review, Computers & Security 122 (2022) 102875. doi:https://doi.org/10.1016/j.cose.2022.102875.
URL https://www.sciencedirect.com/science/article/pii/S0167404822002693

[5] Z. Chen, J. Liu, Y. Shen, M. Simsek, B. Kantarci, H. T. Mouftah, P. Djukic, Machine learning-enabled iot security: Open issues and challenges under advanced persistent threats, ACM Comput. Surv. 55 (5). doi:10.1145/3530812.
URL https://doi.org/10.1145/3530812

[6] V. Kampourakis, V. Gkioulos, S. Katsikas, A systematic literature review on wireless security testbeds in the cyber-physical realm, Computers & Security 133 (2023) 103383. doi:https://doi.org/10.1016/j.cose.2023.103383.
URL https://www.sciencedirect.com/science/article/pii/S0167404823002936

[7] Y. Ki, E. Kim, H. K. Kim, A novel approach to detect malware based on api call sequence analysis, Int. J. Distributed Sens. Networks.

[8] J. Coordination, Detecting lateral movement through tracking event logs (June 2017).
URL https://www.jpcert.or.jp/english/pub/sr/20170612ac-ir_research_en.pdf

[9] J. Coordination, Detecting lateral movement through tracking event logs (version 2) (December 2017).
URL https://blogs.jpcert.or.jp/en/2017/12/research-report-released-detecting-lateral-movement-through-
html

[10] M. Russinovich, T. Garnier, Sysmon v13. 22, Retrieved June 28 (2021) 2021.
URL https://learn.microsoft.com/en-us/sysinternals/downloads/sysmon

[11] V. Mavroeidis, A. Jøsang, Data-driven threat hunting using sysmon, in: Proceedings of the 2nd International Conference on Cryptography, Security and Privacy, 2018, pp. 82–88.

[12] V. Mavroeidis, S. Bromander, Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence, in: 2017 European Intelligence and Security Informatics Conference (EISIC), 2017, pp. 91–98. doi:10.1109/EISIC.2017.20.

[13] A. Berady, M. Jaume, V. V. T. Tong, G. Guette, From ttp to ioc: Advanced persistent graphs for threat hunting, IEEE Transactions on Network and Service Management 18 (2) (2021) 1321–1333. doi:10.1109/TNSM.2021.3056999.

[14] M. Labs, "apt29" mordor labs dataset collections (2020).
URL https://github.com/OTRF/detection-hackathon-apt29

[15] W. Matsuda, M. Fujimoto, T. Mitsunaga, Real-time detection system against malicious tools by monitoring dll on client computers, in: 2019 IEEE Conference on Application, Information and Network Security (AINS), 2019, pp. 36–41. doi:10.1109/AINS47559.2019.8968697.

[16] U. Jain, et al., Lateral movement detection using elk stack, Ph.D. thesis, University of Houston (2018).

[17] P. Rajesh, M. Ismail. Ismail. B, M. Alam, M. Tahernezhadi, Network forensics investigation in virtual data centers using elk, in: 2021 International Symposium on Electrical, Electronics and Information Engineering, 2021, pp. 175–179.

[18] M. G. El-Hadidi, M. A. Azer, Detecting mimikatz in lateral movements using mutex, in: 2020 15th International Conference on Computer Engineering and Systems (ICCES), 2020, pp. 1–6. doi:10.1109/ICCES51560.2020.9334643.

[19] S. Agarwal, A. Sable, D. Sawant, S. Kahalekar, M. K. Hanawal, Threat detection and response in linux endpoints, in: 2022 14th International Conference on COMmunication Systems & NETworkS (COMSNETS), 2022, pp. 447–449. doi:10.1109/COMSNETS53615.2022.9668567.

[20] A. Niakanlahiji, J. Wei, M. R. Alam, Q. Wang, B.-T. Chu, ShadowMove: A stealthy lateral movement strategy, in: 29th USENIX Security Symposium (USENIX Security 20), USENIX Association, 2020, pp. 559–576.
URL https://www.usenix.org/conference/usenixsecurity20/presentation/niakanlahiji

[21] C. Smiliotopoulos, K. Barmpatsalou , G. Kambourakis, Revisiting the detection of lateral movement through sysmon, Applied Sciences 12 (15). doi:10.3390/app12157746.
URL https://www.mdpi.com/2076-3417/12/15/7746

[22] MITRE, Lateral movement - the adversary is trying to move through your environment. (July 2019).
URL https://attack.mitre.org/tactics/TA0008/

[23] C. Smiliotopoulos, G. Kambourakis, "lmd" sysmon dataset collections (2023).
URL https://github.com/ChristosSmiliotopoulos/Lateral-Movement-Dataset--LMD_Collections.git

[24] N. Michael, J. Mink, J. Liu, S. Gaur, W. U. Hassan, A. Bates, On the forensic validity of approximated audit logs, in: Annual Computer Security Applications Conference, ACSAC '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 189–202. doi:10.1145/3427228.3427272.
URL https://doi.org/10.1145/3427228.3427272

[25] D. A. R. P. Agency, "darpa" transparent computing engagement 5 data release (2023).
URL https://github.com/darpa-i2o/Transparent-Computing

[26] M. Guri, Usbculprit: Usb-borne air-gap malware, in: Proceedings of the 2021 European Interdisciplinary Cybersecurity Conference, EICC '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 7–13. doi:10.1145/3487405.3487412.
URL https://doi.org/10.1145/3487405.3487412

[27] M. Mundt, H. Baier, Threat-based simulation of data exfiltration towards mitigating multiple ransomware extortions, Digital ThreatsJust Accepted. doi:10.1145/3568993.
URL https://doi.org/10.1145/3568993

[28] MITRE, Mitre att&ck - the adversary is trying to move through your environment. (July 2019).
URL https://attack.mitre.org/

[29] M. Mahmoud, M. Mannan, A. Youssef, Apthunter: Detecting advanced persistent threats in early stages, Digital Threats 4 (1). doi:10.1145/3559768.
URL https://doi.org/10.1145/3559768

[30] N.-E. Park, Y.-R. Lee, S. Joo, S.-Y. Kim, S.-H. Kim, J.-Y. Park, S.-Y. Kim, I.-G. Lee, Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks, Computers and Electrical Engineering 105 (2023) 108548. doi:https://doi.org/10.1016/j.compeleceng.2022.108548.
URL https://www.sciencedirect.com/science/article/pii/S0045790622007637

[31] P. Bajpai, R. Enbody, Know thy ransomware response: A detailed framework for devising effective ransomware response strategies, Digital ThreatsJust Accepted. doi:10.1145/3606022.
URL https://doi.org/10.1145/3606022

[32] R. S. Marques, H. M. Al-Khateeb, G. Epiphaniou, C. Maple, APIVADS: A novel privacy-preserving pivot attack detection scheme based on statistical pattern recognition, IEEE Trans. Inf. Forensics Secur. 17 (2022) 700–715. doi:10.1109/TIFS.2022.3146076.
URL https://doi.org/10.1109/TIFS.2022.3146076

[33] S. Xiao, Y. Ye, N. Kanwal, T. Newe, B. Lee, Sok: context and risk aware access control for zero trust systems, Security and Communication Networks 2022.
URL https://doi.org/10.1155/2022/7026779

[34] S. Ahamed, L. Ramanathan, Real-time heuristic-based detection of attacks performed on a linux machine using osquery, SN Comput. Sci. 3 (5) (2022) 405. doi:10.1007/s42979-022-01288-6.
URL https://doi.org/10.1007/s42979-022-01288-6

[35] D. J. Bianco, Enterprise detection and response : "the pyramid of pain" (2014).
URL http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html

[36] D. Weissman, A. Jayasumana, Integrating iot monitoring for security operation center, in: 2020 Global Internet of Things Summit (GIoTS), 2020, pp. 1–6. doi:10.1109/GIOTS49054.2020.9119680.

[37] D. R. dos Santos, M. Dagrada, E. Costante, Leveraging operational technology and the internet of things to attack smart buildings, J. Comput. Virol. Hacking Tech. 17 (1) (2021) 1–20. doi:10.1007/s11416-020-00358-8.
URL https://doi.org/10.1007/s11416-020-00358-8

[38] E. Süren, F. Heiding, J. Olegård, R. Lagerström, Patriot: practical and agile threat research for iot, International Journal of Information Security 22 (1) (2023) 213–233.
URL https://doi.org/10.1007/s10207-022-00633-3

[39] A. Nappa, M. Z. Rafique, J. Caballero, The malicia dataset: Identification and analysis of drive-by download operations, Int. J. Inf. Secur. 14 (1) (2015) 15–33. doi:10.1007/s10207-014-0248-7.
URL https://doi.org/10.1007/s10207-014-0248-7

[40] G. Kaiafas, G. Varisteas, S. Lagraa, R. State, C. D. Nguyen, T. Ries, M. Ourdane, Detecting malicious authentication events trustfully, in: NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium, 2018, pp. 1–6. doi:10.1109/NOMS.2018.8406295.

[41] A. D. Kent, Cybersecurity Data Sources for Dynamic Network Research, in: Dynamic Networks in Cybersecurity, Imperial College Press, 2015.

[42] H. Bian, T. Bai, M. A. Salahuddin, N. Limam, A. A. Daya, R. Boutaba, Host in danger? detecting network intrusions from authentication logs, in: 2019 15th International Conference on Network and Service Management (CNSM), 2019, pp. 1–9. doi:10.23919/CNSM46954.2019.9012700.

[43] T. Bai, H. Bian, A. A. Daya, M. A. Salahuddin, N. Limam, R. Boutaba, A machine learning approach for rdp-based lateral movement detection, in: 2019 IEEE 44th Conference on Local Computer Networks (LCN), 2019, pp. 242–245. doi:10.1109/LCN44214.2019.8990853.

[44] H. Bian, T. Bai, M. A. Salahuddin, N. Limam, A. A. Daya, R. Boutaba, Uncovering lateral movement using authentication logs, IEEE Transactions on Network and Service Management 18 (1) (2021) 1049–1063. doi:10.1109/TNSM.2021.3054356.

[45] C.-M. Chen, G.-H. Syu, Z.-X. Cai, Analyzing system log based on machine learning model, International Journal of Network Security 22 (6) (2020) 925–933.

[46] M. Narouei, M. Ahmadi, G. Giacinto, H. Takabi, A. Sami, Dllminer: structural mining for malware detection, Security and Communication Networks 8 (18) (2015) 3311–3322.

[47] J. T. Juwono, C. Lim, A. Erwin, A comparative study of behavior analysis sandboxes in malware detection, in: International Conference on New Media (CONMEDIA), 2015, p. 73.

[48] K. Rieck, P. Trinius, C. Willems, T. Holzaff, Automatic analysis of malware behavior using machine learning, J. Comput. Secur. 19 (4) (2011) 639–668.

[49] C. Smiliotopoulos, G. Kambourakis, K. Barbatsalou, On the detection of lateral movement through supervised machine learning and an open-source tool to create turnkey datasets from sysmon logs, International Journal of Information Security 22 (2023) 1893–1919. doi:https://doi.org/10.1007/s10207-023-00725-8.
URL https://link.springer.com/article/10.1007/s10207-023-00725-8

[50] D. He, H. Gu, S. Zhu, S. Chan, M. Guizani, A comprehensive detection method for the lateral movement stage of apt attacks, IEEE Internet of Things Journal (2023) 1–1doi:10.1109/JIOT.2023.3322412.

[51] A. Bohara, M. A. Noureddine, A. Fawaz, W. H. Sanders, An unsupervised multi-detector approach for identifying malicious lateral movement, in: 2017 IEEE 36th Symposium on Reliable Distributed Systems (SRDS), 2017, pp. 224–233. doi:10.1109/SRDS.2017.31.

[52] D. C. Le, N. Zincir-Heywood, Anomaly detection for insider threats using unsupervised ensembles, IEEE Transactions on Network and Service Management 18 (2) (2021) 1152–1164. doi:10.1109/TNSM.2021.3071928.

[53] C. R.Trzeciak, The CERT Insider Threat Database, Carnegie Mellon University's Software Engineering Institute Blog, 2011.

[54] A. Harilal, F. Toffalini, J. Castellanos, J. Guarnizo, I. Homoliak, M. Ochoa, Twos: A dataset of malicious insider threat behavior based on a gamified competition, in: Proceedings of the 2017 International Workshop on Managing Insider Security Threats, MIST '17, Association for Computing Machinery, New York, NY, USA, 2017, p. 45–56. doi:10.1145/3139923.3139929.
URL https://doi.org/10.1145/3139923.3139929

[55] M. Chen, Y. Yao, J. Liu, B. Jiang, L. Su, Z. Lu, A novel approach for identifying lateral movement attacks based on network embedding, in: 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom), 2018, pp. 708–715. doi:10.1109/BDCloud.2018.00107.

[56] U. Noor, Z. Anwar, A. W. Malik, S. Khan, S. Saleem, A machine learning framework for investigating data breaches based on semantic analysis of adversary's attack patterns in threat intelligence repositories, Future Generation Computer Systems 95 (2019) 467–487. doi:https://doi.org/10.1016/j.future.2019.01.022.
URL https://www.sciencedirect.com/science/article/pii/S0167739X18306708

[57] B. A. Powell, Role-based lateral movement detection with unsupervised learning, Intelligent Systems with Applications 16 (2022) 200106. doi:https://doi.org/10.1016/j.iswa.2022.200106.
URL https://www.sciencedirect.com/science/article/pii/S2667305322000448

[58] M. Imran, H. Sidd, A. Raza, M. Raza, F. Rustam, I. Ashraf, A performance overview of machine learning-based defense strategies for advanced persistent threats in industrial control systems, Computers & Security 134 (2023) 103445. doi:10.1016/j.cose.2023.103445.

[59] L. González-Manzano, J. M. de Fuentes, F. Lombardi, C. Ramos, A technical characterization of apts by leveraging public resources, International Journal of Information Security (2023) 1–18doi:https://doi.org/10.1007/s10207-023-00706-x.

[60] M. Arifeen, A. Petrovski, S. Petrovski, Automated microsegmentation for lateral movement prevention in industrial internet of things (iiot), in: 2021 14th International Conference on Security of Information and Networks (SIN), Vol. 1, 2021, pp. 1–6. doi:10.1109/SIN54109.2021.9699232.

[61] N. Moustafa, J. Slay, Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set), in: 2015 Military Communications and Information Systems Conference (MilCIS), 2015, pp. 1–6. doi:10.1109/MilCIS.2015.7348942.

[62] I. Ullah, Q. Mahmoud, A scheme for generating a dataset for anomalous activity detection in iot networks (2020) 508–520doi:10.1007/978-3-030-47358-7_52.

[63] N. Koroniotis, N. Moustafa, J. Slay, A new intelligent satellite deep learning network forensic framework for smart satellite networks, Computers and Electrical Engineering 99 (2022) 107745. doi:https://doi.org/10.1016/j.compeleceng.2022.107745.
URL https://www.sciencedirect.com/science/article/pii/S0045790622000556

[64] C. I. of Cyber Security, Nsl-kdd dataset (2020).
URL https://www.unb.ca/cic/datasets/nsl.html

[65] N. Moustafa, The bot-iot dataset (2019). doi:10.21227/r7v2-x988.
URL https://dx.doi.org/10.21227/r7v2-x988

[66] H. C. Altunay, Z. Albayrak, A hybrid cnn+lstm-based intrusion detection system for industrial iot networks, Engineering Science and Technology, an International Journal 38 (2023) 101322. doi:https://doi.org/10.1016/j.jestch.2022.101322.
URL https://www.sciencedirect.com/science/article/pii/S2215098622002312

[67] M. Al-Hawawreh, E. Sitnikova, N. Aboutorab, X-iiotid: A connectivity- and device-agnostic intrusion dataset for industrial internet of things (2021). doi:10.21227/mpb6-py55.
URL https://dx.doi.org/10.21227/mpb6-py55

[68] M. Sarhan, S. Layeghy, N. Moustafa, M. Portmann, Cyber threat intelligence sharing scheme based on federated learning for network intrusion detection, Journal of Network and Systems Management 31 (1) (2023) 3. doi:https://doi.org/10.1007/s10922-022-09691-3.

[69] P. Jayalaxmi, R. Saha, G. Kumar, M. Alazab, M. Conti, X. Cheng, Pignus: A deep learning model for ids in industrial internet-of-things, Computers & Security 132 (2023) 103315. doi:https://doi.org/10.1016/j.cose.2023.103315.
URL https://www.sciencedirect.com/science/article/pii/S0167404823002250

[70] H. I. F.-L. D. (HIFLD), Natural gas pipelines dataset (2019).
URL https://hifld-geoplatform.opendata.arcgis.com/datasets/geoplatform::natural-gas-pipelines/about

[71] W. W. Corporation, Water tank (wcorp-076) (2023).
URL https://catalogue.data.wa.gov.au/it/dataset/water-tank

[72] E. Purvine, J. R. Johnson, C. Lo, A graph-based impact metric for mitigating lateral movement cyber attacks, in: Proceedings of the 2016 ACM Workshop on Automated Decision Making for Active Cyber Defense, SafeConfig '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 45–52. doi:10.1145/2994475.2994476.
URL https://doi.org/10.1145/2994475.2994476

[73] Q. Liu, J. W. Stokes, R. Mead, T. Burrell, I. Hellen, J. Lambert, A. Marochko, W. Cui, Latte: Large-scale lateral movement detection, in: MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM), 2018, pp. 1–6. doi:10.1109/MILCOM.2018.8599748.

[74] G. Ho, M. Dhiman, D. Akhawe, V. Paxson, S. Savage, G. M. Voelker, D. Wagner, Hopper: Modeling and detecting lateral movement, in: 30th USENIX Security Symposium (USENIX Security 21), USENIX Association, 2021, pp. 3093–3110.
URL https://www.usenix.org/conference/usenixsecurity21/presentation/ho

[75] Y. Fang, C. Wang, Z. Fang, C. Huang, Lmtracker: Lateral movement path detection based on heterogeneous graph embedding, Neurocomputing 474 (2022) 37–47. doi:https://doi.org/10.1016/j.neucom.2021.12.026.
URL https://www.sciencedirect.com/science/article/pii/S0925231221018646

[76] C.-K. Chen, S.-C. Lin, S.-C. Huang, Y.-T. Chu, C.-L. Lei, C.-Y. Huang, Building machine learning-based threat hunting system from scratch, Digital Threats 3 (3). doi:10.1145/3491260.
URL https://doi.org/10.1145/3491260

[77] H. Haddadpajouh, A. Azmoodeh, A. Dehghantanha, R. M. Parizi, Mvfcc: A multi-view fuzzy consensus clustering model for malware threat attribution (2020).
URL https://cybersciencelab.org/advanced-persistent-threat-apt-malware-dataset/

[78] N. Agmon, A. Shabtai, R. Puzis, Deployment optimization of iot devices through attack graph analysis, in: Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks, WiSec '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 192–202. doi:10.1145/3317549.3323411.
URL https://doi.org/10.1145/3317549.3323411

[79] X. Yang, G. Peng, D. Zhang, Y. Lv, et al., An enhanced intrusion detection system for iot networks based on deep learning and knowledge graph, Security and Communication Networks 2022. doi:https://doi.org/10.1155/2022/4748528.
URL https://www.hindawi.com/journals/scn/2022/4748528/

[80] Y. Wang, Y. Guo, C. Fang, An end-to-end method for advanced persistent threats reconstruction in large-scale networks based on alert and log correlation, Journal of Information Security and Applications 71 (2022) 103373. doi:https://doi.org/10.1016/j.jisa.2022.103373.
URL https://www.sciencedirect.com/science/article/pii/S2214212622002186

[81] MITRE, Lateral movement - the adversary is trying to move through your environment. (July 2019).

[82] C. S. E. C. . the Canadian Institute for Cybersecurity (CIC), Cse-cic-ids2018 dataset (2018).
URL https://www.unb.ca/cic/datasets/ids-2018.html

[83] S. H. Javed, M. B. Ahmad, M. Asif, W. Akram, K. Mahmood, A. K. Das, S. Shetty, Apt adversarial defence mechanism for industrial iot enabled cyber-physical system, IEEE Access 11 (2023) 74000–74020. doi:10.1109/ACCESS.2023.3291599.

[84] S. Myneni, A. Chowdhary, A. Sabur, S. Sengupta, G. Agrawal, D. Huang, M. Kang, Dapt 2020 - constructing a benchmark dataset for advanced persistent threats, in: G. Wang, A. Ciptadi, A. Ahmadzadeh (Eds.), Deployable Machine Learning for Security Defense, Springer International Publishing, Cham, 2020, pp. 138–163.

[85] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, H. Janicke, Edge-iiotset: A new comprehensive realistic cyber security dataset of iot and iiot applications for centralized and federated learning, IEEE Access 10 (2022) 40281–40306. doi:10.1109/ACCESS.2022.3165809.

[86] A. A. M. Sharadqh, H. A. M. Hatamleh, A. M. A. Alnaser, S. S. Saloum, T. A. Alawneh, Hybrid chain: Blockchain enabled framework for bi-level intrusion detection and graph-based mitigation for security provisioning in edge assisted iot environment, IEEE Access 11 (2023) 27433–27449. doi:10.1109/ACCESS.2023.3256277.

[87] A. Kumar, V. L. L. Thing, Raptor: Advanced persistent threat detection in industrial iot via attack stage correlationarXiv:2301.11524, doi:https://doi.org/10.48550/arXiv.2301.11524.
URL https://arxiv.org/abs/2301.11524

[88] C. Smiliotopoulos, K. Barbatsalou, G. Kambourakis, Python_evtx_analyzer (pex - v1) (2022).
URL https://github.com/ChristosSmiliotopoulos/Python_Evtx_Analyzer.git

[89] C. Smiliotopoulos, G. Kambourakis, evtx_to_csv_export tool (etcexp) (2023).
URL https://github.com/ChristosSmiliotopoulos/Python-Projects-Repository.git