Golden Sun-Rise
International Journal of Multidisciplinary on Science and Management
ISSN: 3048-5037/ Volume 1 Issue 4 December 2024 / Page No: 66-78
Paper Id: IJMSM-V1I4P109 / Doi:10.71141/30485037/V1I4P109

Original Article

Machine Learning Applications in Intrusion Detection: A Comprehensive Review

Anitha Mareedu

Electrical engineering Texas A&M University - Kingsville 700 University Blvd, Kingsville.

Received: 19 October 2024 Revised: 29 October 2024 Accepted: 09 November 2024 Published: 25 November 2024

Abstract: *The rapid growth of digital technologies and interconnected systems has significantly expanded the attack surface for cybercriminals, leading to a surge in sophisticated cyber threats such as advanced persistent threats (APTs), ransomware, and zero-day exploits. The existing scheme of IDS is the so-called Traditional Intrusion Detection Systems, based mainly on signature-based detection, but with the limitations concerning the capability to detect new attacks and the necessity to update them on a regular basis. In regard to these difficulties, Machine Learning (ML)-based IDS was observed as a potential paradigm and provided greater levels of both adaptability and scalability, as well as the capacity to identify threats even in instances when they have never been observed before. This survey investigates the use of ML in the detection of intrusions in a detailed manner that considers three fundamental aspects, namely, feature engineering, supervised-learning models, and benchmarking practices of IDS/IPS. It is discussed in the way of exploring the classic and new IDS datasets, methods of selecting and extracting features, and the effectiveness of the supervised algorithms directives, e.g., use of decision trees, use of support vector machines, and use of deep neural networks. It also singles out benchmarking tools such as Snort, Suricata, and Zeek, as well as principal evaluation metrics such as accuracy, precision, recall, and latency. Nevertheless, ML-based IDS still have serious problems, such as data imbalance, adversarial attacks, and implementation in real-time IoT and clouds. New research topics like federated learning to train privacy-preserving models, explainable AI to make models understandable and blockchain/quantum-resistant IDS designs are also brought up. Through the solutions of these issues and through the application of the latest advancements in technologies, ML-based IDS may become powerful and intelligent tools able to protect a network against the dynamic threat that continuously spreads on modern networks.*

Keywords: *Intrusion Detection System (IDS), Machine Learning, Feature Engineering, Benchmarking, Supervised Learning.*

I. INTRODUCTION

The dependence on digital technologies and interconnected systems that depend on them has substantially increased the attack surface of cybercriminals [1]. Modern networks are encountering a surge in the rate and complexity of attacks, such as advanced persistent threats (APTs), ransomware, and zero-day exploits [2]. Cybersecurity is a topic of what enterprise infrastructures worldwide should be concerned about since these threats also attack such vital areas as healthcare, finance, and the IoT ecosystem. In this new world of threat, Intrusion Prevention Systems (IPS) and Intrusion Detection Systems (IDS) are becoming an essential element in detecting and preventing attacks. IDS is dedicated to network traffic monitoring and analysis tasks in searching for anomalies [3] , whereas IPS advances its capability of blocking suspicious activity. Although critical, traditional practices to IDS have great shortcomings in attempting to tackle the sophisticated and dynamic cyber threats.

Traditional approaches use IDS solutions that are largely signature-based, in which the detection of an intrusion relies on predetermined signatures of attacks [4]. There is always a mismatch in these systems, as they effectively detect known threats but cannot detect zero-day attacks or some advanced adversarial methods. Moreover, signature database systems are less flexible to quick-evolving environments since updating them is a manual task that should be carried out frequently [5]. Such weaknesses have also contributed to the emergence of Machine Learning (ML)-based IDS, which proposes the provisions of flexibility and foreknowledge due to the data-driven solutions. With massive concurrency, a real-time IDS based on ML can be used to process huge traffic quantities, determine patterns that suggest malicious behaviour, and learn historical patterns of attacks to reduce false positivity. Due to this versatility, they are suitable for the modern cloud and distributed environments.

An ML-based IDS usually has the form of a multi-stage pipeline aimed at increasing detection efficiency[6]. To give a clear picture as represented in Figure 1, data collection starts with feature derivation on network sources; this is followed by feature engineering, where raw data is converted into attributes that signify the results in enhancing the model. Such features are subsequently trained on the supervised machine learning systems, and this helps the model to differentiate between the bad and harmful actions. A model, after being trained, works in conditions of real time, and it detects anomalies and launches alerts or automated response when necessary[7]. The pipeline shows how ML could be incorporated in IDS without any modification, making it more adaptable, scalable, and accurate [8].

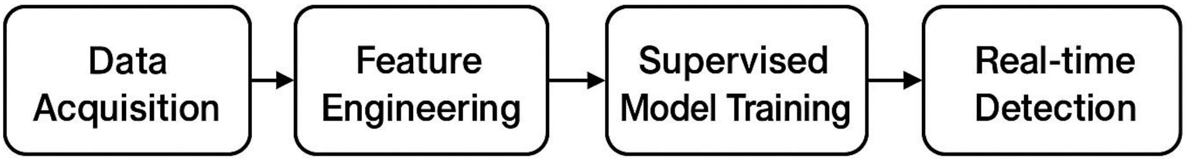

Figure 1. Architecture of a Machine Learning-based Intrusion Detection System (IDS) pipeline

A detailed review of the application of machine learning in intrusion detection is provided with consideration to the three main areas, namely feature engineering, supervised learning, and IDS/IPS benchmarking practices. It discusses some of the core concepts, popular datasets, recent algorithm developments, evaluation methods, and dynamic areas of research, such as federated learning, explainable AI, and compatibility with blockchain-based security.

II. FOUNDATIONS AND RELATED WORK

One of the most common components of cybersecurity has been the Intrusion Detection Systems (IDS) that have kept up with trends in the complexity of cyber threats. With the growth in scope of networks and the increase in the expertise of attacks, the research for IDS has evolved away from the static mindset of either a reactive or a rule-based approach towards a more dynamic data-driven approach. It is of vital importance to comprehend this evolution to place in the context the importance of machine learning (ML) in the current IDS design and determine the ways in which the intelligent systems target the correction of the limitations of their ancestors. This section provides an overview of the historical development of IDS, contrasting signature-based and anomaly-based approaches, and highlights the integration of ML for adaptive and scalable threat detection in contemporary environments.

A. Evolution of Intrusion Detection

The discipline of intrusion detection has experienced one paradigm: the transition of the initiative axioms of a very primitive intrusion defence, namely, to a more versatile framework[9]. Early implementations of IDS technology were rule-based, allowing expert analysts to code known signatures of an attack into detection systems. These systems were effective at defending against documented threats yet had few capabilities to defend against emerging attacks. As network protocols and attack vectors become increasingly diverse, researchers created anomaly-based IDS, which tried to profile the normal behaviour of networks and mark exceptions as possible intrusions.

Being innovative, however, this strategy also created problems with differentiating between malicious anomalies (e.g., populating a site with a significant traffic spike because of a genuine reason) and attacks, which resulted in high levels of false positives. Since then the emergence of big data and improvements in processing capabilities have led to the development of machine learning methods[10], allowing IDS to operate similarly to how IDS learn the complicated traffic and adapt to it in almost real time as well as identify threats that have never been seen before. Figure 1 presents the ML-based IDS pipeline, and now we explore the historically comparative discussion of these methods.

a) Signature-based IDS

Signature-based IDS finds the intrusion by comparing the observed traffic with the predefined attack signatures stored in the database [11]. Examples of this type, in widespread use, include Snort and Bro (now Zeek), algorithms that detect malicious payloads or port scans or protocol violations.

The advantages of this approach include high detection accuracy for known attacks and low false positive rates, making them suitable for environments where stability and predictability are paramount[12]. However, the limitations are equally significant:

- Inability to detect zero-day exploits or polymorphic malware.
- The database needs to be updated quite regularly to address the new threats.
- They can be evaded by small manipulations of the signatures of attacks.

These shortcomings necessitated alternative strategies for detecting novel attacks.

b) Anomaly-based IDS

Anomaly-based IDS overcome the problems of signature-based systems by taking models of normal system functioning and raising the possibility of an intrusion when there is a deviation [13]. Heuristic methods and clustering algorithms have been frequently involved, as have statistical methods.

The pros of anomaly-based IDS include their capacity to:

- Detects previously unseen attacks without requiring prior knowledge.
- Offer better protection in dynamic environments where threats evolve rapidly.

Such systems, however, have a tendency of high false positives since benign but unusual user activities (e.g. large file transfers) may cause alerts. Also, creating a resilient baseline in highly heterogeneous networks is not easy, disregarding the fact that attackers are able to execute slow or quiet attacks that can only be carried out within the normal traffic.

B. Machine Learning in IDS

Recent developments have also added support to machine learning in the design of IDS, which provides a data-backed mechanism for adaptive and scalable surveillance of threats. The models of ML can process large volumes of data within the network, find peculiarities that may be an indicator of an attack, and generate their detection levels. Supervised learning techniques, which include decision trees, support vector machines (SVM), and deep neural networks (DNN)[14], have demonstrated tremendous success in the classification of network traffic as either benign or malicious. In the meantime, the unsupervised and semi-supervised methods have the promise of working in situations where there is restricted labelled data.

The main differences between traditional and ML-based IDS are summarised in Table 1 below, which gives a comparison perspective on how ML has solved most of the issues that have plagued the design of IDS over the years.

Table 1. Evolution of Intrusion Detection Systems: From Rule-Based to Machine Learning-Based Paradigms

Approach	Working Principle	Advantages	Limitations
Signature-based	Matches traffic patterns to a database of known attack signatures	High accuracy for known attacks; low false positives	Cannot detect zero-day attacks; frequent updates required
Anomaly-based	Identifies deviations from established normal behavior	Detects unknown threats; adaptable to dynamic networks	High false positives; baseline definition is complex
ML-based	Learns patterns in network data using supervised/unsupervised learning	Detects novel and complex threats; scalable and adaptive	Needs large, quality datasets; explainability and adversarial robustness are concerns

ML-based IDS represents a crucial step towards proactive and adaptive threat management, but it also raises policy and governance challenges around explainability, trust, and security in AI-driven systems [15].

III. IDS DATASETS FOR MACHINE LEARNING RESEARCH

Machine Learning (ML)-based Intrusion Detection Systems (IDS) are highly reliant on the quality and the variety of datasets extracted to train and evaluate their effectiveness. A number of test datasets have been created over the years to simulate network traffic and attacks aimed at IDS research. Datasets on this are widely divided into simpler (legacy) and more modern (contemporary) collections, with the newer ones being optimised to work in an IoT and cloud setup. Although the resources have brought about a great deal of improvement, they also have their underlying problems, like class imbalance, the lack of realism, and the inability to model encrypted traffic.

A. Classic Datasets

a) KDDCup99 and NSL-KDD

One of the first and the most commonly used benchmarks in the study of IDS research was the KDDCup99 dataset, created during the DARPA intrusion detection evaluation program [16]. It has huge amounts of the network traffic data either in its normal state or in one of several types of attacks, such as Denial of Service (DoS), probe, User-to-Root (U2R), and Remote-to-Local (R2L), among others.

Despite its historical importance, KDDCup99 exhibits significant drawbacks:

- Redundant records that bias ML algorithms.
- Outdated attack patterns that no longer reflect modern network environments.
- Insufficient representation of encrypted or obfuscated traffic.

To reduce some of these problems, the NSL-KDD dataset was proposed as a better one. NSL-KDD eliminates repeating records and equalises the proportions of geometric abuses so that the model will be trained better. Nevertheless, it does not reflect the complexity and heterogeneity of real-life traffic, which restricts the use of this tool in scenarios of checking up on a sophisticated IDS solution.

B. Modern Benchmark Datasets

a) UNSW-NB15

UNSW-NB15 has a more realistic combination of benign and malicious network traffic. It is generated with IXIA PerfectStorm tool and contains modern forms of attacks, including fuzzers, worms, shellcode, and exploits [17]. UNSW-NB15 is well-equipped with features that aid in the testing of ML models on varying types of traffic as well as multi-class classification issues.

b) CICIDS2017

CICIDS2017 is an initiative that was developed by the Canadian Institute of Cybersecurity and is described as a realistic model of real-world networks and incorporates various attack possibilities such as botnets, distributed denial of service (DDoS), Heartbleed, and infiltration attempts [13]. The dataset includes timestamped flows and application-layer features, which is especially helpful in the development of ML models to solve high-dimensional, time-sensitive traffic.

C. Dataset Challenges

While these datasets have advanced IDS research, they are not without challenges:

a) Class Imbalance

Observably, benign traffic prevails in most datasets and there are very few attack instances[18]. Such imbalance may bias ML models such that there is a low rate of detecting rare and critical intrusions.

b) Dynamic Attack Patterns

Cyber threats change on a daily basis and current data repositories usually do not keep pace with the situation in the actual world [19]. It should be updated periodically or ambiguous traffic should be generated to stay relevant.

c) Encrypted Traffic

Increasing encryption use (e.g., TLS 1.3, HTTPS) in general limits visibility of the packet contents, making feature extraction and detection of attacks more difficult [20].

Figure 2 below shows a taxonomy of the IDS datasets based on the number of legacy, contemporary, and IoT/cloud-orientated datasets to demonstrate the trend of how the benchmark datasets evolved in the field of intrusion detection.

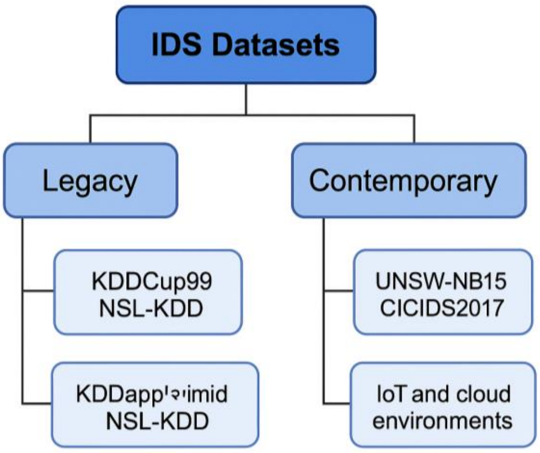

Figure 2. Taxonomy of IDS datasets by category. Legacy datasets

Recently, researchers have subscribed to the idea that high-quality and scalable datasets are essential to training ML models successfully. Specifically, when dealing with distributed environments such as IoT and edge computing, the size of big data network data has to be looked upon with great control regarding data caching and the distribution of dataset patterns to address the issues of bandwidth and data storage limitations. All these points are important when it comes to the creation of efficient and flexible IDS solutions that can be deployed in real life scenarios [25].

IV. FEATURE ENGINEERING IN IDS

In the structural development of Machine Learning (ML) Intrusion Detection Systems (IDS), feature engineering is an important aspect. It also entails the process of translation of raw network traffic data into informative and meaningful attributes that greatly promote the capacity of ML models in identifying malicious operations. The ill mechanisms of features may result in underfitting, overfitting, and generalisation of the network in various network environments. By contrast, well-designed features enhance the accuracy, robustness, and compute efficiency of models. The significance of feature engineering is discussed in this section and the essential features of feature selection and extraction are followed. The impact that feature selection and extraction has on the performance of ML-driven IDS is also mentioned.

A. Feature Engineering

In specialised network security applications, particular raw traffic data may carry many more irrelevant or duplicate fields, such as redundant protocol headers, tormented timestamps, or statistical chaff. Having these irrelevant features may confuse ML algorithms, resulting in higher false positives and poor ability to capture unseen patterns of attacks [21]. Good feature engineering will result in retention of the most discriminative attributes and elimination of others, thus reducing dimensionality and calculation complexity and enhancing detection. As an illustration, within the IDS datasets like the CICIDS2017, researchers have demonstrated that choosing the flow-based metrics (e.g., average packet size, flow duration, and inter-arrival times) can perform better than giving raw information at the packet level. Also, to detect such subtle behavioral patterns of threats, domain-specific attributes created in feature engineering may be useful to capture them.

B. Feature Selection Techniques

Feature selection refers to the procedure of defining and keeping the best applicable features of a set of datum and removing monotonous and noisy features. This will help in enhancing the accuracy of the IDS detection and finally be scalable [22]. The most common method of feature selection falls under three categories: filter methods, wrapper methods and embedded methods.

a) Filter Methods

Filter methods rank features based on statistical measures, independent of any ML algorithm. Common examples include:

- **Information Gain (IG):** Measures the reduction in entropy when a feature is known.
- **Chi-square Test:** Evaluates the dependency between categorical features and class labels.

These methods are computationally efficient and suitable for high-dimensional datasets. However, they may overlook interactions between features, which can limit their effectiveness in complex IDS scenarios.

b) Wrapper Methods

Wrapper methods evaluate different subsets of features using a specific ML model to determine the optimal feature set. A notable technique is:

- **Recursive Feature Elimination (RFE):** Iteratively removes the least important features based on model performance until an optimal subset is achieved.

Although wrapper methods often deliver higher accuracy than filter methods, they are computationally intensive, especially for large-scale IDS datasets.

c) Embedded Methods

Embedded methods incorporate feature selection directly into the model training process. Examples include:

- **Lasso (L1 Regularization):** Encourages sparsity in feature weights by penalizing the absolute magnitude of coefficients.
- **Tree-Based Selection:** Decision trees and ensemble methods (e.g., Random Forest) naturally rank features by their contribution to classification decisions.

These methods provide a balance between computational efficiency and detection accuracy, making them well-suited for IDS applications. Table 2 highlights and compares these feature selection approaches with commonly used feature extraction techniques, summarizing their strengths and impact on IDS performance.

Table 2. Feature selection and extraction techniques in ML-based IDS and their impact on detection accuracy

Technique Type	Methods	Impact on IDS Accuracy
Filter Methods	Information Gain, Chi-square	Fast but may overlook feature interactions; moderate gains
Wrapper Methods	Recursive Feature Elimination (RFE)	High accuracy; computationally intensive
Embedded Methods	Lasso (L1 Regularization), Random Forest	Good balance of accuracy and efficiency
Dimensionality Reduction	PCA, Autoencoders	Effective for high-dimensional data; aids anomaly detection

C. Feature Extraction Techniques

Feature extraction differs from selection in that it creates new features by transforming the original variables into a reduced-dimensional space while retaining critical information. This is particularly valuable for IDS datasets that are high-dimensional and complex.

a) Dimensionality Reduction

Two widely used dimensionality reduction techniques in IDS research include:

- **Principal Component Analysis (PCA):** Projects original features onto orthogonal components that capture the maximum variance in data. PCA is useful for simplifying datasets while preserving essential information.
- **Autoencoders:** Neural networks designed to reconstruct inputs through a bottleneck layer, enabling them to learn compressed representations of data. Autoencoders have shown effectiveness in capturing hidden patterns and detecting subtle anomalies in network traffic.

Effective feature engineering, as outlined above and summarised in Table 2, is crucial for developing IDS capable of not only detecting known attack patterns but also generalising to novel and evolving cyber threats in dynamic network environments.

V. SUPERVISED LEARNING IN INTRUSION DETECTION

Supervised learning has come out as one of the strong paradigms in the design of Machine Learning (ML)-based Intrusion Detection Systems (IDS). In supervised models, the datasets are labelled with both good and bad traffic data and the algorithm is trained on this set of data to determine future network activity. Such models have proven to be extremely accurate at both detecting known and previously unobserved attack patterns when trained on good data. This section will give the overview of supervised models in IDS and repeat about the principal families of algorithms currently in use along with the often-used performance indicators to help quantify how effective each model is.

A. Description of supervised Models

Supervised learning is based on the dataset in which every case is given its label: normal or presenting a particular kind of attack. These labels are vital in training the IDS models due to the fact that they give the algorithm an opportunity to train on the unique patterns that distinguish between genuine traffic and malicious traffic. Labelled data can be used to form decision boundaries in feature space, upon which the classification of new, never-seen-before traffic occurs as propagated within the supervised algorithms [23] . But high-quality labelled data is difficult to come by, as manual annotation is expensive as well as the privacy and dynamic nature of attack techniques. This has inspired the development of benchmark datasets, e.g., NSL-KDD and CICIDS2017, with structured labelled data to train and test.

B. Family of Algorithms

Supervised algorithms in IDS may be classified into three major categories, including tree-based models, margin-based models, and neural networks, all of which imply different trade offs in terms of accuracy and interpretability and running time.

a) Tree-Based Models

Tree-based methods are among the most widely used algorithms for IDS because of their interpretability and strong performance:

- **Decision Trees:** These models split the dataset into subsets based on feature thresholds, forming a hierarchical structure that is easy to interpret.
- **Random Forests:** An ensemble of decision trees that improves classification performance by averaging predictions and reducing overfitting. Random Forests are particularly effective for high-dimensional IDS datasets such as CICIDS2017, providing robust detection capabilities across multiple attack categories.

b) Margin-Based Models

Support Vector Machines (SVM): These models find the optimal hyperplane that maximises the margin between benign and malicious classes in feature space [23]. SVMs are highly effective in binary classification scenarios and can handle non-linear patterns using kernel functions. However, their performance can degrade with very large IDS datasets due to computational demands.

c) Neural Networks and Deep Learning

Deep learning approaches have gained prominence in IDS research [26]:

- **Artificial Neural Networks (ANN):** Multi-layer perceptrons capable of capturing complex relationships in network traffic.
- **Convolutional Neural Networks (CNN):** Used to identify spatial patterns in traffic data, such as packet sequences or byte-level features.
- **Recurrent Neural Networks (RNN):** Designed to capture temporal dependencies, making them well-suited for time-series IDS data.

These models excel in environments with large-scale traffic and complex attack vectors but require significant computational resources and large labelled datasets for effective training.

C. IDS Performance Metrics

Evaluating the performance of supervised IDS models requires metrics that capture various aspects of classification accuracy:

- **Accuracy:** The proportion of correctly classified instances out of all instances.
- **Precision:** The ratio of true positives to all instances classified as positive, reflecting the model’s ability to avoid false alarms.
- **Recall (Sensitivity):** The ratio of true positives to all actual positive instances, indicating how well the model detects attacks.
- **F1-Score:** The harmonic mean of precision and recall, providing a balanced metric for imbalanced datasets

a) Area under the Curve (AUC)

Measures the trade-off between true positive and false positive rates across classification thresholds. These metrics are essential for understanding not only the overall accuracy of an IDS but also its ability to detect rare and critical attacks in real-world network environments. The choice of algorithm and evaluation metric must align with the operational constraints and risk tolerance of the target environment, as no single model excels universally across all attack scenarios [27].

VI. IDS/IPS BENCHMARKING AND EVALUATION

The critical aspect in the assessment of Intrusion Detection and Prevention Systems (IDS/IPS) is benchmarking. Its benchmarking offers a common platform on how to gauge the effectiveness, efficiency, and scalability of such systems with conditions that could be controlled and reproduced. In the absence of adequate benchmarking, it is very difficult to compare various IDS/IPS implementations and thus there will be fragmented research and researchers with little scope of applying them to real-world scenarios. It explains why the standardised benchmarks are necessary, discusses the popular IDS/IPS tools and identifies the important metrics that can be applied to measure their performance.

A. *Need for Standardized Benchmarks*

With the development of the IDS technologies, the following issue of reproducibility and comparability of research is relevant. The benchmarking framework is standardised so that researchers and practitioners can provide a fair assessment of the IDS solutions on various datasets, algorithms, and traffic. Variability in benchmarking procedures does not enable researchers to justify reported increases in detection performance due to changes in algorithms or in experimental conditions. Structured datasets to evaluate against are available as benchmarks like NSL-KDD, CICIDS2017, etc., and popular system-level benchmarks include Snort, Suricata, and Zeek.

B. *Common Benchmarking Tools*

The solutions to benchmarking of the IDS/IPS systems necessitate potent tools that can recreate different network settings and give the most precise assessments of intrusion detection systems functioning in highly differentiated traffic conditions and scenarios of attack. A number of open-source tools are especially popular and used both in academic studies and industrial implementations because of their flexibility and full-featured resource set as well as community support. Among them, one can distinguish such dominating platforms as Snort, Suricata, and Zeek (or Bro), which have enormously influenced IDS benchmarking processes.

a) Snort

Snort is an early and most generally used open-source intrusion detection system with signatures. One example is Snort, developed by Sourcefire and maintained under Cisco, which analyses traffic and logs packets in real-time in an attempt to detect known attacks by matching against predefined rules [28]. Its modular structure enables security researchers and practitioners to design and swap their own rules, and it has a huge and actively upgraded signature store. Snort has been a capability-of-choice among those organisations that wish to deploy a reliable IDS on traditional network deployments given this comprehensive rule base.

Although Snort is highly effective at detecting known threats, the restriction of Snort to predetermined signatures makes it ineffective at detecting zero-day exploits and highly clever variations of the attacks that do not correlate with the preset signatures. In addition, performance of Snort could also be an issue in high-performance or large-scale network systems that are connected together, as single-threaded architecture could cause a bottleneck and it thus needs to be optimised well.

Strengths:

- Mature and well-documented with extensive community support.
- Rich signature database covering numerous attack types.
- Effective for detecting known threats with low false positive rates.

Limitations:

- Limited capability to detect zero-day attacks or polymorphic malware.
- Requires frequent signature updates to remain effective.
- Scalability challenges in high-throughput networks.

b) Suricata and Zeek

In order to overcome some of the drawbacks of the traditional IDS, such as Snort, new open-source technologies such as Suricata and Zeek have emerged. The tools include sophisticated capabilities, and hence they are highly adapted to be used in ML-based IDS benchmarking under modern network conditions [29]. Suricata is a high-performance IDS/IPS engine that is capable of multi-threading over packet processing; it can more easily process gigabit-scale traffic when compared to Snort. It integrates signature-based detection with protocol analysis and supports intrusion prevention capability, which makes it a very flexible real-time monitor and response tool on enterprise networks. Its compatibility with new protocols and use of deep packet inspection also make Suricata even more applicable in infrastructures with much complexity nowadays.

Zeek (formerly Bro), in its turn, is based on the completely different approach of network behaviour and traffic analysis instead of signature matching only. It supports user-defined detection logic in its powerful scripting language, allowing researchers to define policy on anomaly detection and application-layer inspection. Zeek is highly capable of supplying context-rich information about network events and therefore finds a lot of use in forensic cases as well as in complex attack behaviour analysis.

Suricata:

- Multi-threaded architecture enabling high-speed packet processing.
- Combines IDS and IPS features with support for modern protocols.
- Offers efficient real-time detection in large-scale networks.

Zeek:

- Focuses on deep network traffic analysis and application-layer inspection.
- Highly customisable through its scripting language for creating advanced detection policies.
- Better suited for anomaly detection and behavioural analytics.

Figure 3 provides a visual comparison of these IDS/IPS benchmarking tools, summarizing their core capabilities, strengths and limitations.

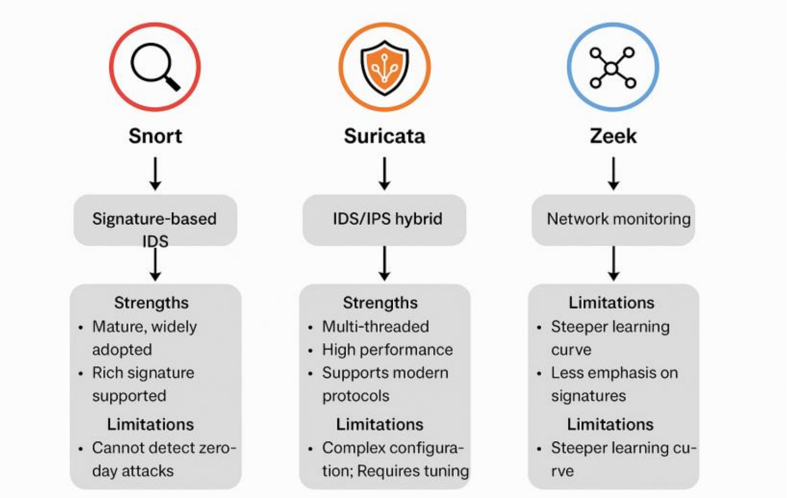

Figure 3. Core capabilities and comparison of IDS/IPS benchmarking tools

C. Benchmarking Metrics

Several metrics are commonly used to evaluate the performance of IDS/IPS tools:

- **Detection Rate (True Positive Rate):** Measures the system's ability to correctly identify actual attacks.
- **False Positives:** Indicates the frequency of benign traffic being incorrectly flagged as malicious.
- **Latency:** Reflects the time taken to detect and respond to an intrusion, which is critical in real-time systems
- **Throughput:** The volume of traffic the system can analyse without performance degradation.
- **Resource Utilization:** Evaluates CPU, memory, and storage consumption during operation.

These metrics provide a holistic view of IDS/IPS performance, balancing detection effectiveness with operational efficiency. As modern IDS benchmarking must also account for cloud-native deployment challenges, including elasticity, container orchestration, and integration with distributed infrastructures [30].

VII. CHALLENGES IN ML-BASED IDS

Although, Machine Learning (ML)-based Intrusion Detection Systems (IDS) have posed considerable promise in terms of improving cybersecurity, this has come with a number of issues that stand in the way of employing them in practice and their success. Such problems are caused by the character of data, by ML models limitations, as well as by the difficulties of practical work on IDS solutions introduction in large-scale environments and in dynamic conditions. It is important to deal with these challenges to enhance ML-based IDSs beyond experimental prototypes to real-life systems. In this section, important challenges related to data, and models and ML-based IDS deployment are discussed.

A. Problems in Data

a) Imbalanced Datasets

Class imbalance is presented by most IDS datasets, where benign traffic is very large compared to malicious cases. This unbalance makes the ML algorithms favor the majority thus having low detection rates of infrequent but serious attacks like zero-day exploits or invisible breaches.

b) Lack of tagged Data

The ML methods that are supervised are done through the utilization of many huge datasets of superior valuable labeled information. Nonetheless, it is difficult to obtain such datasets in the real-world scenarios since there are concerns regarding privacy, labeling expenses, and the ever-changing nature of cyber threats. To reduce such issues, synthetic data generation and semi-supervised learning methods are suggested, which bring about their specific complexities.

B. Modeling Problems

a) Overfitting

ML models, especially the deep learning architecture, are highly affected by overfitting on training sets especially when they are small or not indicative of actual traffic patterns. Models that have been overfit thus fail and respond badly to untested attacks or novel network states.

b) Adversarial Attacks on ML Models

Adversarial attacks also apply to ML-based IDS, where the attackers manipulate the inputs (craft) to insinuate a false positive detection in the system. Examples are evasion attacks, where malicious traffic is altered to look benign and poisoning attacks, where training data is poisoned so as to decrease the model's efficacy. The robust and explainable ML model design is still the research issue.

C. Deployment Problem

a) Real-Time Detection

This entails the use of ML-based IDS in a real-time environment, which mandates the need to have models that handle high traffic at a low latency. Most of the algorithms with high complexities have difficulty in achieving these performance requirements without some high computational means.

b) Scalability IoT/ Cloud Environments

The problems related to modern networks, especially those that are IoT and cloud-based ones, present new difficulties like distributed design, heterogeneous devices, and encrypted traffic. IDS based on ML should be scalable and flexible to be able to monitor such environments and not overload the resource potential and not create bottlenecks. This has to do with interdisciplinary work that takes the form of ML algorithm improvements, data engineering, and systems design. Development of approaches like federated learning, adversarial robustness, and lightweight models of the IDS to edge computing that cope with these constraints is a promising area.

VIII. FUTURE RESEARCH DIRECTIONS

Currently, the cyber threat is increasing both in size and complexity, posing the necessity to enhance Machine Learning (ML) Intrusion Detection Systems (IDS) to be more adaptive, scalable, and trustworthy. The existing weaknesses of these systems, including data privacy issues, the absence of interpretability, etc., have encouraged scholars to come up with new methods and combine new technologies. Next-generation IDS should not only prevent sophisticated attacks with high precision but also perform viably in the decentralised, restricted, and dynamic settings.

Federated learning (FL) is one of the promising research directions in terms of developing IDS. Thereby, FL allows training the model on distributed devices and organisations without relaying raw traffic information to a central classifier. The specified method of privacy protection is especially applicable in the context of the Internet of Things (IoT) and multi-cloud settings where one cannot openly share sensitive data due to privacy laws or company policies. With the help of FL, IDS solutions are capable of attaining collective intelligence without compromising the confidentiality of the data and the exposure risk. Nevertheless, FL applied in the IDS also has disadvantages like communication overhead, heterogeneous data distributions, and the susceptibility to poisoning attacks, and these aspects need to be discussed in the future.

A second major direction that needs investigation is how explainable artificial intelligence (XAI) should be inserted into IDS. Although traditional ML and deep models usually have high detection accuracy, they are black-box, minimising interpretability and explainability. XAI methods attempt to close this divide through offering explanations of the decisions made by IDS, which can be understood by human beings. As an example, a feature attribution technique can indicate what features of the network were most useful in assigning a flow as being malicious, thereby enabling the security analysts to test and be assured with system results. As well as promoting trust in the user, explainability is crucial to being able to address necessary compliance aspects of regulated industries, including healthcare, finance, and critical infrastructure.

Besides the ML innovations, emerging technologies integration can radically change the capabilities of IDS solutions. IDS logs stored on blockchains can be unalterable and immutable records of intrusion and can be used to improve accountability and forensics analysis and investigation after a breach. Moreover, as quantum computers will be coming in the future, researchers are already starting to investigate the idea of quantum-resistant IDS designs that would defend themselves against attacks that potentially use quantum algorithms to attack cryptographic defences. The farsighted methods state the necessity of the forward-compatible and robust IDS designs that can work in the context of extremely dynamic technological change. Combined, these future directions of research point to a paradigm transformation in IDS advancements, that is, to intelligent, privacy-aware, explainable, and technologically synergistic systems. These emerging directions are very important in order to address IDS capable of solving security challenges of a technically connected world in the future.

IX. CONCLUSION

ICS-based computer intrusion detection systems have also transformed the management of cybersecurity threats due to their intelligent and flexible solutions that address the evolving and effective threats. In contrast to the classical IDS solutions based on the static signature-based detection, they enable the processing of huge amounts of data of the heterogeneous traffic of the networks, the detection of complex attack patterns, and even real-time adaptation in response to previously unknown threats. The current review has undertaken an in-depth study of the IDS technologies development, feature engineering to improve detection performance, and the application of supervised learning algorithms with an emphasis on properly evaluating the latter using benchmarking tools. Investigating old well-known examples of datasets, like KDDCup99, and recent benchmarks, like CICIDS2017 and UNSW-NB15, we demonstrated how data quality and diversity directly determine the success of ML models. The platforms that are in the benchmarking, including Snort, Suricata, and Zeek, were also touched upon as the platforms essential in encompassing the performance of the IDS under circumstances that are realistic and reproducible amid an attempt to ensure the solutions provided are viable in the deployment of the modern complex infrastructure in networks these days.

Nonetheless, implementation of ML-based IDS is likely to be a tricky business, and areas of challenges that need to be countered should be tackled to ensure their potential is achieved. The model training and evaluation are complicated due to data issues, such as severely skewed datasets, lack of labelled sample sets of attacks and encryption of the majority of traffic. Operationally effective limits are hindered by model-specific property, including overfitting, generalisation to dynamic environments and susceptibility to adversarial attacks. Moreover, the implementation of these systems in environments where they would have to operate in practice, such as IoT environments and multi-cloud environments, involves the limitation of latency, resource consumption, and scalability. The requirement of explainable artificial intelligence (XAI) in making IDS decisions transparent and interpretable by the security analysts was also stressed upon in this review, which is essential in providing trust to security analysts and regulatory compliance required in sensitive areas like the health care and finance industries.

Moving forward, the unification of the new technologies has its potential solutions that can be applied to such problems and propel the potential of ML-based IDS. Federated learning offers a privacy-conservative training system of distributed IDS into the various domains without presenting personal data or information. IDS logging can be based on blockchain to provide tamper-resistant records, which enhance auditability and forensics and quantum-resistant IDS architectures are under consideration to be ready when quantum computing causes disruption. Future studies should also work on the lightweight and energy-constrained model that can be used in low-resource environments, e.g., edge devices and IoT networks. Academia, industry, and policymakers will play critical roles in streamlining datasets, measures of evaluation, and deployment. When such technical and systemic challenges are tackled, ML-based IDS stands a chance of being developed as powerful, expandable, and reliable lines of defence that may be used to secure the ever-growing networked digital environment against the ever-changing cyberattacks.

X. REFERENCES

[1] D. Dimitrov and W. Willian, “The impact of the advanced technologies over the cyber attacks surface,” in Computer Science On-line Conference, Cham: Springer International Publishing, 2020.

[2] I. Stellios, P. Kotzanikolaou, and M. Psarakis, “Advanced persistent threats and zero-day exploits in industrial Internet of Things,” in Security and Privacy Trends in the Industrial Internet of Things, Cham: Springer International Publishing, 2019, pp. 47–68.

[3] K. Coulibaly, “An overview of intrusion detection and prevention systems,” arXiv preprint, arXiv:2004.08967, 2020.

[4] K. I. Iyer, “From signatures to behavior: Evolving strategies for next-generation intrusion detection,” Eur. J. Adv. Eng. Technol., vol. 8, no. 6, pp. 165–171, 2021.

[5] M. Weqar, S. Mehfuz, and D. Gupta, “Autonomous device discovery for IoT: Challenges and future research directions,” in Internet of Things, Chapman and Hall/CRC, 2023, pp. 257–276.

[6] M. Verkerken et al., “A novel multi-stage approach for hierarchical intrusion detection,” IEEE Trans. Netw. Serv. Manag., vol. 20, no. 3, pp. 3915–3929, 2023.

[7] F. Van Wyk et al., “Real-time sensor anomaly detection and identification in automated vehicles,” IEEE Trans. Intell. Transp. Syst., vol. 21, no. 3, pp. 1264–1276, 2019.

[8] J. Jangid and S. Dixit, The AI Renaissance: Innovations, Ethics, and the Future of Intelligent Systems, vol. 1. Technoscience Academy, 2023.

[9] M. G. Yaseen and A. S. Albahri, “Mapping the evolution of intrusion detection in big data: A bibliometric analysis,” Mesopotamian J. Big Data, 2023, pp. 138–148.

[10] A. A. Aburomman and M. B. I. Reaz, “A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems,” Inf. Sci., vol. 414, pp. 225–246, 201.

[11] Y. Otoum and A. Nayak, “As-ids: Anomaly and signature based ids for the internet of things,” J. Netw. Syst. Manag., vol. 29, no. 3, p. 23, 2021.

[12] A. Khraisat and A. Alazab, “A critical review of intrusion detection systems in the internet of things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges,” Cybersecurity, vol. 4, no. 1, p. 18, 2021.

[13] Z. K. Maseer et al., “Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset,” IEEE Access, vol. 9, pp. 22351–22370, 2021.

[14] M. Dua, “Machine learning approach to ids: A comprehensive review,” in Proc. 3rd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA), 2019.

[15] Asharf, J., Moustafa, N., Khurshid, H., Debie, E., Haider, W., & Wahab, A. (2020). *A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, Solutions and Future Directions.* Electronics, 9(7), 1177. https://doi.org/10.3390/electronics9071177

[16] D. D. Protić, “Review of KDD Cup ‘99, NSL-KDD and Kyoto 2006+ datasets,” Vojnotehnički Glasnik/Mil. Tech. Courier, vol. 66, no. 3, pp. 580–596, 2018.

[17] N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” in 2015 Mil. Commun. Inf. Syst. Conf. (MilCIS), IEEE, 2015.

[18] D. Ramyachitra and P. Manikandan, “Imbalanced dataset classification and solutions: A review,” Int. J. Comput. Bus. Res. (IJCBR), vol. 5, no. 4, pp. 1–29, 2014.

[19] A. Wang et al., “A data-driven study of DDoS attacks and their dynamics,” IEEE Trans. Dependable Secure Comput., vol. 17, no. 3, pp. 648–661, 2018.

[20] S. Rezaei and X. Liu, “Deep learning for encrypted traffic classification: An overview,” arXiv preprint, arXiv:1810.07906, 2018.

[21] F. Nargesian et al., “Learning feature engineering for classification,” in Proc. Int. Joint Conf. Artif. Intell. (IJCAI), vol. 17, 2017.

[22] S. Visalakshi and V. Radha, “A literature review of feature selection techniques and applications: Review of feature selection in data mining,” in 2014 IEEE Int. Conf. Comput. Intell. Comput. Res., 2014.

[23] A. Tharwat and W. Schenck, “A survey on active learning: State-of-the-art, practical challenges and research directions,” Mathematics, vol. 11, no. 4, p. 820, 2023.

[24] V. Jakkula, “Tutorial on support vector machine (SVM),” School of EECS, Washington State Univ., vol. 37, no. 2.5, pp. 3, 2006.

[25] J. Jangid, “Efficient training data caching for deep learning in edge computing networks,” Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., vol. 7, no. 5, pp. 337–362, 2020, doi: 10.32628/CSEIT20631113.

[26] C. C. Aggarwal, Neural Networks and Deep Learning, vol. 10, no. 978. Cham: Springer, 2018.

[27] S. Dixit, “AI-powered risk modeling in quantum finance: Redefining enterprise decision systems,” Int. J. Sci. Res. Sci. Eng. Technol., vol. 9, no. 4, pp. 547–572, 2022, doi: 10.32628/IJSRSET221656.

[28] W. Park and S. Ahn, “Performance comparison and detection analysis in snort and suricata environment,” Wireless Pers. Commun., vol. 94, no. 2, pp. 241–252, 2017.

[29] A. Waleed, A. F. Jamali, and A. Masood, “Which open-source IDS? Snort, Suricata or Zeek,” Comput. Netw., vol. 213, p. 109116, 2022.

[30] Tait, K.-A., Sher Khan, J., Alqahtani, F., Shah, A. A., Khan, F. A., Ur Rehman, M., Boulila, W., & Ahmad, J. (2021). *Intrusion Detection using Machine Learning Techniques: An Experimental Comparison.* arXiv. https://doi.org/10.48550/arXiv.2105.13435