Mesopotamian journal of Cybersecurity
Vol.5, No.3, **pp**. 1184–1198
DOI: https://doi.org/10.58496/MJCS/2025/063; ISSN: 2958-6542
https://mesopotamian.press/journals/index.php/cybersecurity

Research Article

Mitigating Zero-Day Vulnerabilities in IIoT Systems: Challenges and Advances in AI-Powered Intrusion Detection Systems

Khalid Asaad Hashim[1], Yusnani Binti Mohd Yussoff[1,*], Shahrani Binti Shahbudin[1]

[1] *Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia*

ARTICLEINFO

Article History

Received 12 Dec 2024
Revised 15 Jul 2025
Accepted 28 Sep 2025
Published 15 Oct 2025

Keywords

Intrusion Detection Systems
Zero-day attacks
False Negative Rate
IIoT Security

ABSTRACT

Zero-day attacks are one of the great challenges that intrusion detection systems have been facing and keep on facing today, especially worsening within Industrial Internet of Things environments since their ability to take advantage of unknown vulnerabilities results in a high rate of false negatives. It is under this framework that this paper presents a set of experiments that have been carried out with the objective of analyzing the consequences of zero-day attacks with regard to performance degradation in Intrusion Detection System (IDS) and, secondly, and with greater emphasis, about those failings which have been pointed out up to now as those affecting precision in detection. This has been done through the systematic review of 200 research papers published from the years 2023 to 2024, further categorized into the four main focus areas: general AI-based IDS, Machine learning (ML)-based IDS, Deep Learning (DL)-based IDS, and Deep Reinforcement Learning (DRL)-based IDS. Accordingly, 45% were DL-based IDS reviews; 35% related to machine learning; 15% consisted of the ones about DRL-based ones, while 5% pertain to the General AI-based ones. Results show that the approaches with DL-based systems will come up with extensive promises, reducing the impact brought by false negatives, besides extending the issues even when one considers a background of adversarial attack issues. It underlines that, in IDS, apart from accuracy, detection specificity and recall are also of essence for dealing with low frequent but high-impact zero-day threats. These techniques further make the following proposal: the use of both machine learning and deep learning techniques should be improved in enhancing the performance of IDS.

1. INTRODUCTION

The Industrial Internet of Things (IIoT) started a revolution that seamlessly connects and automates most sectors, from manufacturing to healthcare. At the same time, however, the IIoT presents several challenges related to cybersecurity, especially zero-day attacks. Zero-day attacks take advantage of previously unknown vulnerabilities, making traditional security measures such as Intrusion Detection System (IDS) increasingly inadequate. Most of these new threats can't be picked up by IDS in real time and that results in very high false negative rates (FNR) when malicious activities pass unnoticed. Zero-day attacks are especially devastating in IIoT environments, where a breach can result in the compromise of critical infrastructure and put at risk both data integrity and operational safety. Since IIoT networks are dynamic and distributed, this has made the detection of such attacks more challenging compared to traditional Information Technology (IT) systems. Most of the existing IDS frameworks are signature-based and designed to detect either known attack signatures or deviations from predefined norms, and hence, they cannot handle the variability and complexity of IIoT traffic, leading to a high number of false negatives. The high-end complexity involved with zero-day attacks and also the critical situation of IoT environments impose greater challenges. More advanced methods are necessitated. The last couple of years' research directions put more significance into incorporating IDS frameworks with AI and ML for developing better traditional approaches in detection methods. These intelligent systems have the capability for learning from large datasets, adapting to ever-evolving threats, and recognizing patterns that might otherwise go unnoticed using signature-based approaches.

Prior to more details, some important terminologies that are widely used are presented in this paragraph. The false negative rate (FNR) signifies the proportion of bad things that happen and go undetected by an Intrusion Detection System (IDS), thereby being misidentified as good things. This is a particularly alarming problem in Industrial IoT (IIoT) settings, where something like a service disruption could quite literally cause a big problem. A zero-day attack targets a vulnerability that is not known to the software vendor or to antivirus vendors. By definition, a zero-day exploit is one that is not yet known and for which no detection method or patch exists. Conventional intrusion detection systems are simply not effective with zero-day vulnerabilities. A final sophisticated class of threat is represented by adversarial attacks. These are inputs that have been carefully manipulated to look normal to humans but are designed to cause machine learning or deep learning models to make

*Corresponding author. myyusnani@gmail.com

mistakes. When that happens, the model may label harmful data as benign. Together, these limitations highlight the shortcomings of traditional intrusion detection systems and provide the motivation for this paper to explore alternative, more intelligent, and adaptable means of detecting intrusions. In general, the effectiveness of such methods for zero-day attack detection is still an open problem in resource-constrained IIoT systems.

This paper reviews 25 paper of 200 research papers published between the year 2023 and 2024, focusing on recent developments in IDS technology. The reviewed literature has been classified based on their focus and broadly segregated into four categories: general AI-based IDS, Machine Learning-based IDS, Deep Learning-based IDS, and Deep Reinforcement Learning-based IDS. Figure 1 illustrates the distribution of papers in these categories, which indicates that most of the research in recent years has focused on DL-based IDS, followed by ML-based approaches [1],[2],[3].

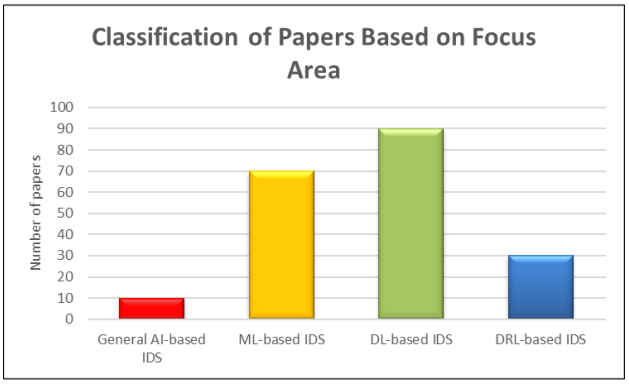

Fig. 1. Classification of Papers Based on Focus Area

This taxonomy reflects the technological evolution of IDS research, progressing from rule-based reasoning to data-driven learning, then to automated feature extraction, and finally to dynamic policy adaptation. General AI approaches often use expert systems and fuzzy logic without learning capabilities, while Machine learning (ML)-based IDS rely on algorithms like SVM and decision trees. Deep learning (DL)-based IDS leverage neural networks to model complex traffic behaviors. Deep reinforcement learning (DRL)-based IDS integrates deep learning with reinforcement learning to enable continuous adaptation in evolving threat landscapes. This categorization supports the study's objective of evaluating the effectiveness and limitations of each approach in detecting zero-day attacks in IIoT environments.

Table I presents a comparative assessment of the four categories of IDS (General AI-based, Machine learning (ML)-based, Deep learning (DL)-based and Deep reinforcement learning (DRL)-based) against key technical criteria pertinent to IIoT settings. The technical criteria that these categories of IDS are evaluated against include adaptability, scalability, computational efficiency, integration ease, flexibility, response time, real-time reliability, multi-layered security support, and maintainability.

TABLE I. COMPARATIVE EVALUATION OF IDS TECHNIQUES BASED ON KEY OPERATIONAL CRITERIA IN IIOT ENVIRONMENTS

Criterion	General AI-based IDS	ML-based IDS	DL-based IDS	DRL-based IDS
Adaptability	***Poor***	***Moderate***	***High***	***Very High***
Scalability	***Limited***	***Good***	***Good to Excellent***	***Excellent***
Computational Cost	***Low***	***Moderate***	***High***	***Very High***
Ease of Integration	***High***	***Moderate***	***Low***	***Low***
Flexibility	***Low***	***Good***	***Good***	***Excellent***
Response Time	***Fast***	***Fast to Moderate***	***Moderate to Slow***	***Moderate***
Real-time Reliability	***Poor***	***Good***	***Good***	***Excellent***
Multi-layer Security Support	***Limited***	***Moderate***	***Very Good***	***Excellent***
Maintainability	***High***	***Moderate***	***Low***	***Low***

This distribution shows an increasing interest in leveraging deep learning for the improvement of IDS capabilities, led by DL-based IDS, which stands at 45%, followed by ML-based IDS at 35%, DRL-based IDS at 15%, and finally General AI-based IDS at 5%. While deep learning approaches have enormous potential and great scope for improvement in their detectability accuracy and reduction of false negatives a particular aspect concerning the detection of few and complex attack-patterns more limitations remain with regard to adversarial attacks and high demands for computation, which IIoT environments are highly opposed to. This paper describes what zero-day attacks impede in the performance of any IDS and then proposes strategies with intent to enhance the detection of zero-negative cases.

2. ZERO-DAY ATTACKS IN IoT

Zero-day attacks take advantage of unknown vulnerabilities in either the software or hardware and leave all systems defenseless until it is patched. The reason one should be highly concerned relates to the IIoT, where this is more than a nightmare due to the extent of complexity and diversity related to connected devices. In all IIoT environments, devices run very critical operations, and if there was a successful zero-day exploit, it could disrupt, cause physical damage, or even a loss of life. The stealth factor in zero-day attacks, put together with the inability of traditional IDS to detect threats when predefined signatures are not ready, makes zero-day vulnerabilities such a great cybersecurity risk that has to be taken account of [4].

This increase in attacks has already been well-documented through several reports by Kaspersky, observing IoT devices and infrastructure online against cyberattacks [1]. Symantec's Internet Security Threat Report confirms an upward trend in attacks owing to the fact that vulnerabilities used for gaining access into those devices are usually left unpatched, thus becoming real goldmines for most hackers [2]. Fairly speaking, this denotes one of the critical pieces of evidence for growing momentum, complexity, and finally sophistication of attacks targeting IIoT ecosystems. Figure 2 are projections based on analysis of industry reports and trends, combined with vulnerability disclosures by major IIoT vendors such as Siemens and Schneider Electric. They have reported increasing numbers of vulnerabilities in their IIoT devices and systems [4],[5].

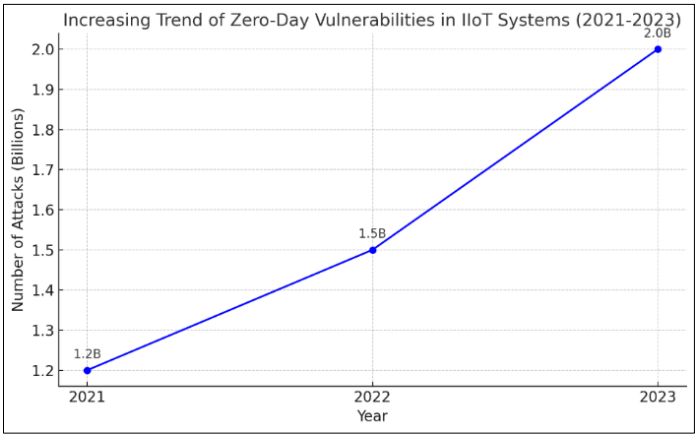

Fig. 2. Classification of Papers Based on Focus Area

Figure 3 illustrates the attack sequence and detection limitations in IIoT environments. It highlights the complexities involved in identifying these attacks, given the heterogeneous nature of IIoT devices and networks.

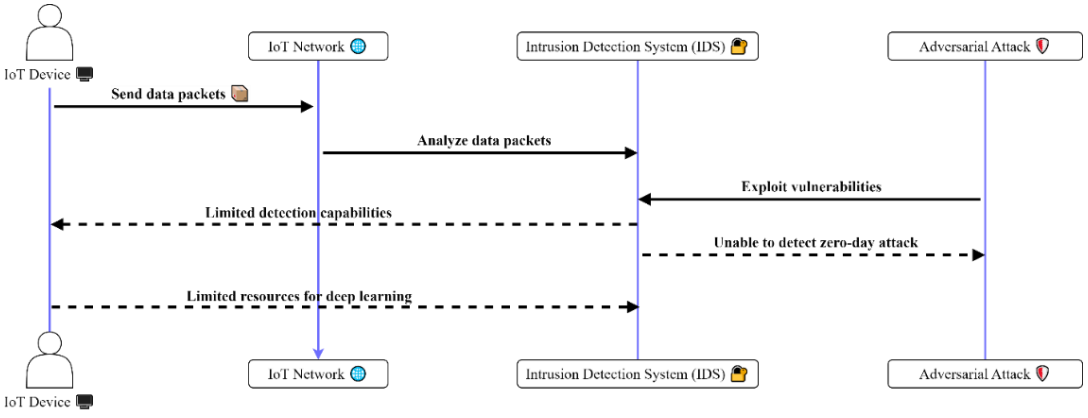

Fig. 3. Sequence of IoT/IIoT Attack and Detection Limitations

Coupled with the increase in attacks, this consistent growth in zero-day vulnerabilities underpins the increasing drive toward using advanced IDS propelled by AI as a modern strategy for maintaining security across IIoT networks. This is where the role of zero-day vulnerabilities treads on risky ground: not only do these vulnerabilities offer ways through which attackers may intrude upon systems unnoticed, but are usually those wherein exploits to vital infrastructure take place before some defenses could be granted. Besides, operational constraints in critical infrastructure make it very hard to patch IIoT systems, which, in case of exposure to a vulnerability, may remain exposed for a long period of time. According to the report by the Ponemon Institute in 2023, the average time required for detection of such an attack has increased, hence allowing the IIoT system to be susceptible for long-term exploitation [6],[7]. Further, lack of visibility and control over IoT devices further aggravates the problem. Many IoT devices are deployed in isolated networks or environments where IT personnel have limited access and oversight [8]. This makes it challenging to monitor for suspicious activities and implement timely security updates [8]. Finally, the increasing number of zero-day vulnerabilities and attacks reveals that current security defenses are not able to cope with the rapidly evolving threat landscape. More proactive detection mechanisms using artificial intelligence and machine learning are needed to improve resilience in IIoT systems and reduce the risks related to zero-day vulnerabilities.

3. LIMITATIONS OF EXISTING IDS IN DETECTING ZERO-DAY ATTACKS

Intrusion Detection Systems (IDS) is one of the most important tools that help ensure the security of IIoT environments. It continuously monitors network traffic and system behaviors to flag suspicious activities that might be indicative of a security breach. Generally, IDS can be classified into three categories: signature-based IDS, anomaly-based IDS, and hybrid systems. Each of these types has its relative merits and challenges with regard to threat detection, especially zero-day attacks. However, despite the importance of IDS in IIoT, there is a significant limitation in detecting and mitigating sophisticated threats correctly, such as zero-day vulnerabilities [9]. Figure 4 shows the main types of IDS:

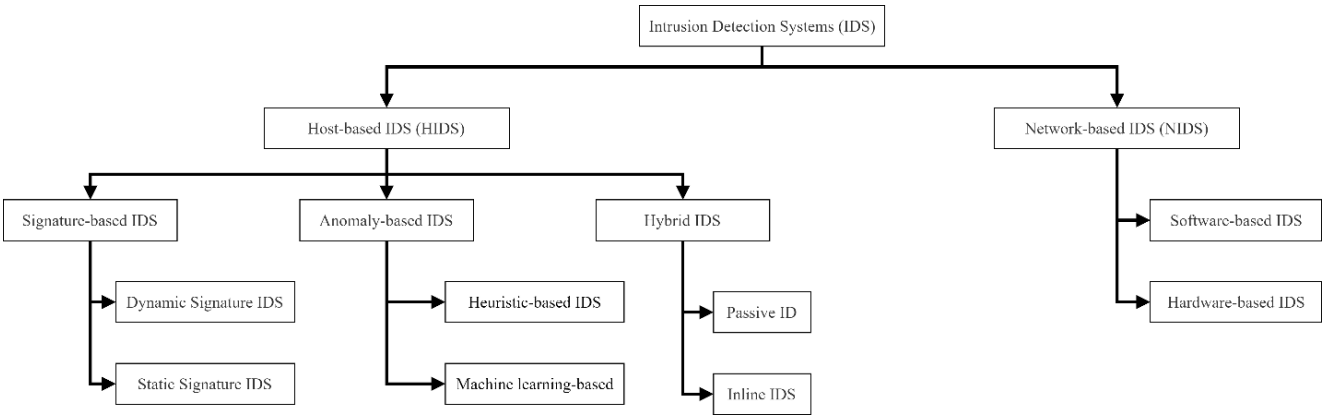

Fig. 4. Types of Intrusion Detection Systems (IDS)

- Signature-based IDS: They completely rely on predefined patterns of a security attack or predefined signatures and hence are quite efficient in identifying already-known vulnerabilities. However, detecting zero-day attacks is difficult because no prior signature exists for newly occurring vulnerabilities, leading to a high false negative rate where novel attacks remain undetected [9].
- Anomaly-based IDS: Instead of relying on a database of known attack patterns, anomaly-based IDS establish a baseline of normal system behavior. Any deviations from this baseline are flagged as potential threats. While this approach can detect unknown threats, it has a high false positive rate, particularly in complex and dynamic IIoT environments where establishing an accurate baseline is challenging [9].
- Heuristic-based IDS: These systems use rule-based methods to identify suspicious behavior and are more flexible compared to signature-based systems. However, heuristics are limited due to the evolving nature of new threats. Sophisticated attackers can create zero-day exploits that bypass known heuristics, rendering these systems ineffective. Additionally, heuristic-based systems can quickly become outdated as new attack techniques emerge [10].
- Machine learning-based IDS: These systems represent a more advanced approach, capable of adapting to new threats by learning from data. However, they face significant challenges in detecting zero-day attacks with high accuracy. Effective machine learning models require extensive training on diverse datasets, which is difficult given the scarcity of labeled data for zero-day attacks. These systems also require continuous retraining to adapt to evolving attack vectors and are vulnerable to adversarial attacks where attackers manipulate input data to evade detection[11].
- Hybrid IDS: Hybrid systems combine the strengths of signature-based and anomaly-based approaches. These systems aim to achieve better detection coverage by leveraging both known attack signatures and behavioral analysis.

However, even hybrid IDS face challenges in detecting zero-day attacks, especially in IIoT environments where resource constraints and the diverse nature of devices introduce significant complexities [9].

While each type of IDS plays a vital role in detecting and mitigating cyber threats, their performance in detecting zero-day attacks remains insufficient. IDS face many challenges in 2024 that lead to high false negative rates (FNR) impacting the reliability and performance significantly [12] as shown in Figure 5. The limitations of current IDS approaches can be summarized as follows:

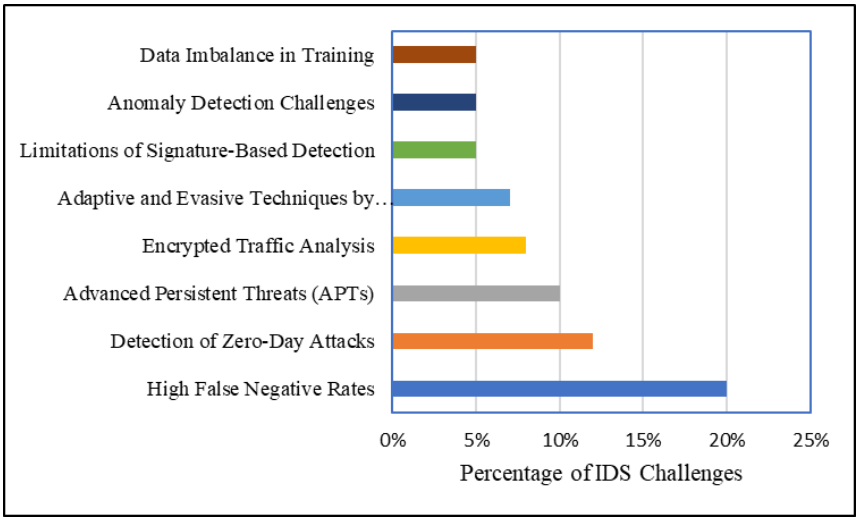

Fig. 5. Challenges in Intrusion Detection Systems (IDS) Related to False Negative Rates (FNR) in 2024

- High False Negative Rates (FNR): Signature-based IDS struggle to detect zero-day attacks due to the absence of known signatures. This results in many zero-day vulnerabilities remaining undetected, compromising the security of IIoT systems.
- False Positives Rates (FPR) in Dynamic Environments: Anomaly-based IDS face difficulties in dynamic IIoT environments, where defining a consistent "normal" state is challenging. This leads to a high number of false positives, overwhelming security teams with non-threatening alerts.
- Resource Constraints: Many IDS are computationally intensive, making them difficult to deploy effectively in IIoT devices that have limited processing power and memory. The complexity of analyzing large volumes of network traffic in real-time further compounds this issue.
- Inability to Handle Encrypted Traffic: The rise of encrypted communication protocols poses a challenge for IDS, as encrypted traffic hides malicious activities, making it difficult for traditional IDS to inspect the data and detect attacks[13].

Finaly in 2024, it has become crystal clear that traditional IDS lags behind the unprecedented growth in complexity and sophistication of cyber-attacks, especially zero-day. Traditional approaches to detection, including signature-based, anomaly-based, or even AI algorithms, have failed to provide accurate detection of the unknown threats, mostly leading to a higher number of false negatives. As the leveraging of encrypted traffic by attackers, use of imbalanced datasets, and constantly evolving techniques have gone up, it's pretty evident that IDS cannot sustain themselves on a mere static and predefined way of detection. Intrusion detection in the near future would have to be performed using algorithms that can scale and be resource-efficient and can adapt easily to changes. DRL combines strengths of DL learning with anomaly detection and real-time analysis; therefore, this can prove to be one of the bright directions ahead. However, such systems first need to overcome key challenges in terms of data imbalance, scalability requirements, and pressure to maintain realtime performance if they are to really help enhance the detection capabilities.

4. RESEARCH TREND IN IDS

Intrusion Detection Systems (IDS) have seen notable advancements and gained significant research attention in recent years. By analyzing publications from 2023 and 2024, we can categorize the research into four main areas: General AI-based IDS (excluding Machine Learning (ML), Deep Learning (DL), and Deep Reinforcement Learning (DRL)), ML-based IDS, DL-based IDS, and DRL-based IDS. . The selection of 25 representative papers from a corpus of 200 studies was performed for a focused, deeper evaluation. Nine critical criteria were used for selection, essential to the assessment of IDS applicability in

IIoT environments. These are adaptability, scalability, computational efficiency, ease of integration, flexibility, response time, real-time reliability, multi-layered security , and maintainability. These dimensions reflect real-world deployment concerns and make the most prominent papers in the selected subset representative of a comprehensive evaluation of the technological and operational requirements of IIoT-based IDS frameworks.

Figure 6 offers a comparative glance at how the distribution of IDS research publications shakes out across four main methodological types. These are General AI, Machine Learning (ML), Deep Learning (DL), and Deep Reinforcement Learning (DRL). The year 2023 is contrasted with 2024 for this overview. The crude data, as it were, comes from the manual classification of 200 peer-reviewed studies that are in our possessive grasp and focus on the kind of work that concerns usspecifically, studies that look into the kinds of things that could happen to an Industrial Internet of Things (IIoT) environment. In 2023, research predominantly focused on DL-based IDS, which made up approximately 45% of the 200 published studies. This strong interest highlights the growing use of neural networks to detect complex intrusion patterns. ML-based IDS accounted for 35% of the research, with studies utilizing algorithms like support vector machines, decision trees, and ensemble methods to improve accuracy and reduce false positives. General AI-based IDS, which do not rely on ML, DL, or DRL, comprised about 10% of the research and primarily involved rule-based and expert systems. Meanwhile, DRL-based IDS represented 10% of the publications, emphasizing the potential of adaptive learning for real-time intrusion detection (as shown in Figure 6) [14],[15].

In 2024, the research trends largely continued. Of the 200 papers published, DL-based IDS remained the dominant focus at 45%, reflecting sustained interest in deep learning techniques. ML-based IDS maintained a significant presence, representing 35% of the research. General AI-based IDS saw a slight decline to 6%, as more advanced methods gained attention. DRLbased IDS increased to 15% of the publications, signaling growing recognition of its adaptive capabilities in managing evolving threats (as shown in Figure 6) [14],[15].

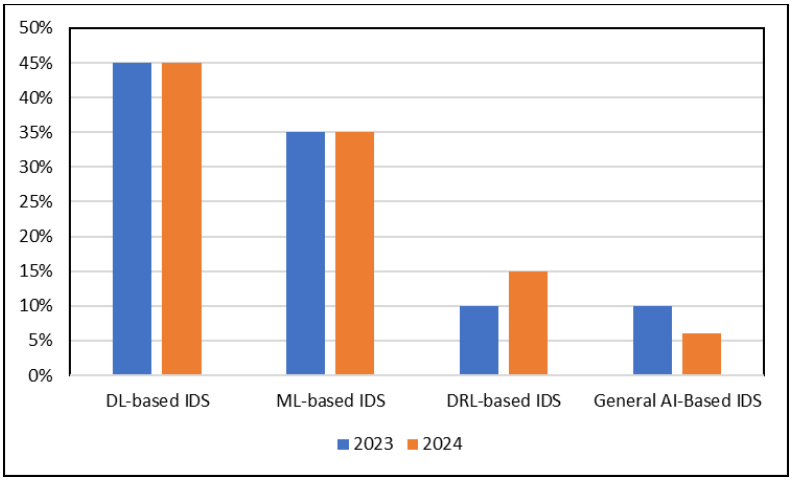

Fig. 6. Intrusion Detection System Research Statistics In 2023 and 2024

Overall, the increased focus on DL and DRL-based IDS highlights the importance of advanced learning methods in developing robust intrusion detection systems, while ML approaches continue to play a critical role in improving detection performance. . Even though large intrusion detection system (IDS) studies depend on widely accepted benchmark datasets for evaluation like UNSW-NB15, CICIDS2017, WUSTL-IIoT-2021, and NSL-KDD they share crucial limitations that can profoundly affect the external validity and real-world relevance of research findings. Indeed, several recent papers have already pointed out that these datasets suffer from very severe class imbalance, overly simplistic and static definitions of what constitutes an attack, and a very poor representation of the kinds of sophisticated, modern zero-day threats that are common in today's world. Additionally, even the best of these datasets do not come even close to capturing the dynamic traffic behavior of real IIoT environments, where the volume, timing, and even the topology of communication can vary all over the place from one moment to the next.

Deep Learning (DL) and Deep Reinforcement Learning (DRL) offer significant advantages, including higher accuracy, fewer false negatives, and adaptive learning capabilities. However, these methods introduce substantial challenges. First, DL models require extensive labeled datasets for training, which are costly and labor-intensive to produce. Second, despite their inherently parallel architectures, DL systems suffer from computational inefficiencies; training and inference demand heavy computing resources, and scaling remains problematic. This computational burden also translates into high energy consumption, limiting practical deployment. Most critically, while DL models excel at handling natural variations in real-world data, their sensitivity to adversarial perturbations makes them vulnerable to evasion attacks. Unlike traditional hand-

engineered systems, even imperceptible input modifications can induce misclassification, posing security risks in critical applications.

This research looks into developing more realistic, heterogeneous, and continuously updated datasets to support robust evaluations. In addition, the author emphasizes how important it is to design lightweight deep learning/deep reinforcement learning models, to enhance their adversarial robustness, and to explore hybrid frameworks that combine learning-based strategies with rule-based or federated systems to overcome existing deployment constraints.

4.1 General AI-based Intrusion Detection Systems

The reviewed studies highlight various AI methods used in Intrusion Detection Systems (IDS), each with specific strengths and limitations as shown in Table II. A specific AI-based IDS was developed by Gökçe Karacayılmaz et al.[16] to protect devices in Industrial Internet of Things (IIoT) against Man-in-the-Middle and Start-Stop attacks as well Distributed Denial of Service. The system integrates neural networks (NN) with the ReLU activation function and uses a continuous network monitor in order to improve efficiency for identifying threats. The system was tested on real-world PLC traffic, and it achieved an accuracy of 99.7% with a low false positive rate of just.002%, showcasing its ability to effectively distinguish between malicious and non-malicious traffic.

Building upon AI approaches, they devise a solution to distribute COVID-19 cases across hospitals in Rajdeep Borgohain et al.[17] presented IDS in which fuzzy logic is used in combination with genetic algorithms to improve network anomaly detection. This new behavior helps the system to handle uncertain data and adaptively optimize detection rules, which will mitigate false positives that are common cases in rule-based systems. While the team behind AI Breaker were not specific on performance results, they did mention that genetic algorithms allow their system to adapt and as a result provide some resiliency in the face of complex, ambiguous attacks.

Similarly, Shao-Shin Hung et al.[18] introduced an ontology-based IDS framework to address the issue of customizing multiple domain-specifics IDS solutions by non-experts. Using domain-specific knowledge and intelligent reasoning, the system allows users to design IDS applications without detailed technical knowledges. The application of IDS alone can not deliver the required hit rates except in a ontology based approach that which outperformed all tested methods on DoS and U2R attacks with higher detection rate((hitrate) for both 0.9028 superior to any other methods tried.

Philokypros P. Ioulianou et al.[19]: IoTCrawler: Browsing the Internet of Things (IoT), focused on IoT networks Developed a signature based IDS, which can detect known attacks especially (Denial of Service) DoS and routing attack patterns. They offer a hybrid nature of detection, combining centralized and distributed methods in order to provide full protection at the network level. Although their simulations were successful in detecting attacks, they raised concerns of the power consumption by low-power IoT devices during long-lasting DoS flooding.

Extending the use of fuzzy logic, Mohammad Almseidin et al.[20] proposed a Fuzzy Logic-based IDS for detecting DDoS attacks. In this system the membership functions are trapezoid and it is used to Mamdani inference method for precise response conclusion. With a true positive rate of 91.1% and a corresponding false positive rate as low as just 0.006%.

Expanding the focus on cybersecurity, Pranita Binnar et al.[21] examined the security challenges of IIoT systems with Digital Forensic Incident Response (DFIR)models. Their research underscores the significance of incident response in managing cyberattacks. The performance of the forensic tools, though not in numbers based on specific given results gives a mechanism idea to how such tool may help administrators for ensuring security and integrity from more complex deeply interconnected system like Industrial Control Systems (ICS).

Finally the last study is Sulyman Age Abdulkareem et al.[22] organized an automatic SEL (Stack Ensemble Learning) mechanism to localize all sorts of a network-like attack in IIoT surroundings. In their system, data dimensionality is reduced using Feature Importance (FI) to increase the detection accuracy yet keep high computational efficiency. It was tested on the Edge-IIoTset, achieving an accuracy of 87.37%, where performance is traded off with lower computational cost.

TABLE II. SUMMARY OF AI METHODS, PERFORMANCE, AND DRAWBACKS

Author	*Year*	*Type of Publication*	*AI Method*	*Performance Matrix*	*Drawback*
Gökçe Karacayılmaz [16]	*2024*	*Journal Article (Cluster Computing)*	*A hybrid approach combining rule-based reasoning, anomaly detection, and reinforcement learning*	*accuracy:99.7% precision:0.993 recall:0.993 F1-score:0.993.*	*The system's performance and adaptability in larger and more complex industrial settings may need further testing and refinement, especially for real-time processing and handling extreme data flow*

Author	Year	Type of Publication	AI Method	Performance Matrix	Drawback
Rajdeep Borgohain [17]	2024	Journal Article (International Journal of Advance Networking and Applications)	Fuzzy Logic and Genetic Algorithms (FuGeIDS - Fuzzy-Genetic Intrusion Detection System)	The paper discusses reducing false positives but does not provide a specific accuracy or performance metric in number form	The labor-intensive process of rule generation in systems like FIRE (Fuzzy Intrusion Recognition Engine) and challenges in initializing agents for training in genetic algorithm models
Shao-Shin Hung [18]	2008	Journal Article (Computer Standards & Interfaces)	Ontology-based model for network intrusion detection.	DoS Detection Rate (PD): 90.12% False Alarm Rate (FAR) for DoS: 0.0037 U2R Detection Rate (PD): 98.67% False Alarm Rate (FAR) for U2R: 0.0029 R2L Detection Rate (PD): 15.02% False Alarm Rate (FAR) for R2L: 0.0000027 Probing Detection Rate (PD): 60.01% False Alarm Rate (FAR) for Probing: 0.0000053	The model had issues with detecting rare categories of attacks like R2L and U2R, which still suffer from higher false negative rates and lower detection accuracy compared to common attack categories
Philokypros P. Ioulianou [19]	2018	Conference Paper	Signature-based Intrusion Detection System (IDS) with centralized and distributed components	No specific numerical performance metrics like accuracy or F1 score were provided, but the system was tested on DoS attacks like "Hello" flooding and version number modification using the Cooja simulator.	The system requires installation of detection modules near devices, which can add complexity. Also, high energy consumption was noted in certain nodes during attack scenarios
Mohammad Almseidin [20]	2019	Journal Article	Anomaly-based Intrusion Detection System (IDS) using Fuzzy Logic	True Positive Rate (TPR): 91.1% False Positive Rate (FPR): 0.006 False Negative Rate (FNR): 0.089 True Negative Rate (TNR): 99.4%	The system may suffer from limitations in detecting new types of DDoS attacks, such as HTTP flood and SQL injection, due to the dataset used for training and testing (DDoS-2016 dataset). It relies heavily on feature selection and may not generalize well to all intrusion types
Pranita Binnar [21]	2024	Journal Article (Cyber Security and Applications)	o specific AI method was applied; the paper focuses on using Digital Forensic Incident Response (DFIR) for securing Industrial IoT (IIoT) systems, including Cyber-Physical Systems (CPS), Industrial Control Systems (ICS), and SCADA.	The document does not provide numerical performance metrics for this framework.	The paper highlights the challenges in integrating DFIR with existing IIoT security solutions due to the dynamic and resource-constrained nature of IIoT systems. Additionally, there is a lack of standard forensic tools specifically designed for SCADA systems
Sulyman Age Abdulkareem [22]	2024	Journal Article (Journal of Network and Computer Applications)	Stacking Ensemble Learning (SEL) with Feature Importance (FI) for dimensionality reduction.	Accuracy: 87.37% Precision: 90.65% Recall: 77.73% F1 Score: 80.88% Training Time: 16.18 seconds Testing Time: 0.10 seconds	The system may face challenges in detecting more complex, zero-day attacks due to limitations in signature-based detection. There is also a trade-off between feature reduction and maintaining detection accuracy for high-dimensional datasets

4.2 Machine Learning-based Intrusion Detection Systems

Machine learning is crucial for detecting intrusions in IoT networks by analyzing system logs and identifying user and device behavior patterns. Various models and techniques have been proposed to enhance intrusion detection as shown in Table III. Yakub Kayode Saheed et al.[23] used the UNSW-NB15 dataset and applied PCA for feature selection, training classifiers like XGBoost and SVM on the reduced dataset. Aliaa Al-Bakaa et al.[24] utilized redundancy quantitative analysis (RQA) with the UNSW-NB15 dataset to detect intrusions. Rehab Alanazi et al.[25] employed anomaly-based machine learning with the X-IIoTID dataset, incorporating Neighborhood Components Analysis and Minimum Redundancy Maximum Relevance for feature selection. R. Gopi et al.[26] proposed the CCSOA-OWKELM technique, combining chaotic cuckoo search

optimization and optimal wavelet kernel extreme learning machine for feature selection and classification. Mojtaba Eskandari et al.[27] developed Passban, an IDS for IoT devices using edge computing. Ghada Abdelmoumin et al.[28] explored optimizing single-learner AML-IDS with PCA and 1-SVM AML-IDS models. Mohammed M. Alani et al.[29] introduced a two-layer IDS integrating machine learning with flow-based and packet-based features. Zhihan Lv et al.[30] proposed an IDS using a weighted sample and class C support vector machine (CSWC-SVM) and tested it with KDD CUP 1999 data in a simulated environment. Dataset selection remains critical for building effective IDS models, influencing the performance of machine learning algorithms. While Machine Learning is a very promising approach to develop intrusion detection systems and other cybersecurity mechanisms, a lot of stubborn issues significantly reduce their effectiveness. One such serious issue is that the number of false negatives, especially in anomaly-based IDSs, is high. Instead, an anomaly-based IDS is designed to raise the red flag in case of suspect behavior; it could just poorly detect malicious activities. That is, the threats that are not detected and pass through the system leave the networks exposed to possible attacks. The false negatives pose a serious risk where operators may be confident in the safety of their network, thereby overlooking or underestimating a real threat . Another key disadvantage is the need for large amounts of high-quality labeled data to train ML models effectively. In cybersecurity, such datasets are hard to come by, and static datasets used in training often fail to reflect the dynamic nature of real-world networks. The biggest barrier still remains the inability to keep pace with the changing nature of the threats, say zero-day attacks. Because of this, many ML models easily fall into overfitting, where they perform excellently on training data but generalize very poorly to new, unseen data in live environments. Scalability is another concern, especially for resource-intensive models such as deep learning, probably requiring huge computational power. This turns out to be a serious bottleneck in resource-constrained environments, such as the IIoT. It is often impractical to deploy ML models in such scenarios due to their high processing demands and latency issues that hinder real-time threat detection. Another problem with ML models is that they require extensive manual intervention in the form of feature engineering and parameter tuning. These processes may be cumbersome and hence become very time-consuming something that reduces the overall throughput of IDS deployments and affects the scalability of IDS in large diverse environments.

TABLE III. SUMMARY OF ML METHODS, PERFORMANCE, AND DRAWBACKS

Author	Year	Type of Publication	AI Method	Performance Matrix	Drawback
Yakub Kayode Saheed [23]	*2022*	*Journal Article (Alexandria Engineering Journal)*	*Multiple machine learning algorithms were used, including CatBoost, XGBoost,*	*PCA-XGBoost: Accuracy = 99.99%, MCC = 99.97% PCA-CatBoost: Accuracy = 99.99%, MCC = 99.97% PCA-KNN: Accuracy = 99.98%, MCC = 99.96% PCA-SVM: Accuracy = 99.98%, MCC = 99.96% PCA-QDA: Accuracy = 99.97%, MCC = 99.94% PCA-NB: Accuracy = 97.14%, MCC = 93.41%*	*The Naïve Bayes model had lower accuracy compared to other models (97.14%), and the MCC was also relatively lower at 93.41%. Additionally, some previous studies focused on older datasets like NSL-KDD, which may not reflect present-day IoT attacks.*
Aliaa Al-Bakaa [24]	*2022*	*Journal Article (Computers & Security)*	*Recurrence Quantification Analysis (RQA) combined with Machine Learning classifiers (Logistic Regression, K-Nearest Neighbors, Decision Tree, Random Forest).*	*Logistic Regression with RQA: Accuracy = 94.71%, F-score = 0.8720 KNN with RQA: Accuracy = 96.24%, F-score = 0.9121 Decision Tree with RQA: Accuracy = 94.82%, F-score = 0.8728 Random Forest with RQA: Accuracy = 96.28%, F-score = 0.9124*	*RQA significantly improves the detection accuracy but shows less improvement with certain features (e.g., the srcip feature), and the encoding technique used may affect performance*
Rehab Alanazi [25]	*2023*	*Journal Article (Computer Systems Science & Engineering)*	*The modle employed multiple machine learning classifiers: Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), and Linear Discriminant Analysis (LDA). Feature selection methods used include Minimum Redundancy Maximum Relevance (MRMR)*	*Decision Tree (DT) with MRMR: Accuracy = 99.58%, Sensitivity = 99.59%, Specificity = 99.58%, F1-score = 99.59%, False Positive Rate (FPR) = 0.42% K-Nearest Neighbors (KNN) with MRMR: Accuracy = 98.65%, Sensitivity = 98.93%, Specificity = 98.37%, F1-score = 98.67%, FPR = 1.63% SVM with MRMR: Accuracy = 85.81%, Sensitivity = 73.80%, Specificity = 97.99%, F1-score = 83.97%, FPR = 2.01% LDA with MRMR: Accuracy = 85.58%, Sensitivity = 73.76%, Specificity = 97.57%, F1-score = 83.75%, FPR = 2.42%*	*Some classifiers (SVM and LDA) showed lower performance compared to DT and KNN, especially in terms of accuracy and false positive rate. Also, the feature selection method MRMR yielded better performance than NCA*

Author	Year	Type of Publication	AI Method	Performance Matrix	Drawback
R. Gopi [26]	2023	Journal Article (Computer Systems Science & Engineering)	Chaotic Cuckoo Search Optimization Algorithm (CCSOA) with Optimal Wavelet Kernel Extreme Learning Machine (OWKELM). The CCSOA is used for feature selection, and OWKELM is applied for intrusion detection and classification. The method also includes hyperparameter tuning using the Sunflower Optimization (SFO) algorithm.	NSL-KDD Dataset: Precision = 99.98%, Recall = 99.97%, Accuracy = 99.97%, F-Score = 99.97% CICIDS2017 Dataset: Precision = 99.91%, Recall = 99.91%, Accuracy = 99.92%, F-Score = 99.91%	The CCSOA-OWKELM technique, although highly accurate, involves high computational complexity due to the multiple layers of feature selection and hyperparameter optimization, making it resource-intensive
Mojtaba Eskandari [27]	2020	Journal Article (IEEE Internet of Things Journal)	Anomaly-based intrusion detection system (Passban IDS) using one-class classification algorithms, including Isolation Forest and Local Outlier Factor (LOF)	Isolation Forest: F1 Score = 0.99 (Port Scanning), 0.96 (HTTP Brute Force), 0.96 (SSH Brute Force), 0.90 (SYN Flood) Local Outlier Factor (LOF): Lower performance compared to Isolation Forest in several attack detection scenarios.	The system may produce false positives when encountering traffic that deviates from the routine but is not an actual attack. Additionally, when the system is under SYN flood attacks, its detection rate decreases due to the computational resources being overwhelmed
Ghada Abdelmoumin [28]	2022	Journal Article (IEEE Internet of Things Journal)	Anomaly-based machine learning-enabled intrusion detection system (AML-IDS) models using Principal Component Analysis (PCA) and One-Class Support Vector Machine (1-SVM) with ensemble learning (Stacking technique) for optimization.	PCA AML-IDS Model: AUC = 0.139, F1-Score = 0.313 (optimized with IoT_Botnet dataset) 1-SVM AML-IDS Model: AUC = 0.472, F1-Score = 0.968 (optimized with IoT_Botnet dataset) Ensemble models combining PCA, SVM, and DL-IDS achieved AUC = 1 and F1-Score = 0.966 (IoT_Botnet dataset)	Single-learner models such as PCA and 1-SVM AML-IDS exhibit low detection rates and high false positives when compared to DL-IDS. The models struggle with dissimilarity between training and testing data, imbalanced datasets, and the use of single-learner methods, which affects their prediction accuracy
Mohammed M. Alani [29]	2023	Journal Article (IEEE Transactions on Industrial Informatics)	A two-layer Intrusion Detection System (IDS) using machine learning with flow-based and packet-based classifiers. The classifiers used are XGBoost (XGB) for packet-based detection and Random Forest (RF) for flow-based detection	Packet-based classifier (XGB): Accuracy = 99.15%, FPR = 0.96%, FNR = 0.61% Flow-based classifier (RF): Accuracy = 99.66%, FPR = 0.26%, FNR = 0.40%	The main drawback is the computational complexity of the system due to handling both packet-level and flow-level features, which may affect deployment efficiency in resource-constrained IoT environments
Zhihan Lv [30]	2021	Journal Article (IEEE Internet of Things Journal)	Class and Sample Weighted C-Support Vector Machine (CSWC-SVM) algorithm with SVM for intrusion detection in industrial control systems.	Polynomial Kernel Function: Accuracy = 85.7%, False Positive Rate = 3.8% Radial Basis Kernel Function: Accuracy = 86.2%, False Positive Rate = 2.8% Sigmoid Kernel Function: Accuracy = 86.7%, False Positive Rate = 2.3%	Although the CSWC-SVM algorithm improves classification accuracy and reduces false positives, the model's performance may degrade when dealing with new, unknown attack types due to the limitations of the dataset used (KDD CUP 1999), which lacks up-to-date attack patterns

4.3 Deep Learning-based Intrusion Detection Systems

Deep learning significantly enhances intrusion detection by enabling systems to recognize patterns and detect abnormal behaviors, crucial for securing computer systems against cyberattacks. Various models and techniques have been proposed for different scenarios as shown in Table IV.

Mohamed Abdel-Basset et al.[31] proposed Deep-IFS, which relies on Local Gated Recurrent Units and Multi-Head SelfAttention to reduce computation time significantly with substantial scalability. On the Bot-IoT dataset, the model was able to give an accuracy of 99.77%, precision-100%, and F1 score-99.88%. This model is highly suitable for performing intrusion detection tasks in real-time in IIoT. Izhar Ahmed Khan et al.[32], inspired by such performance-oriented models, proposed a deep autoencoder model of IDS using LSTM for ICS network traffic monitoring. High accuracy rates were achieved, with over 97.95% for the GP dataset and 97.62% for the UNSW-NB15 dataset, which are quite impactful for real-time traffic monitoring. With a similar focus on improving detection accuracy, Pampapathi B M et al.[33] propose the Filtered Deep Learning Neural Network-FDLNN-relying on techniques like Modified-K-Means and Elephant Herding Optimization. Using classic approaches in testing on the ToN-IoT dataset, the proposed model produces an accuracy of as high as 96.12%. Following the trend of model optimization, Jawad Ahmad et al.[34] presented a Deep Random Neural Network trained by Particle Swarm Optimization and Sequential Quadratic Programming. Performance was astonishing, with accuracy up to 99.57%, on datasets like UNSW-NB15 and ToN-IoT. It was presented with belief in its robustness regarding the detection of cyber-attacks. Onto this Rasheed Ahmad et al.[35] conducted benchmarking in order to compare

various deep learning methods against each other. Comparing the performance of the MLP, CNN, LSTM, and TCN models using several datasets has shown some of these models, such as the TCN and LSTM models, yielding as high as 100% accuracy and therefore proving to be viable for IoT IDS. Coming back to practical scenarios, Dusan et al.[36] proposed a semi-supervised CNNbased methodology to detect cyberattacks on the communication links in ICS. On a real-world scenario test of the SWaT testbed, the proposed approach returned high values of precision and recall with low TPR, hence suitable for real-time anomaly detection. Still widening the boundaries of semi-supervised learning, Mohammad Mehedi et al.[37] combined unsupervised and supervised in their intrusion detection model. The tested model on an IIoT testbed reached 96.63% accuracy, which outperformed the state-of-the-art traditional detection systems. Meanwhile, Akbar Telikani et al.[38] proposed the EvolCostDeep model, which hybridized the autoencoder with CNN. The performance of this model-when tested on the ToN-IoT dataset-showed a mean recall of 93.3% and precision of 97.6%, thus balancing the detection accuracy with scalability. Bhawana Sharma et al.[39] implemented GANs over DNNs to improve the generalization of deep learning models. The model fared well, enhancing the accuracy 7% on the UNSWNB15 dataset. Lastly but not least, Xiaofeng Wang et al.[40] studied federated learning for anomaly detection in IoT networks. Their proposed decentralized model, which keeps the data private and employs mutual information, outperforms the benchmark with a quality score of 99.4% on accuracy for the IOT-Botnet 2020 dataset, thereby indicating that federated learning does work to enhance security in IoT without giving away sensitive data.

Deep learning (DL) techniques are undeniably powerful, but they come with a few significant challenges, especially in certain situations. One of the main issues is how sensitive these models are to imbalanced datasets. Imagine you have a dataset that's full of normal behavior data but only a small portion showing actual cyberattacks. Deep learning models tend to focus on the majority of the data the normal behavior leaving them less capable of spotting the rarer, more complex attack patterns. This might result in top-shelf overall accuracy on paper but, in reality, these models may lose sight of the rare and much more dangerous attacks, which increase false negatives. Another challenge is regarding computational power that DL models need. Models such as deep neural networks and CNN require a lot of computing power, memory, and storage to work efficiently. It is pretty tough to deploy these types of models on devices or any system with limited resources, as is usually found in IoT networks. Besides, setting up and fine-tuning these models for the perfect output is a complex and resource-consuming task, often requiring experienced hands. This is the careful tuning that, if not done, makes models either overfitting that is, performing well on training data but poorly on new, unseen data or underfitting, where they don't even capture the patterns in the training data.

Other weaknesses in DL include adversarial attacks tiny, hardly noticeable modifications to input data send the model's predictions completely askew. This can amount to a big problem in the cybersecurity world where high-stakes decisions are made day by day. Finally, to be appropriately optimized, these models require vast amounts of labeled data, which are not easily available at all times. Another challenge in working with deep learning is that collecting and labeling data is time-consuming and expensive

TABLE IV. SUMMARY OF DL METHODS, PERFORMANCE, AND DRAWBACKS

Author	Year	Type of Publication	AI Method	Performance Matrix	Drawback
Mohamed Abdel-Basset [31]	*2021*	*IEEE Transactions on Industrial Informatics*	*Deep-IFS (Intrusion Detection Approach using a LocalGRU and Multihead Attention in Fog Computing)*	*Accuracy: 99.77% Precision: 99.99% Recall: 99.77% F1-Measure: 99.88%*	*The Deep-IFS model is trained in a supervised manner, limiting its ability to learn from unlabeled data. There are also concerns about data privacy and the potential for inefficiency in handling large volumes of IIoT traffic due to its centralized training structure*
Izhar Ahmed Khan [32]	*2022*	*Ad Hoc Networks (Elsevier Journal)*	*Deep Autoencoder-based Long Short-Term Memory (LSTM) model*	*Gas Pipeline Dataset: Accuracy: 97.95% Precision: 98% Recall: 96.63% F1 Score: 97.89% UNSW-NB15 Dataset: Accuracy: 97.62% Precision: 97.69% Recall: 96.01% F1 Score: 97.55%*	*The proposed model struggles with processing multiclass problems, as it is optimized for binary classification (normal vs. attack). It also requires improvements to detect specific attack classes more accurately*
pathi B M [33]	*2022*	*elematics and Informatics Reports*	*Filtered Deep Learning Neural Network (FDLNN)*	*Accuracy: 96.12% Precision: 94.75% Recall: 95.23% F1-Measure: 94.98%*	*The model has limitations in handling heterogeneous data sources and lacks efficiency in detecting zero-day attacks*

Author	Year	Type of Publication	AI Method	Performance Matrix	Drawback
Jawad Ahmad [34]	2022	Journal of King Saud University Computer and Information Sciences	Deep Random Neural Network with Particle Swarm Optimization (DRaNN-PSO)	DS2OS Dataset: Training Accuracy: 98.97% Testing Accuracy: 98.64% Precision: 99.11% Recall: 99.03% F1-Score: 99.06% UNSW-NB15 Dataset: Training Accuracy: 99.48% Testing Accuracy: 99.12% Precision: 99.53% Recall: 99.50% F1-Score: 99.51% ToN_IoT Dataset: Training Accuracy: 99.72% Testing Accuracy: 99.57% Precision: 99.66% Recall: 99.59% F1-Score: 99.62%	The model has higher computational complexity due to the combination of PSO and SQP, making it less efficient for real-time intrusion detection applications
Rasheed Ahmad [35]	2021	Future Generation Computer Systems (Elsevier)	Convolutional Neural Network (CNN) with Transfer Learning	Accuracy: 98.52% Precision: 98.65% Recall: 97.89% F1-Score: 98.27%	The model struggles with high-dimensional feature spaces, leading to increased training time and computational resource requirements. Additionally, its detection accuracy for complex, multi-stage attacks remains an area for improvement
Dusan Nedeljkovic [36]	2022	Computers & Security (Elsevier)	Convolutional Neural Network (CNN) for cyber-attack detection in industrial control systems	Accuracy: 97.85% Precision: 93.8% Recall: 96.8% F1-Score: 95.3%	The method faces challenges with real-time deployment in industrial systems due to its computational complexity, especially when scaling to large, distributed systems
Mohammad Mehedi [37]	2021	IEEE Transactions on Industrial Informatics	Deep-Learning Feature-Extraction-Based Semisupervised Model	Accuracy: 96.63%	The model faces challenges with real-time deployment due to computational complexity and the large trust boundaries of IIoT networks, which require rapid adaptation to dynamic attack patterns
Akbar Telikani [38]	2022	IEEE Internet of Things Journal	EvolCostDeep (A hybrid model of stacked autoencoders and convolutional neural networks)	UNSW-NB15 Dataset: Accuracy: Not explicitly mentioned Precision: 98% Recall: 90.6% F1-Score: 93.8% ToN-IoT Dataset: Accuracy: Not explicitly mentioned Precision: 97.3% Recall: 96.1% F1-Score: 96.6%	The model faces scalability issues when handling large datasets, which results in high computational time. The cost-sensitive learning model requires further optimization to reduce training time
Bhawana Sharma [39]	2023	Computers and Electrical Engineering	Deep Neural Network (DNN) with GAN (Generative Adversarial Networks) for class balancing	Accuracy: 91% Precision: 1 (DoS class), 0.84 (Exploits class), 0.76 (Fuzzers class) Recall: 1 (DoS class), 0.92 (Exploits class), 0.86 (Fuzzers class) F1-Score: 1 (DoS class), 0.88 (Exploits class), 0.81 (Fuzzers class)	he model's main challenge is dealing with imbalanced datasets, which require additional techniques like GANs to balance minority classes. Additionally, the complexity of the model increases, which may affect real-time deployment and computational efficiency
Xiaofeng Wang [40]	2023	Computers and Electrical Engineering	Federated Deep Neural Network (Federated MI-DNN) for anomaly detection	Accuracy: 99.4% F1-Score: 99.3% True Positive Rate: 99.3% True Negative Rate: 99.2%	The model has high communication costs during federated learning rounds and requires further optimization to reduce the communication overhead during data transfer between clients and the server

5. Practical Implementation of IDS Technologies in Industrial Contexts

In today's interconnected industrial landscape, cybersecurity solutions are no longer a luxury they are foundational. Different enterprises adopt different paradigms of Intrusion Detection Systems (IDS), not just as technical tools, but as strategic choices that reflect their risk profiles, operational environments, and technological maturity.

At PetroSafe Energy, a petrochemical corporation operating in a stable infrastructure with known network behaviors, the security team has implemented a General AI-based IDS built on expert systems and rule-based logic. This legacy approach thrives in low-variability environments, where threats follow predictable patterns and the cost of computational overhead must be minimal. Although limited in adaptability, these systems remain relevant for contexts where system transparency and determinism are critical [35].

Meanwhile, FlexiTech Smart Devices, a manufacturer of modular consumer IoT equipment, employs a Machine Learning-based IDS powered by classical classifiers such as Random Forest and Support Vector Machines. These models, trained on semi-supervised logs, can detect a broad range of known attacks, and adapt quickly when retrained with new data. For mid-scale environments with partially dynamic networks, ML offers a reliable balance between interpretability and automation [36].

In a more technologically demanding sector, NeuroGrid Robotics, a high-precision robotics firm, deploys a Deep Learning-based IDS using Convolutional and Recurrent Neural Networks. With millions of data packets flowing from sensors, actuators, and cloud nodes, traditional ML techniques become insufficient. DL models here are capable of capturing latent dependencies and nonlinear relationships within traffic features, enabling high-accuracy detection even for polymorphic or stealthy attacks [37].

At the edge of innovation stands AutoSecure AI Systems, a pioneer in autonomous cybersecurity solutions for Industry 4.0. Their IDS integrates Deep Reinforcement Learning (DRL) to continuously interact with the network environment, learn optimal detection policies over time, and adapt defenses based on new, zero-day attack patterns. DRL's ability to balance exploration and exploitation makes it especially suited for highly dynamic and adversarial IIoT settings [38]. While powerful, the DRL model still faces challenges such as explainability (the "black-box" problem) and computational cost, which are active research areas [39].

6. CONCLUSION

Zero-day attacks keep being a big problem for intrusion detection systems (IDS), leading to high false negative rates as these new threats go undetected. While traditional signature-based IDS have a hard time identifying such emerging threats, machine learning and deep learning-based approaches show promise by learning to recognize anomalies and behavioral deviations. Machine learning models can generalize from known data and infer unseen threats. Deep learning models extract complex feature representations that allow for more nuanced identification of new attack vectors. Still, it's worth noting that both techniques are limited in their effectiveness against adversarial inputs, and that the amount of power they require for training tends to hobble their real-world usefulness in settings like the IIoT. One up-and-coming alternative that is worth watching is an IDS that uses Deep Reinforcement Learning (DRL) to ensure zero-day attacks don't lead to big problems. Right now, an IDS that uses DRL to keep up can learn from all the new business data that's coming into the IIoT and can learn in such a way that it keeps a balance between exploiting Detection Policies it already knows about and exploring new strategies it hasn't seen before.

This study of General AI, ML, DL, and DRL-based IDS frameworks and comparative analysis of offers provides a direct answer to the research question guiding this survey and the "how" of the research: How can we enhance existing IDS frameworks to reduce the risks of zero-day attacks in Industrial IoT? This is a question of great practical significance. However, the central thesis that consolidates the comparative analysis and offers a meaningful answer to the research question guiding the study is this: No paradigm, be it based on General AI, ML, or DL, offers a universal solution or a clear path forward for enhancing existing IDS frameworks in this manner. Instead, the optimal match for the target problem seems to be IDS frameworks based on Deep Reinforcement Learning.Even so, DRL-based IDS frameworks have many practical shortcomings that can and should be addressed by future research. For instance, the high computational cost of training and inference may make it impossible to deploy these systems in the resource-restrained edge environments that make up the IIoT. Then there's the black-box problem. Black-box systems are, by definition, inscrutable. This means the decision-making logic of a black-box system can be understood only through the use of another system like a human with a lot of knowledge about IDS solutions if it can be understood at all. And that's a big if, for which there are many good reasons to be skeptical.

7. ACKNOWLEDGMENT

The authors would like to express sincere appreciation to the Universiti Teknologi MARA (UiTM), 40450, Selangor, Malaysia for the support provided throughout this project. The availability of research facilities, technical tools, and a conducive academic environment greatly contributed to the successful completion of this work. The authors also acknowledge the assistance and cooperation of the academic and technical staff involved.

References

[1] D. Brecht, 'Network Intrusion Detection Systems: a 101', *Http://Www.Brighthub.Com/*, 2010, [Online]. Available: http://www.brighthub.com/computing/smb-security/articles/38389.aspx

[2] B. Alotaibi, 'A Survey on Industrial Internet of Things Security: Requirements, Attacks, AI-Based Solutions, and Edge Computing Opportunities', *Sensors*, vol. 23, no. 17, 2023, doi: 10.3390/s23177470.

[3] Y. Wu, B. Zou, and Y. Cao, 'Current Status and Challenges and Future Trends of Deep Learning-Based Intrusion Detection Models', 2024.

[4] F. Luo, J. Wang, X. Zhang, Y. Jiang, Z. Li, and C. Luo, 'In-vehicle network intrusion detection systems: a systematic survey of deep learning-based approaches', *PeerJ Comput. Sci.*, vol. 9, 2023, doi: 10.7717/peerj-cs.1648.

[5] N. Alam and M. Ahmed, 'Zero-day Network Intrusion Detection using Machine Learning Approach', *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. October, pp. 194–201, 2023, doi: 10.17762/ijritcc.v11i8s.7190.

[6] Z. Dai *et al.*, 'An intrusion detection model to detect zero-day attacks in unseen data using machine learning', *PLoS One*, vol. 19, no. 9, pp. 1–25, 2024, doi: 10.1371/journal.pone.0308469.

[7] M. A. Alsoufi *et al.*, 'Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review', *Appl. Sci.*, vol. 11, no. 18, 2021, doi: 10.3390/app11188383.

[8] 'Cost of a data breach 2024 | IBM'. Accessed: Oct. 25, 2024. [Online]. Available: https://www.ibm.com/reports/data-breach

[9] 'Home | Ponemon Institute'. Accessed: Oct. 25, 2024. [Online]. Available: https://www.ponemon.org/

[10] 'ICS Patch Tuesday: Siemens, Schneider Electric, CISA Issue Advisories - SecurityWeek'. Accessed: Oct. 25, 2024. [Online]. Available: https://www.securityweek.com/ics-patch-tuesday-siemens-schneider-electric-cisa-issue-advisories/

[11] S. Soliman, W. Oudah, and A. Aljuhani, 'Deep learning-based intrusion detection approach for securing industrial Internet of Things', *Alexandria Eng. J.*, vol. 81, no. May, pp. 371–383, 2023, doi: 10.1016/j.aej.2023.09.023.

[12] G. Karacayılmaz and H. Artuner, 'A novel approach detection for IIoT attacks via artificial intelligence', *Cluster Comput.*, vol. 27, no. 8, pp. 10467–10485, 2024, doi: 10.1007/s10586-024-04529-w.

[13] R. Borgohain, 'FuGeIDS: Fuzzy Genetic paradigms in Intrusion Detection Systems', *arXiv Prepr. arXiv1204.6416*, pp. 1–7, 2012, [Online]. Available: http://arxiv.org/abs/1204.6416

[14] S. S. Hung and D. Shing-Min Liu, 'A user-oriented ontology-based approach for network intrusion detection', *Comput. Stand. Interfaces*, vol. 30, no. 1–2, pp. 78–88, 2008, doi: 10.1016/j.csi.2007.07.008.

[15] M. A. Jabbar and R. Aluvalu, 'Intrusion detection system for the internet of things: A review', *IET Conf. Publ.*, vol. 2018, no. CP747, 2018, doi: 10.1049/cp.2018.1419.

[16] M. Almseidin, J. Al-Sawwa, and M. Alkasassbeh, 'Anomaly-based Intrusion Detection System Using Fuzzy Logic', *2021 Int. Conf. Inf. Technol. ICIT 2021 - Proc.*, pp. 290–295, 2021, doi: 10.1109/ICIT52682.2021.9491742.

[17] P. Binnar, S. Bhirud, and F. Kazi, 'Security analysis of cyber physical system using digital forensic incident response', *Cyber Secur. Appl.*, vol. 2, no. December 2023, 2024, doi: 10.1016/j.csa.2023.100034.

[18] S. Age, C. Heng, F. Carrez, and K. Moessner, 'Journal of Network and Computer Applications A lightweight SEL for attack detection in IoT / IIoT networks', vol. 230, no. February, 2024.

[19] Y. Kayode Saheed, A. Idris Abiodun, S. Misra, M. Kristiansen Holone, and R. Colomo-Palacios, 'A machine learning-based intrusion detection for detecting internet of things network attacks', *Alexandria Eng. J.*, vol. 61, no. 12, pp. 9395–9409, Dec. 2022, doi: 10.1016/j.aej.2022.02.063.

[20] A. Al-Bakaa and B. Al-Musawi, 'A new intrusion detection system based on using non-linear statistical analysis and features selection techniques', *Comput. Secur.*, vol. 122, Nov. 2022, doi: 10.1016/j.cose.2022.102906.

[21] R. Alanazi and A. Aljuhani, 'Anomaly Detection for Industrial Internet of Things Cyberattacks', *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. 2361–2378, 2023, doi: 10.32604/csse.2023.026712.

[22] R. Gopi *et al.*, 'Intelligent Intrusion Detection System for Industrial Internet of Things Environment', *Comput. Syst. Sci. Eng.*, vol. 44, no. 2, pp. 1567–1582, 2023, doi: 10.32604/csse.2023.025216.

[23] M. Eskandari, Z. H. Janjua, M. Vecchio, and F. Antonelli, 'Passban IDS: An Intelligent Anomaly-Based Intrusion Detection System for IoT Edge Devices', *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6882–6897, 2020, doi: 10.1109/JIOT.2020.2970501.

[24] G. Abdelmoumin, D. B. Rawat, and A. Rahman, 'On the Performance of Machine Learning Models for Anomaly-Based Intelligent Intrusion Detection Systems for the Internet of Things', *IEEE Internet Things J.*, vol. 9, no. 6, pp. 4280–4290, 2022, doi: 10.1109/JIOT.2021.3103829.

[25] M. M. Alani and A. I. Awad, 'An Intelligent Two-Layer Intrusion Detection System for the Internet of Things', *IEEE Trans. Ind. Informatics*, vol. 19, no. 1, pp. 683–692, 2023, doi: 10.1109/TII.2022.3192035.

[26] Z. Lv, D. Chen, R. Lou, and H. Song, 'Industrial Security Solution for Virtual Reality', *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6273–6281, 2021, doi: 10.1109/JIOT.2020.3004469.

[27] M. Abdel-Basset, V. Chang, H. Hawash, R. K. Chakrabortty, and M. Ryan, 'Deep-IFS: Intrusion Detection Approach for Industrial Internet of Things Traffic in Fog Environment', *IEEE Trans. Ind. Informatics*, vol. 17, no. 11, pp. 7704–7715, Nov. 2021, doi: 10.1109/TII.2020.3025755.

[28] I. A. Khan, M. Keshk, D. Pi, N. Khan, Y. Hussain, and H. Soliman, 'Enhancing IIoT networks protection: A robust security model for attack detection in Internet Industrial Control Systems', *Ad Hoc Networks*, vol. 134, Sep. 2022, doi: 10.1016/j.adhoc.2022.102930.

[29] P. B M, N. G. M, and M. S. Hema, ‘Towards an effective deep learning-based intrusion detection system in the internet of things’, *Telemat. Informatics Reports*, vol. 7, Sep. 2022, doi: 10.1016/j.teler.2022.100009.

[30] J. Ahmad, S. A. Shah, S. Latif, F. Ahmed, Z. Zou, and N. Pitropakis, ‘DRaNN_PSO: A deep random neural network with particle swarm optimization for intrusion detection in the industrial internet of things’, *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8112–8121, Nov. 2022, doi: 10.1016/j.jksuci.2022.07.023.

[31] R. Ahmad, I. Alsmadi, W. Alhamdani, and L. Tawalbeh, ‘A comprehensive deep learning benchmark for IoT IDS’, *Comput. Secur.*, vol. 114, 2022, doi: 10.1016/j.cose.2021.102588.

[32] D. Nedeljkovic and Z. Jakovljevic, ‘CNN based method for the development of cyber-attacks detection algorithms in industrial control systems’, *Comput. Secur.*, vol. 114, Mar. 2022, doi: 10.1016/j.cose.2021.102585.

[33] M. M. Hassan, S. Huda, S. Sharmeen, J. Abawajy, and G. Fortino, ‘An Adaptive Trust Boundary Protection for IIoT Networks Using Deep-Learning Feature-Extraction-Based Semisupervised Model’, *IEEE Trans. Ind. Informatics*, vol. 17, no. 4, pp. 2860–2870, Apr. 2021, doi: 10.1109/TII.2020.3015026.

[34] A. Telikani, J. Shen, J. Yang, and P. Wang, ‘Industrial IoT Intrusion Detection via Evolutionary Cost-Sensitive Learning and Fog Computing’, *IEEE Internet Things J.*, vol. 9, no. 22, pp. 23260–23271, Nov. 2022, doi: 10.1109/JIOT.2022.3188224.

[35] A. Patcha and J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Computer Networks*, vol. 51, no. 12, pp. 3448–3470, 2007. DOI: 10.1016/j.comnet.2006.11.001

[36] M. Thaseen and C. A. Kumar, "Intrusion detection model using fusion of PCA and optimized SVM," *Procedia Computer Science*, vol. 78, pp. 26–31, 2016. DOI: 10.1016/j.procs.2016.02.006

[37] K. Kim et al., "Long Short-Term Memory Recurrent Neural Network Classifier for Intrusion Detection," *IEEE Access*, vol. 6, pp. 21954–21963, 2018. DOI: 10.1109/ACCESS.2018.2820191

[38] Q. Wang et al., "A deep reinforcement learning-based framework for lightweight intrusion detection in Industrial IoT," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14470–14480, 2021. DOI: 10.1109/JIOT.2021.3050451

[39] A. Lin et al., "Explainable Reinforcement Learning: A Survey," *Neurocomputing*, vol. 537, pp. 250–268, 2023. DOI: 10.1016/j.neucom.2023.03.003