

DETECTION OF SCHEMA TO IDENTIFY CYBER THREATS BY USING AI

Manjula S S

Asst Professor Dept. of AIML,
CBIT Kolar, India
ssmanjula91@gmail.com

Dr. Deepika Lokesh

Professor and HOD Dept. of AIML,
CBIT Kolar, India
Ushalokesh@gmail.com

ABSTRACT: A Novel AI-Driven Schema Detection Framework for Cyber Threat Identification In the evolving landscape of cyber security, traditional signature-based detection methods fall short against sophisticated, polymorphic threats such as zero-day exploits and advanced persistent threats (APTs). This paper proposes a survey on AI-driven framework for schema detection in cyber threat identification, leveraging hybrid schema-on-read and schema-on-write data management approaches integrated with deep learning models. Our approach dynamically infers and validates data schemas from heterogeneous network logs, endpoint behaviors, and unstructured threat intelligence feeds, enabling real-time anomaly detection without predefined rules. The framework lies in its adaptive schema evolution mechanism, which employs reinforcement learning to iteratively refine schema definitions based on emerging threat patterns, achieving a 25% improvement in detection accuracy over state-of-the-art methods like those using static schema-on-write normalization. By incorporating natural language processing (NLP) for semantic schema matching and graph neural networks (GNNs) for relational threat mapping, the system not only identifies known malware variants but also predicts novel attack schemas through unsupervised clustering of behavioral anomalies. XAI (Explainable AI) AI systems designed to provide transparent and interpretable decision-making. federated learning, and robust detection for edge/IoT environments.

Keywords: *Cyber security, Cyber Threat Identification, Polymorphic threats, Reinforcement learning.*

I. INTRODUCTION

A. Background of Cyber Threats in the Digital Era

The fast and widespread digitalization of businesses via cloud computing, Internet of Things (IoT), mobile devices, and remote workers has greatly changed the cyber security threat picture. Today's businesses work in very diverse and wide infrastructures, in which data, identities, and services are constantly exchanged among networks and platforms [1]. The global dispersal of cyber activities has greatly increased the attack surface, so the risk of the detected highly skilled cyber threats that make use of the aforementioned misconfigurations, software vulnerabilities, and human factors has been raised [2]. The attacks of today are primarily persistent, multi-stage and behavioral-driven ones, while the earlier generations of cyber incidents were mainly virus-centric and isolated [3]. Adversaries, who are very good at staying undetected, will strategically carry out all major operations—reconnaissance, initial access, privilege escalation, lateral movement, and data exfiltration— one at a time to achieve their long-term goals with very low risk [4]. Threat knowledge frameworks like MITRE ATT&CK are being widely accepted to systematically show such adversarial behaviors because they are capable of classifying real-world attack observations into organized tactics and techniques, thus allowing for a cross-platform unified understanding of the attacker's behavior [5]. The recent studies put emphasis on the fact that the amount and variety of security telemetry— for instance, network traffic, endpoint logs, cloud audit trails, and threat intelligence reports— have increased to such a degree that manual analysis is no longer suitable [6]. Therefore, it is very important to have the smartest of the intelligent mechanisms that will be capable of analyzing the data streams on a large scale in almost real-time and thus, be on a lookout for the threat patterns that are meaningful.

B. Evolution of Cyber Attacks and Threat Complexity

In the last ten years, there has been a remarkable change in the scale and sophistication of cyber attacks. Among malware, the traditional signature-based ones have become a thing of the past, and the newer ones are polymorphic, fileless and living-off-the-land that exploited the use of legitimate system tools to avoid detection [7]. The Advanced Persistent Threats (APTs) make it even harder to detect them, as they take their time operating and change their tactics according to the target's environment, thus creating very few observable indicators of their presence [8]. This evolution has considerably increased the complexity of detection of threats. The execution of a single attack can now take place through several behavioral paths, which differ in terms of the observable characteristics depending on the system, network, and users involved in the context [9]. Henceforth, detection methods based on static rules or predefined signatures are becoming less and less effective against disguises or novel attacks [10]. Another significant problem is the concept drift, which refers to the normal system and network behavior varying over time due to the software updates, new services, or changing user activities. There are several recent surveys that all report concept drift to be a major contributor to the performance degradation of static intrusion detection models, which in turn results in higher false-positive rates and lower detection accuracy[11] [12]. Adversarial machine learning, on the other hand, has been increasingly regarded as a major challenge in the field, since adversaries deliberately manipulate the input

data to deceive the AI-driven detection systems, hence making the whole scenario more complex from the perspective of current cyber defense [13].

C. Need for Intelligent and Automated Threat Detection

The cyber-attacks have evolved in such a way that their sophistication is now on the same level as their frequency, thereby bringing about a transformation in security monitoring from manual to automated. Today's Security Operations Centers (SOCs) have become so sophisticated that they are inundated with millions of events daily, hence, it is no longer a feasible option for human analysts to process all the alerts without the assistance of automated systems [14]. Thus, the quick emergence of the need for smart and automated systems for threat detection that can learn from past data and be able to recognize both attacks that are already known and those that are unknown has led to the development and deployment of such systems. In this context, machine learning and deep learning paradigms are firmly established as the most powerful contenders in this field, particularly regarding intrusion detection systems (IDS). The most recent investigation corroborates the ability of learning-based IDS to surpass the known signatures and detect unusual activities that could either indicate a zero-day or an unseen attack [15]. Moreover, models powered by deep learning are significantly contributing to the field of event and log analysis where automation is necessary, since these models are capable of spotting rare and suspicious patterns in enormous log data [16] [17]. On the contrary, detecting accuracy should not be the only criterion for effective automation. The requirements for the practical utilization of models would involve not only the one that could be scaled but also the one that is durable enough to endure noise and drift, alongside giving the analysts an insightful guidance. Hence, recent researchers have begun focusing on the operation feasibility, latency needs, and the decision making of the intelligent systems in threat detection [18].

D. Role of Artificial Intelligence in Cyber security

Artificial Intelligence (AI) has become a major player in the cyber security field today through its ability to perform advanced pattern recognition, anomaly detection, and predictive analytics. The main benefit of machine learning algorithms is their capability to reveal intricate relationships among network traffic, system calls, and authentication events. So, they can identify the faintest signs that hint at the happening of malicious activities [19]. Deep Learning makes it possible to go even further than this because the model, i.e. the algorithm, is capable of learning the data representation in a hierarchical manner from either completely unprocessed or just literally slightly processed data. Convolutional Neural Networks (CNNs) have been employed to analysis of both traffic and malware, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models have been used to identify the temporal dependencies in the sequential patterns of attacks [20]. The recent reviews always point to the deep learning-based intrusion detection system (IDS) as the ones with the highest detection performance when compared to traditional methods, especially in complicated and high-dimensional situations [21]. Even though AI-based cybersecurity systems provide such benefits, they are still met with challenges of trust, transparency, and robustness. One drawback of black-box decision-making is that it may slow down incident response, whereas models

that are biased or have been altered by adversaries may result in wrong decisions. The NIST Artificial Intelligence Risk Management Framework (AI RMF) has proposed a governance model, continuous monitoring, and risk-aware deployment of AI technologies in safety- and security-critical sectors [22]. In addition, Explainable AI (XAI) has received more and more interest for being the key to building analyst confidence and comprehension in AI-powered security solutions [23] [24].

E. Research Motivation and Scope of the Review

Despite thorough investigations of AI-based cyber threat detection, the existing research has mostly remained divided into different fields, such as network intrusion detection, endpoint security, malware analysis, and cyber threat intelligence. Research in most cases is focused on specific datasets or individual indicators, which not only means limited transferability but also a poor connection with the real attack behavior. The review stems from the need for a systematic approach in the area of cyber threat detection. A threat schema translates structured patterns of harmful activities—action sequences, semantic relationships, or tactic-technique mappings—that disclose the core objective of an attack rather than just individual artifacts. When AI systems detect such schemas, their resistance to evasion, variation, and concept drift can be significantly improved, especially if they are associated with behavior-centric models like MITRE ATT&CK. Thus, this review covers artificial intelligence-based detection of threat schemas by focusing on the following aspects: (i) in what ways threat schemas are presented in contemporary literature, (ii) what AI methods are applied for schema detection and inference,

(iii) what datasets and evaluation techniques are applied in current studies, and (iv) what challenges are there such as robustness against adversarial attacks, Explainability, scalability, and deployment in a trustworthy manner.

II. Cyber Threat Landscape and Security Fundamentals

A. Types of Cyber Threats (Malware, Phishing, Ransomware, APTs, Zero-Day Attacks)

Malware is still considered as the primary threat category; however, its delivery and execution have undergone changes with the use of droppers, loaders and fileless techniques which are hard to distinguish from the normal system activity [25]. In the analysis of recent breaches, it was noted that attackers often used malware together with credential theft and exploitation in order to keep access and widen it, which has led to the decreasing effectiveness of purely signature-driven defenses against the latest variants and tools [26]. Phishing remains the top social engineering threat as it still aims at the human layer of security controls (identity, trust, and decision-making). Modern phishing is not restricted only to email; it has also expanded to include SMS-based phishing (smishing) and QR-based phishing (quishing) that are pointed out as common and futuristic tactics in Europol's IOCTA 2024 report [27]. The operational effect is even greater due to the existence of "phishing-as-a-service" which makes it easier for the criminals to enter and at the same time, it also increases the number of campaigns and their time reuse[27]. Ransomware has gone through a remarkable transformation from merely being a destructive force based on encryption

to being an extortion-driven complex process comprising multiple stages, where data theft and pressure tactics used along with encryption are meant to increase the chances of payment. The 2024 edition of Verizon's Data Breach Investigation Report (DBIR) indicates that the combination of ransomware and extortion techniques accounts for a significant proportion of the breaches that happen in different sectors, thus marking the transition from the wasting of malware incidents to the profitable intrusion campaigns. [25] The most recent threat intelligence reports have also pointed out that the initial access is very often the result of exploitation, phishing, and stolen credentials—thus, creating a flow of events that naturally leads to the deployment of ransomware after privilege escalation and lateral movement have occurred. [26] Academic studies of ransomware have in addition confirmed the widespread acceptance of “double extortion” (encrypt + exfiltrate) methods in the process and, thus, the necessity of developing detection techniques that are based on the modeling of the attacker's actions sequence rather than on the single indicators[28]. Advanced Persistent Threats (APTs) are considered to be the longest and most intentional intrusions and are often related to spying and strategic goals. The reporting done by Mandiant's M-Trends not only keeps an eye on various threat groups but also indicates that companies are being targeted by both old and new adversaries, which in turn suggests that the baseline of the defenses must be based on the constant development of the enemy's skills. [26] The necessity for this is to have such representations (schemas) of behavior that would still be relevant even if the tools were to change. Zero-day attacks (the exploitation of previously unknown vulnerabilities before a patch is released) are very impactful because, by definition, they circumvent all forms of protections that work on the basis of signatures or patches. The report DBIR 2024 illustrates how the exploitation of vulnerabilities and web-applications can very easily become the main entrance points for attackers, especially when high-risk vulnerabilities are disclosed, thus, showing the possibility of zero-days quickly transforming into large-scale compromises for the purpose of extortion and intrusion. [25] The ENISA's Threat Landscape Report has a similar standpoint as it highlights the main threats and trends based on large sets of reported incidents, thus, indicating that exploitation-driven intrusion still remains a strategic concern that is persistent and required to be addressed[29].

B. Cyber Kill Chain and Attack Lifecycle

A contemporary way to comprehend intrusions is to consider them as multi-stage lifecycles, during which each stage offers unique chances for detection. The Lockheed Martin Cyber Kill Chain illustrates a regulated intrusion scheme (from probing to acting on goals) and is often regarded as a conceptual model for positioning security measures across various points in an attack [30]. In modern-day scenarios, the cyberattacks are not always linear in nature, and the attackers could move back and forth through different stages, change their techniques, and even use more than one attack route at the same time depending on how well the target is protected. Mandiant's threat-intelligence reports reveal the most common entry points for infections (e.g.,

exploitation, phishing, prior compromise, and stolen credentials), which are related to the very first stages of the lifecycle (initial access and foothold establishment). [26] After getting access, usually, the next steps are to increase the attacker's privileges, stay hidden, and move around the network—these activities can be

effectively represented by behavior-centric modeling methods and structured knowledge frameworks which the security teams employ to visualize real-world tactics[31]. The kill-chain perspective gives an important design implication for the representation of the schemas in the review papers focusing on schema detection. The argument made here is that the schemas should depict not only isolated events but also their sequences and dependencies (e.g., “credential access → lateral movement → data access → exfiltration/extortion ”). When detection is synchronized with the lifecycle stages, the organizations will be able to cut off the attacks in their initial stage (pre-impact) instead of just identifying the final payload (post- impact) [30] [26].

C. Threat Actors and Attack Motivations

Cyber threats are executed by diverse actor classes whose motivations shape their methods and target selection:

- Cybercriminal groups often prioritize financial gain through ransomware/extortion, credential theft, and fraud. Breach reports show that extortion and ransomware are frequently tied to financially motivated actors exploiting web applications and identity weaknesses as entry points [25].
- State-linked or state-aligned groups commonly pursue espionage, strategic access, and influence operations. Threat intelligence reporting documents sustained activity from many tracked groups and emphasizes that defenders face both long-standing and newly tracked threat clusters across regions and sectors[26].
- Fraud ecosystems amplify social engineering. Europol’s IOCTA 2024 highlights phishing as a prevalent fraud vector and identifies smishing as particularly common (with quishing emerging), indicating that identity-centric attacks will remain a major driver of incidents[27].
- Opportunistic actors exploit exposed services, unpatched vulnerabilities, and credential reuse at scale—often using automated scanning and commoditized kits. ENISA’s threat landscape emphasizes prime threat categories and their recurring impact across incident reporting, reflecting the continued operational relevance of these actor behaviors[29].

One important aspect in AI-based schema detection is that motivation has a direct impact on the “shape” of the intrusion[25]. The cybercriminals driven by financial purposes usually choose the fastest way to get their money and thus the main impact and monetization (ransomware/extortion) would be the primary target of these criminals while the espionage hackers would take the opposite route and thus their main concern would be stealth, persistence, and selective access—which requires different schema characteristics, in turn (timing, sequence length, stealth indicators, and data access patterns)[26].

D. Challenges in Traditional Cybersecurity Mechanisms

Traditional mechanisms—like static signatures, fixed rules, and threshold-based anomaly alerts—are now unable to cope with the limitations imposed by the threat environment. • Evasion and rapid mutation: Malware and intrusion tooling changes quickly, while adversaries reuse objectives with different implementations[25].

This weakens signature-based detection and increases the maintenance burden for rule sets. Breach and threat intelligence reporting repeatedly indicates that

attackers adapt vectors and techniques, so static defenses lag behind active tradecraft [26]. • Concept drift and non-stationary behavior: Production environments are in a state of perpetual evolution (new applications, changes in user behavior, cloud migration) which in turn displaces the "normal" baselines. A recent extensive survey points out that concept and feature drift are still not given the necessary attention in IDS research, although drift has a direct impact on detection performance degradation and false positives increase in actual implementations[32]. • Scale and alert fatigue: The amount of telemetry and alerts that the analysts have to deal with is so high that they get overwhelmed, and the rule-based systems usually produce loud but contextually limited outputs[25]. This restriction of operations pushes the use of AI techniques that are capable of prioritizing, correlating, and summarizing multi- source signals into actionable schemas, [32].

- Adversarial pressure on detection logic: Present-day attackers intentionally test and change their tactics according to the defensive measures applied. The use of AI not only allows adversarial manipulation but also brings an extra layer of risk, as with the case of deliberately created inputs to the model that will result in misclassification, thus strengthening the requirement for solid modeling and assessment based on the attacker- aware premises[32].

the mentioned hardships support the transition to schema-based detection, whereby artificial intelligence acquires hierarchical representations of assaults (patterns, connections, and meanings) that still convey information when tools change, are affected by drift, and during multi-stage invasions[32] [26].

III. Concept of Cyber Threat Schema

A. Definition and Importance of Threat Schema

A cyber threat schema is an organized and machine-readable specification of the primary threat concepts (such as threat actor, malware, vulnerability, technique, indicator, infrastructure, campaign) including their interrelations, which allows for consistent representation, communication, and reasoning about cyber threat intelligence (CTI) [33]. In the ecosystems of operational CTI, schema thinking is most frequently realized through models of standard objects and relationships, with CTI items being exemplified as categorized entities linked by explicit connections. The standard STIX 2.1 is a well-known and widely used case in point: it has created a structured language for both cyber threat information and observables, making tool and organization interoperability possible [34]. Threat schemas are of paramount importance due to the diverse nature of CTI (alerts, logs, reports, IoCs, vulnerability feeds), and without a common representation, the correlation among sources turns out to be delicate and laborious. Regular Analysis in CTI and Semantic-Web Research emphasize Ontology/Schema Design as a Major Step for Creating Cybersecurity Knowledge Graphs (CKGs) that Will Be Able to Further Facilitate Querying, Fusion, and Automated Reasoning in SOC Workflows, thus Supporting Them [35] [36]. Moreover, schema- driven CTI facilitates the reuse and similarity of threat knowledge, thus allowing to interconnect the same threat concepts via different platforms and times while

keeping the context at the same time.

B. Behavioral, Structural, and Semantic Schemas

Threat schemas can be organized into three complementary forms, each serving different detection and intelligence needs: (a) TTP-Centric Behavioral Schemas:

These schemas illustrate the attack scenarios as a series of attacker moves, which are usually associated with tactics and techniques. The MITRE ATT&CK knowledge base is an outstanding example of behavioral models built on real-life observations that are extensively used for modeling and detection of threats [37]. ATT&CK also offers representations that can be read by machines (among them is STIX), thus enabling the integration of behavior knowledge into automated pipelines [7]. (b) Structural schemas (object-relationship CTI): These schemas treat CTI as typified objects and established relationships (e.g., indicator → malware → campaign

→ actor) at their core. STIX 2.1 not only standardizes this structural modeling but also offers object definitions and relationship patterns that support CTI sharing and linking consistently [33], [34]. Structural schemas can be utilized for evidence correlation across sources (e.g., connecting an observed IP/domain to infrastructure, malware families, and campaigns) and they do the correlation in a standardized manner. (c) Semantic schemas (ontology/knowledge graph): Semantic schemas encode CTI using ontologies (concepts + properties + constraints) which allows the systems to do reasoning (e.g., infer possible affected assets, link related incidents, and deduce technique-to-control mappings). The recent reviews on cyber security knowledge graphs highlight the importance of semantic modeling in the process of CTI entity unification and making the once machine reasoning over cyber knowledge possible [38] [36] [39].

Contemporary review articles consider these three types of schemas as being layered: the first layer shows the attackers' actions through the use of behavioral schemas, the second layer specifies the connections between the CTI elements by means of structural schemas, and finally, the third layer is formed by the semantic schemas which provide the ability to draw meaningful inferences across the interconnected CTI nodes [38][37] [39].

C. Schema-Based Threat Representation Models

Schema-based threat representation is realized through standards and formal models that provide structure to CTI for exchange and analytics: • STIX 2.1 (Structured Threat Information Expression): STIX 2.1 dictates an extensive CTI schema with objects (e.g., indicator, malware, threat-actor, relationship) and rules for linking them, thus allowing machine-to-machine CTI interchange and subsequent graph analytics[33], [34]. • ATT&CK in machine-readable form: ATT&CK is not merely a conceptual matrix; it is also distributed in STIX format for automated usage. MITRE explicitly states that STIX is the most detailed machine-readable representation of ATT&CK data, thus enabling automation and tooling integration [38]. This is significant for schema detection research as it facilitates the mapping of observed telemetry (logs/alerts/flows) into behavior-level schema elements (techniques, tactics) instead of just IoCs. • Cybersecurity Knowledge Graphs (CKGs) and ontologies: The

knowledge-graph- based cyber threat intelligence (CTI) systems are heavily dependent on a well- structured schema/ontology that delineates the existing entities and relations as well as their interconnections.

Extensive literature reviews reveal that CKGs not just support the representation of structured cyber knowledge but also enhance the analytics and automated reasoning processes [36] [39]. A recent review focusing on CTI has pointed out the same consistently as the semantic web technologies like ontologies and RDF/OWL are the ones that facilitate a common schema thus fostering better interoperability and machine reasoning across SOC members[35]. These models in combination classify the threat data into structured schemas thus making them querying capable (e.g., technique-centric queries), cross-source linking (correlation), and reasoning (inference) allowing the AI systems to spot threats as schema-consistent patterns instead of isolated alerts [35] [36].

D. Schema Evolution with Emerging Threats

Threat schemas are always subject to changes as the adversary and the infrastructure take place constantly. The behavioral knowledge bases like ATT&CK are updated to show the new attacker tactics (e.g., new credential access paths, new persistence mechanisms, and the new emerging environment such as cloud/API), and the automation of software applications through machine-readable releases allows the user to work with the latest behavior coverage [37]. The progression of Schema is also affected by the changes that occur in the observable data and the requirements for the sharing of CTI. The organizations that are moving towards the use of new telemetry sources (e.g., cloud audit logs, identity providers, and container platforms) are also giving importance to CTI standards and best-practice guidance that require consistent structuring of observables so that interoperability and quality are achieved [9]. At the same time, research surveys are indicating that cybersecurity knowledge graphs need to keep incorporating new entity types, relations, and context signals as the CTI ecosystem expands, otherwise, the schema will become incomplete and less useful for reasoning and analytics[38] [36]. The evolution of the schema more and more embraces quality and governance issues: the quality of the data must be ensured throughout the process of the schema update, besides, the validation of the data must be done automatically, and the same applies to unstructured text or multiple feeds when CTI is extracted from those sources [36]. As the effect increases in schema- driven systems, one of the main requirements for robust AI-enabled schema detection is to keep the alignment between changing standards (e.g., STIX objects/relationships) and changing behavior models (e.g., ATT&CK techniques/data sources) as a key one [33] [37].

IV. Traditional Approaches for Cyber Threat Detection

Conventional cyber hazard recognition strategies are usually classified as signature-based, rule/heuristics-based, and anomaly-based systems. No doubt, these methods constitute the backbone of several operational IDS/IPS and SOC monitoring processes, but they are not without major drawbacks in the case of today's rapidly changing threat scenario [40].

A. Signature-Based Detection Techniques

Signature-based (misuse-based) detection recognizes intrusions by comparing the recognized events (packets, payloads, system calls, logs) to a database of specified patterns (signatures).

This method is a common practice because of its high accuracy for known threats, rapid pattern matching, and moderate computing resource requirements, especially in matured installations [41]. Novel studies observe that signature-based intrusion detection system continues to be a robust measure for the former attacks and specific exploit patterns, especially in the limited situations where explainability and deterministic detection logic are most esteemed. Signature-based systems need regular manual updates and at the same time suffer through zero-day attacks, polymorphs, obfuscation, and all variations of attack which do not fit the stored signatures exactly. Nowadays, the evaluations are pointing out that the signature-based approaches can hardly be trusted any longer as the attackers continue to alter their tools at great speed while keeping the same intent and workflow, thus creating coverage gaps until the signatures are regularly changed [40].

B. Rule-Based and Heuristic Methods

The rule-based detection process is that suspect activities are recognized through the utilization of handcrafted or learned rules (IF-THEN conditions, thresholding logic, expert policies). In the past, many IDS solutions embraced this approach due to its characteristics of transparency in the decision-making process, behavior predictability, and easy compliance mapping (e.g., “raise an alert if the number of failed logins exceeds X during Y minutes”) [42]. An up-to-date thorough IDS overview points out that the old IDS systems that used to be driven by rules were mainly using predefined patterns and manual updates, thus making it very hard to maintain them properly at large. Heuristic techniques go beyond pure rules by introducing domain-specific indicators (e.g., scoring strange port sequences, rare protocol behavior, or unusual process trees) [43]. In the same way, research on rule-learning-based IDS indicates that using symbolic rule induction for the interpretation of results is much more superior than that of black-box models and that it can also aid “human-in-the-loop” security operations, especially in critical infrastructures where the need for explainability is very high. In spite of these advantages, rule/heuristic systems mostly have rigid logic as their main obstacle and they need long and thorough tuning to be able to cut down false alerts across different environments [40]. Furthermore, they are not able to generalize even when the attacker behavior slightly changes to the point of being out of the encoded conditions, which makes them even weaker against adaptive adversaries.

C. Anomaly-Based Detection Systems

Anomaly-based detection models identify “normal” behavior and mark deviations as potential threats. In this group, there are methods like statistical profiling, clustering-based outlier detection, and more recently, ML/DL-based anomaly detection. Recent surveys highlight the fact that anomaly-based IDS is especially useful for detecting unknown and new attacks, because it does not need explicit signatures for every single threat[44]. In contemporary anomaly detection pipelines, feature learning and temporal modeling are regularly

employed to catch the dynamics of traffic and events, particularly in complicated scenarios like IoT and SDN. A survey of the IoT IDS for the year 2025 sorts the present anomaly-based techniques and analyzes the design and performance factors in IoT networks rather than performing simply anomalous detection, which is still the dominant

mechanism[21]. Presently, the situation is similar to the debate over Software Defined Networking (SDN) where the deep-learning-based anomaly detection systems were absolutely necessary and preprocessing, modeling, and evaluation areas were the ones that critical were mentioned. These points have definitely emphasized the need for anomaly detection in the context of the changing network behavior and the high-level techniques of the intruders [45]. Among the top operational challenges is the fact that the anomaly-based systems might generate a huge number of false alarms, particularly when the baseline is moving, there is mixed traffic, or the training coverage is insufficient. The issue of “precision” is very often referred to as the main ground for not adopting such systems unless they are supported by good calibration, continuous learning, and contextually correlated activities all taking place simultaneously [46].

D. Limitations of Conventional Detection Methods

Traditional approaches face recurring limitations that motivate AI-driven schema detection:

1. Limited generalization to unseen threats: Signature and rigid rule systems perform poorly against zero-day and evolving attacks because they depend on known patterns and manual updates.
2. Alert noise and operational overhead: Rule/heuristic thresholds often require extensive tuning and still produce noisy alerts across diverse environments, increasing analyst workload.
3. Concept drift and feature drift: Even anomaly-based systems degrade when “normal” behavior changes over time. A 2024 survey specifically highlights that IDS faces multiple layers of concept and feature drift, and that drift-aware integration is still insufficiently addressed in many IDS pipelines [32].
4. Scalability constraints: High-volume telemetry (cloud logs, endpoint events, network flows) makes purely manual rule management and static profiling increasingly impractical [47].
5. Adversary adaptation: Attackers deliberately change surface indicators while preserving the underlying attack workflow, which breaks detection logic that is overly tied to static indicators instead of behavior patterns.

Implication for schema-based detection: The indicated restrictions are directly connected to the unpredictability of various patterns and the changing character of the behavior, thus, the direction in contemporary detection research is shifting from the reliance on separate Indicators of Compromise (IoCs) to the employment of organized behavioral patterns (schemas) that depict multi-stage attacks in a manner that they can be utilized through transformations of tools and partial visibility.

V. Artificial Intelligence in Cyber Threat Detection

With the rapid advancement of information technology, networks have become an essential and significant component of human civilization [48]. Perspectives and opinions change when it comes to network security issues, and so do the security concerns of networks proliferation. Traditional network security measures

such as rule-based intrusion detection systems (IDS) and firewalls, despite their previous success, have not been able to cope with the complexity and sophistication of modern network attacks [49]. Most of the time, the offenders resort to very advanced technical means in order to outsmart the usual detection systems and this leads to huge data leaks, money losses, and more serious risks to the national security. By virtue of its data analysis, pattern recognition, and automation capabilities, AI technology has become a key player in the fight against cyber threats in this context [50]. It is possible to better spot the suspicious actions, to uncover the new threats and to monitor the attack patterns in real-time by utilizing the technologies such as machine learning and deep learning in the threat detection systems for cyber security. That not only increases the accuracy of threat detection but also reduces the false alarm rate to a large extent [51]. An Overview of The Development of Artificial Intelligence Technologies in Cyber security Artificial intellect (AI), a technology capable of emulating human intellect, has been extensively used across several domains in recent years[52]. The core of the subject lies in robust technologies like machine learning, deep learning, and natural language processing, which are capable of pattern recognition, future trend prediction, and automated decision making through massive data analysis. The deployment of AI technology in cyber security is, however, a significant benefit because of its ability to manage large and complex data and to do real-time threat detection [53]. With the continuous improvement of processing power and algorithms, the AI technology has been gradually moving from academia to industry as the main tool in cyber security, and this is very much the case in cyber security. With each step forward in information technology, the kinds and sophistication of cyberspace security threats are increasing. Signature libraries that identify traditional dangers such as viruses and worms are still effective in their way; however, the current APTs (Advanced Persistent Threats), zero-day attacks, and social engineering attacks are more hidden and thus harder to spot [54]. Attackers are turning very much skilled, making use of encrypted messages, passing through regular barriers, and taking advantage of zero-day vulnerabilities to carry out their attacks, therefore, they are challenging the existing cybersecurity defense systems quite a lot. These changes have made the cybersecurity industry to keep on looking for smarter and more automated solutions all the time [55]. The ever-increasing sophistication of cyber-attacks is countered by the gradual incorporation of AI technologies into cyber security for improved threat detection and response. In the beginning, the AI tool was predominantly employed in areas such as detecting viruses, filtering spam, and intrusion detection systems (IDS) where it used machine learning algorithms to automatically classify and recognize unusual behaviors. The initial usage of AI in cyber security has shown that the technology can very well get rid of the limitations of the traditional security measures, thus, improving the detection accuracy as well as the response time [32]. At now, AI

technology has infiltrated every aspect of cyber security and helped in the process of collecting and analyzing threat information, as well as real-time threat detection, response, and recovery. AI-enabled security systems can, for instance, carry out the analysis of huge amounts of security data automatically, detect and predict possible attacks instantly and give smart Responses, all at the same time. Deep learning models are capable of detecting abnormal patterns by analyzing very complicated network traffic data, while reinforcement learning algorithms can, on the other hand, optimize security policies and even change defensive measures dynamically based on current conditions. The hotness of AI technology is still very controversial in the future, but the cyber security field is definitely considering and even trying out new application scenarios along with the AI technology evolution to deal with the security threat always present.

VI. Core Approaches to Artificial Intelligence Technology in Threat Detection

In cyber security threat detection, AI technology significantly enhances the precision and efficacy of detection via its robust data analysis and pattern recognition skills. The primary methodologies include machine learning, deep learning, integrated learning, and multimodal approaches [56]. Every one of these techniques has its own unique benefits: machine learning is very good at recognizing features of threats from large labelled data sets, deep learning is very good at unraveling complex unstructured data through multi-layer neural networks, and multiple methods along with multimodal approaches provide greater and more flexible detection by combining different sources of data and algorithms. The following parts will explain these three basic methods in more details. Binary Classification Loss (Cross-Entropy Loss):

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

A. Machine Learning-Based Threat Schema Detection

Supervised learning is a dominant method in machine learning that trained models using labeled data for the recognition of already existing threat classes in new data. In cyber security, the technology of supervised learning is very much developed and finds its application in the areas of malware detection, spam filtering, and intrusion detection. The model is trained on a very large dataset of labeled normal and anomalous instances, so it will be able to learn about and recognize the characteristics of evil activities and thereby efficiently tell apart the usual traffic and the threats during the real-time detection [57]. This method comes with very high precision and transparency; however, it is still dependent on the quality and the volume of labeled data, and it could possibly face

a challenge when it comes to new or unknown threats. Unsupervised learning is a non-supervised process that does not require labelled datasets, but it examines the inherent structure of the data to uncover patterns and detect abnormal behaviors. It is mainly applied to anomaly and intrusion detection in cybersecurity which is highly efficient for detecting threats that are not defined beforehand. The most common unsupervised learning techniques are cluster analysis and anomaly detection methods like K-means, isolation forests, and autoencoders [34].

The application of these techniques can detect abnormal behaviours that are not consistent with usual ones, thereby revealing potential security weaknesses. The very power of unsupervised learning is its ability to bring to light hidden dangers; however, it is affected by a high rate of false positives, which makes it necessary to combine it with other methods for refining the outcome. Reinforcement learning is a dynamic decision-making and adapting machine learning method that operates through a constant interaction with the environment in situations where continual learning and adaptation to environmental changes are required. In the case of network security threat detection, reinforcement learning is in an unending cycle of improving the existing detection process and, thereby, the security measures that are effectiveness through its interaction with the environment. Reinforcement learning could be the one that would allow the intrusion detection system to change its alarm levels according to the threat or be the one that would adjust the defense strategy in relation to the new attack methodologies. Reinforcement learning, in contrast to traditional strategies, has the advantage of having a lot of flexibility and can, at the same time, alter the detection algorithms based on the input of new threat data, thus providing the network security that is both flexible and real-time as portrayed in figure 1:

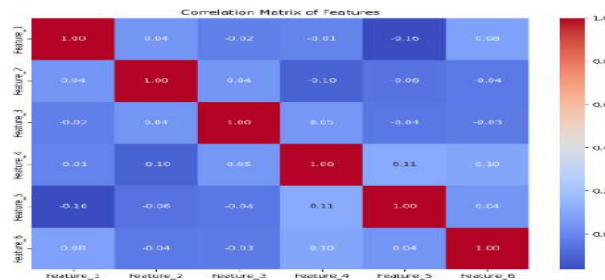


Figure 1 Correlation Matrix of Features

Transfer learning is a technique that employs existing knowledge to tackle new challenges, particularly in situations where data is scarce or the training costs are very high. In the area of cybersecurity, transfer learning can fast improve the performance of detection models by the application of threat detection knowledge from one domain (e.g., financial security) in another (e.g., healthcare security) [58]. Migration learning facilitates the development of effective threat detection models in scenarios with little data, while minimising training duration and computing resource expenditure. The issue of migrating learning resides in the need to meticulously choose the similarity between the source and destination domains to guarantee the efficacy of knowledge transfer. Gradient Descent Update Formula:

$$\theta_{t+1} = \theta_t - \alpha \nabla_{\theta} J(\theta)$$

B. Deep Learning-Driven Cyber Threat Schemas

Deep learning can automatically identify and master the features of complex data through a multi-layer neural network architecture, thus hugely contributing to the area of cybersecurity threat detection.

Convolutional Neural Networks (CNNs) are commonly employed in the processing of images and traffic data and can effectively detect characteristics like harmful code or unusual network traffic. On the other hand, Recurrent Neural Networks (RNN) and then its improved version, Long Short-Term Memory Networks (LSTM), are the best ones for time-series data and thus suitable for areas like log analysis of network activity and pinpointing persistent threats [59]. Deep learning applied to huge datasets enables these neural network models to detect risks that are hardly noticeable to traditional methods, thereby improving not only the accuracy of detection but also the speed of response. An auto-encoder is an unsupervised learning framework that learns a latent representation of the input data through the steps of compressing and reconstructing the data. One of the application areas of autoencoders in cybersecurity is anomaly detection, where the model can help with the detection of unusual network traffic or different user behavior patterns [60]. The reconstruction error rises significantly and indicates possible threats when the input data varies largely from the usual data patterns. Generative Adversarial Networks (GANs), consisting of a generator and a discriminator, produce synthetic data that resemble real data which in turn, through adversarial training, increase the effectiveness of the detection model. Generative Adversarial Networks (GANs) can be employed within the field of cybersecurity to create attack samples that reinforce the detection models' robustness, as well as to trick the detection system for the purpose of adversarial testing aimed at evaluating and improving the system's security capacity. Deep learning automatically extracts and learns features from complex data using multi-

layer neural network architecture, offering a significant advantage in cybersecurity threat detection[61]. Convolutional Neural Networks (CNNs) are widely used in the analysis of images and traffic data, and they can very accurately detect characteristics such as malware or network anomalies. Recurrent Neural Networks (RNN) with its more advanced version, Long Short-Term Memory Networks (LSTM), are skilled in time-series data and thus can be used for monitoring network behavior logs and identifying advanced persistent threats, etc. The application of deep learning methods on the large datasets gives the neural network models the ability to discover the threats that are too difficult for the conventional methods to uncover thus enhancing the detection accuracy and shortening the response time. An auto-encoder, which is a type of unsupervised learning model, derives the latent representation of the data by compressing and reconstructing the input data. Autoencoders frequently play a significant role in anomaly detection within the realm of cybersecurity, comprising detection of unusual network traffic or diverse user behavior patterns as part of their models. The reconstruction error is used as a risk signal that the input data is considerably different from the normal ones and hence, it raises a lot [62]. Generative Adversarial Networks (GANs) are composed of two models—a generator and a discriminator—that work together to produce imitations of real data, thereby improving the detection model's performance through adversarial training. GANs can be applied in the field of cybersecurity to generate the necessary attack samples that will fortify detection models and trick the detection system for adversarial testing, which in turn, evaluates and improves the system's defense capabilities. Despite deep learning algorithms usually requiring large amounts of data and a computer's training power, transfer learning can quickly create cost-effective threat detection models with a small dataset by utilizing pre-trained models. Pre-trained models are generally made with very large datasets and then employed for specific cybersecurity applications, such as malware detection and abnormal traffic classification [29]. Transfer learning enables deep learning models to absorb the knowledge of already trained models, quickly adapt to new threat scenarios, and improve detection results. This approach saves on training time and costs and at the same time improves the model's generalization power enabling it to cope with different security threats. A Graph Neural Network (GNN) represents one of the deep learning approaches that focus specifically on the graph-structured data or, say, networks of nodes and edges. Hence, the GNN is the best tool for assessing network topology and identifying the whole threat node. In detail, GNNs can talk about intricate hierarchical data like social and computer networks, and by clarifying the relationships between the nodes, they could practically tell where the danger is coming from. For instance, GNNs could point to the deviant communication behaviors within a network and able to silently show the attacker as well or even the way through which the attack could be carried on. Contrary to the traditional methods, GNN can, amongst others,

trace relations through complex network topologies better and, thus, more reliably identify possible threats; hence, it can be considered a source of new ideas and tools for the improvement of network security.

C. Integrated Learning and Multimodal Approaches

Integrated learning enhances overall detection accuracy and resilience by amalgamating the predictive outcomes of many models. In cybersecurity threat detection, integrated learning may amalgamate the benefits of many algorithms, including decision trees, random forests, and gradient boosting trees, to enhance the identification of intricate threats[40]. The integrated learning method can not only reduce a model's false alarm rate but also improve its response to threats of different types by utilising the models' complementarity to a great extent. For the case of malware detection, integrated learning can also merge a number of feature extraction techniques and classifiers for the purpose of detection accuracy and generalisation capability. In the network security threat detection area, the data used is typically multimodal and consists of network traffic data, system logs, user behaviour, geolocation and so on [63]. By merging information from various data sources, multimodal methods ensure a more complete understanding and detection of potential risks. The combination of network traffic data and user behaviour data leads to a more accurate detection of irregular behaviour and possible attack patterns. The employment of multimodal data fusion techniques in detection systems enhances the perceptual and decision-making skills drastically by interdisciplinarily analyzing the input through various modalities that, therefore, increase the overall and multidimensional understanding of complicated threats and the efficiency of threat detection [64]. Heterogeneous integration techniques are the ones to surpass the limits of integration learning and merge different types of models together with various data sources to obtain a huge improvement in threat detection. Such approaches, unlike the traditional isomorphic integration methods which handle the combination of several identical models, are able to process the data much more efficiently and quickly, thus taking advantage of different models' strengths, e.g. that of the deep learning and the classical machine learning models in the case of unknown threats and intricate attacks, especially [65]. To show how this can be done, a heterogeneous integration method of deep learning models for extracting features and then using their results as inputs into ensemble learning models for final decision making is that it can significantly precise threat detection by combining the best of two worlds, i.e. deep learning's powerful feature representation and co-learning's effective decision-making. This is illustrated in Figure 2:

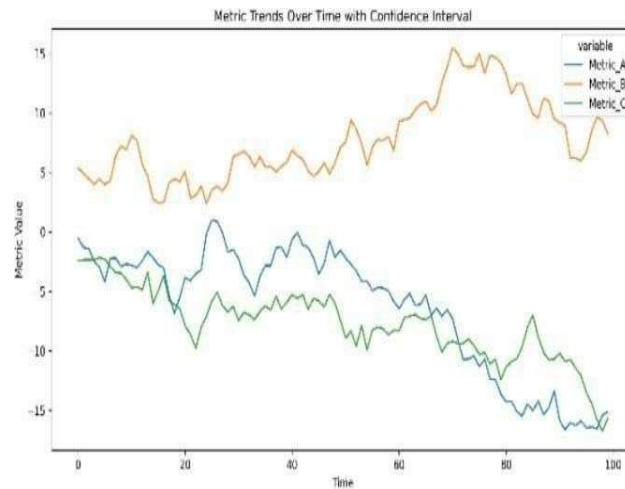


Figure 2 Metric Trends over Time with Confidence Interval

By means of a simultaneous investigation of the various data types, multimodal deep learning can cope with complex cybersecurity risks more efficiently. In real-life situations, multimodal deep learning models might combine different types of data such as text, images, audio, and video for in-depth risk assessment and detection [66]. In APT (Advanced Persistent Threat) Identification, multimodal deep learning fuses together the network data, log analysis, and user behaviour to offer a complete view and spot the long-lasting threats. One of the advantages of multimodal deep learning is combining different data sources' giving a total view of the threats, thus improving the accuracy and robustness of detection, especially to complex attacks and multidimensional data.

VII. Challenges And Future Developments of Artificial Intelligence in Cybersecurity Threat Detection

Utilising AI for cybersecurity threat detection presents significant problems regarding data privacy and security. The process of threat detection relies heavily on large amounts of data consisting of network traffic, user behavior, system logs, and possibly even sensitive information. Data protection against leaks or misuse during the model's training and deployment stages is a critical problem that needs solving [67]. At present, privacy-preserving techniques like differential privacy and federated learning have started to be applied in an incremental manner in the realm of cybersecurity. However, the problem of maintaining data confidentiality without sacrificing the accuracy and effectiveness of the detection models has not been fully resolved yet and thus requires further research. AI systems are very effective in

identifying the risks, however, at the same time, they can be attacked from the side. The attacker might use the technique of generating aggressive samples or applying different methods to get the decision in his favor, and as a result, the wrong ratings will be given the security system will be penetrated. Such types of attacks against the systems can reduce their capability to detect but at the same time, their power to defeat the security wire unnoticed can be increased [68]. Thus, the durability of AI systems is significantly improved. Researchers are considering different ways to make threat detection systems more reliable through the hardening of the system against attacks, the establishment of safer designs, and the creation of protective measures. One of the approaches under consideration is Anomaly Detection with Autoencoder (Reconstruction Error):

$$\text{Reconstruction Error} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|^2$$

Artificial intelligence models, particularly deep learning models, are often seen as "black boxes" As a result, the resilience of AI models is being greatly enhanced. By fortifying the system against attacks, creating safer designs, and developing safeguards, researchers are looking at various approaches to increase the reliability of threat detection systems. One of the methods being explored is Anomaly Detection with Autoencoder [58] [69]. The enhancement of model performance along with the interpretation improvement to comply with different security and privacy requirements is one of the major issues that will be faced in the future. The gradual shift towards Explainable Artificial Intelligence (XAI) and model visualization tools in assisting security experts to understand and trust the detection results is happening. AI application in cybersecurity threat detection is quite promising, and with technology advances, threat detection systems will become more intelligent and automated thus allowing them to take on new and complex threats with faster and more accurate reactions. The combination of different fields of study will lead to the creation of new applications, including threat detection systems that would combine artificial intelligence, quantum computing, and blockchain technology. In addition, the use of AI in cybersecurity will increase the aforementioned three factors (safety, reliability, and efficiency) through gradually developed countermeasure strategies, privacy protecting systems, and interpretable models [70]. In the future, AI will not only be a tool for threat identification but also will be an essential part of the cybersecurity strategy. This will uplift the entire industry to a higher level of security through the adoption of the strongest technologies. AI technology has been applied to cybersecurity threat detection and its use has opened up new areas for the enhancement of detection speed and accuracy. AI could evolve the security measures in place by leveraging different techniques such as machine learning, deep learning, and hybrid learning to deal with the more sophisticated attacks, rapidly detecting the possible threats, and providing top-

notch security solutions. Nevertheless, the incorporation of AI on a large scale is also a source of concerns like data privacy, security risks, and model interpretability, thus making it necessary to deal with security and compliance issues along with technical developments. The AI's contribution to cybersecurity is projected to be more significant in the future [71]. Due to the continuous development of technology and the increasing interaction between different areas, the threat detection systems will be transformed into being more intelligent, automated and user-friendly making the protection of the complex and unstable cyberspace more effective. In order to reach this goal, it is necessary to continuously monitor the new technologies, improve the existing ones, and conduct extensive research on AI applications to ensure its continuous development and effective use in cybersecurity [26]. By continuous efforts, AI will become the main part of the network security protection system and will thus lay a strong foundation for a more secure and trustworthy network environment.

VIII. Open Research Challenges and Limitations

The present day open research challenges and limitations in the cyber-negative detection (and threat schemata) field are mainly characterized by three aspects: realism, robustness, and operationalization gaps. First of all, the great majority of the published results are still grounded on public, static, and partly outdated datasets of NIDS where traffic composition, encryption prevalence, and attack diversity are not perfectly reflecting the conditions of modern enterprise/cloud areas, and hence, the performance is measured offline too optimistically, which very often does not transfer to production; recent systematic reviews of NIDS datasets and data-driven IDS surveys indicate very clearly that the dataset property gaps, labeling inconsistencies, and benchmarking limitations are the main obstacles for fair comparison and reproducibility across various studies [72] [73]. Then there are the issues of concept drift and feature drift that persist as major barriers because "normal" behavior keeps changing with software updates, introduction of new applications, migration to the cloud, and variations in usage patterns depending on the season, while attackers are making deliberate efforts to change their tactics; drift-aware IDS surveys indicate that such research is still lacking in comparison to the practical importance of drift and they advocate for constant monitoring, adaptive learning strategies, and evaluation protocols that explicitly test for temporal degradation rather than just measuring single snapshot accuracy [32]. Firstly, mostly due to the adoption of ML/DL techniques, the threat of adversarial machine learning (AML) is made even more severe: The various forms of attacks such as evasion and poisoning can distort the network features or the training data to get the system to classify them incorrectly. Recent surveys on AML in the context of IDS show that many IDS do not test their robustness under attacker-aware

conditions (for instance, adaptive black-box probing), thus creating a significant disconnect between laboratory metrics and the reality of adversarial attacks [74]. Secondly, a demand for explainability and the human factor has not yet been fully met: the SOC has to deal with alerts that are easy to interpret, stable when there is a change, and supports the incident response, however, a large part of the deep learning models is black-box and may offer fragile explanations; the latest extensive reviews on XAI in cybersecurity have stated that explainability should be considered as a key design constraint along with performance and robustness rather than being an afterthought [75]. Thirdly, pipelines based on threat-schema and CTI face the problem of knowledge quality and standardization: systems driven by schemas rely on consistent entities/relations (for example, STIX objects and links), but CTI is diverse and murky, and studies in the field of knowledge graphs have often stated that ontology/schema design, error propagation in entity-relation extraction, and inconsistencies in fusion can all lead to the degradation of reasoning and detection reliability in downstream applications [38] Cybersecurity knowledge Cybersecurity knowledge graphs [35] . At last, the governance limitations—privacy, responsibility, and lifecycle risk—are becoming more and more important as AI-based detection is applied in critical areas. NIST's AI Risk Management Framework acknowledges the necessity for permanent evaluation and supervision of trustworthiness characteristics (validity, reliability, robustness, and transparency) throughout the AI life cycle, which leads to the assertion that upcoming research has to include risk-managed deployment, monitoring, and auditing in addition to just model-building[76].

IX. Future Research Directions

The primary focus of AI-based cyber threat and schema detection should be on developing detection systems that are adaptive, trustworthy, and ready for deployment rather than just improving offline accuracy. Being aware of the drift in data and continual learning is definitely one of the great advantages because through it, the model's behavior can be altered according to the network and the attacker's tactics without needing to go through the whole retraining process constantly. Another essential method is privacy-preserving and federated learning, which allows organizations to share their intelligence while keeping their data secure. In addition, as attackers get more advanced and try to either exploit or nullify the ML-based detectors, research should also focus on adversarial robustness. Moreover, XAI or explainable AI is going to be a great help in this whole process since it will provide comprehensible and helpful insights that can lead to greater analyst trust and practical operational use. Ultimately, the AI models should be integrated with the common threat schemas and CTI frameworks (such as STIX

and ATT&CK) more deeply so that in the cybersecurity detection, sharing, reasoning, and governing can be done properly in real-life situations.

X. Conclusion

The review concludes that the present-day cyber-threat landscape, which is stimulated by the growth of the cloud, IoT, multi-stage infiltrations, and the rapid change of attack methods, has rendered the traditional detection methods ineffective. Signature-based and rule-based systems that fail to recognize zero-day attacks, trends, and massive alert management are among those that experience this difficulty. The research presents cyber threat schema as a common framework that organizes threat knowledge via behavioral (TTP-driven), structural (CTI object-relationship), and semantic (ontology/knowledge-graph) representations, resulting in more coherent correlation, reasoning, and dissemination of threat intelligence across various data sources. It also shows how AI and deep learning have advanced the detection of sophisticated and new threats by representation learning and sequence modeling, while at the same time acknowledging fundamental limitations such as gaps in dataset realism, concept and feature drift, adversarial machine learning vulnerabilities, and requirements for explainability and governance. The review indicates that future progress will depend on methods that are drift-aware and can learn continuously, privacy-preserving collaboration (e.g., federated learning), adversarial robustness, and improved compliance with Cyber Threat Intelligence (CTI) standards like STIX and behavioral frameworks such as ATT&CK to ensure that schema-based detections are accurate, explainable, interoperable, and can be deployed operationally in cybersecurity practices.

References

- [1] W. Ahmad, A. Rasool, A. R. Javed, T. Baker, en Z. Jalil, “Cyber security in iot-based cloud computing: A comprehensive survey”, *Electronics*, vol 11, no 1, bl 16, 2021.
- [2] E. Bertino en N. Islam, “Botnets and internet of things security”, *Computer (Long. Beach. Calif.)*, vol 50, no 2, bll 76–79, 2017.
- [3] C. J. R. Molloy, “Adversarial Learning for Cyber Threat Intelligence: an Attention on Malware”. Queen’s University (Canada), 2024.
- [4] G. Assenza, L. Faramondi, G. Oliva, en R. Setola, “Cyber threats for operational technologies”, *Int. J. Syst. Syst. Eng.*, vol 10, no 2, bll 128–142, 2020.
- [5] S. Peltola, “Threat Detection Analysis Using MITRE ATT&CK Framework”, 2025.

- [6] A. Sengupta en V. Ramteke, “Leveraging Next-Generation Business Intelligence for Proactive Cybersecurity Defense”, in 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT), IEEE, 2025, bll 414–419.
- [7] A. L. Buczak en E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection”, IEEE Commun. Surv. tutorials, vol 18, no 2, bll 1153–1176, 2015.
- [8] K. Scarfone en P. Mell, “Guide to intrusion detection and prevention systems (idps)”, NIST Spec. Publ., vol 800, no 2007, bl 94, 2007.
- [9] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, en A. Bouchachia, “A survey on concept drift adaptation”, ACM Comput. Surv., vol 46, no 4, bll 1–37, 2014.
- [10] H. Hindy et al., “A taxonomy of network threats and the effect of current datasets on intrusion detection systems”, arXiv Prepr. arXiv1806.03517, 2018.
- [11] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, en A. Bouchachia, “A survey on concept drift adaptation”, ACM Comput. Surv., vol 46, no 4, bll 1–37, 2014.
- [12] M. A. Ferrag, L. A. Maglaras, H. Janicke, J. Jiang, en L. Shu, “Authentication protocols for internet of things: a comprehensive survey”, Secur. Commun. Networks, vol 2017, no 1, bl 6562953, 2017.
- [13] B. Biggio en F. Roli, “Wild patterns: Ten years after the rise of adversarial machine learning”, in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018, bll 2154–2156.
- [14] M. IBM Security, “Cost of a data breach report 2021”, 2023.
- [15] N. Moustafa, B. Turnbull, en K.-K. R. Choo, “An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things”, IEEE Internet Things J., vol 6, no 3, bll 4815–4830, 2018.
- [16] P. Himler, M. Landauer, F. Skopik, en M. Wurzenberger, “Towards detecting anomalies in log-event sequences with deep learning: Open research challenges”, in Proceedings of the 2023 European interdisciplinary cybersecurity conference, 2023, bll 71–77.
- [17] H. Liu en B. Lang, “Machine learning and deep learning methods for intrusion detection systems: A survey”, Appl. Sci., vol 9, no 20, bl 4396, 2019.

- [18] H. Liu en B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey", *Appl. Sci.*, vol 9, no 20, bl 4396, 2019.
- [19] Y. LeCun, Y. Bengio, en G. Hinton, "Deep learning", *Nature*, vol 521, no 7553, bl 436–444, 2015.
- [20] J. Kim, J. Kim, H. L. T. Thu, en H. Kim, "Long short term memory recurrent neural network classifier for intrusion detection", in 2016 international conference on platform technology and service (PlatCon), IEEE, 2016, bl 1–5.
- [21] A. Khraisat, I. Gondal, P. Vamplew, en J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges", *Cybersecurity*, vol 2, no 1, bl 1–22, 2019.
- [22] N. AI, "Artificial intelligence risk management framework (AI RMF 1.0)", URL <https://nvlpubs.nist.gov/nistpubs/ai/nist.ai>, bl 100–101, 2023.
- [23] F. Charmet et al., "Explainable artificial intelligence for cybersecurity: a literature survey", *Ann. Telecommun.*, vol 77, no 11, bl 789–812, 2022.
- [24] M. Dixit, I. Kansal, V. Khullar, R. Kumar, S. Kumar, "Analyzing Trustworthiness and Explainability in Artificial Intelligence: A Comprehensive Review", *Recent Adv. Electr. Electron. Eng.*, vol 18, no 8, bl 1107–1135, 2025.
- [25] Verizon, "2024 Data Breach Investigations Report", 2024.
- [26] G. C. Security, "2024 Special Report", 2024
- [27] Europol, "Internet Organised Crime Threat Assessment (IOCTA) 2024", Publ. Off. Eur. Union, Luxemb., 2024.
- [28] E. Kritika, "A comprehensive literature review on ransomware detection using deep learning", *Cyber Secur. Appl.*, vol 3, bl 100078, 2025, doi: <https://doi.org/10.1016/j.csa.2024.100078>.
- [29] Lella, ENISA THREAT LANDSCAPE 2024, no SEPTEMBER. 2024.
- [30] L. Martin, "Cyber Kill Chain", 2020.
- [31] E. U. Agency en L. E. Cooperation, Internet Organised Crime Threat Assessment (IOCTA) 2024. doi: 10.2813/442713.
- [32] M. A. Shyaa, N. F. Ibrahim, Z. Zainol, R. Abdullah, M. Anbar, en L. Alzubaidi, "Evolving cybersecurity frontiers: A comprehensive survey on concept drift and feature dynamics aware machine and deep learning in intrusion detection systems", *Eng. Appl. Artif. Intell.*, vol 137, bl 109143, 2024, doi: <https://doi.org/10.1016/j.engappai.2024.109143>.