A Literature Review on Lightweight AI-Based Intrusion Detection Systems (LAIDS) for Resource-Constrained Networks

Claire Campbell
University of Cape Town
Cape Town, Western Cape, South Africa

ABSTRACT

The rapid increase of Internet of Things (IoT) devices and wireless sensor networks (WSNs) has introduced significant cybersecurity challenges, particularly for resource-constrained networks. Traditional Intrusion Detection Systems (IDSs) are struggling to keep up with evolving cybersecurity threats, and AI-driven IDS models, while more adaptable and accurate, are resource-intensive and unsuitable for constrained environments. This literature review reveals that Lightweight AI-IDS (LAIDS) offer a promising solution by balancing security effectiveness with computational efficiency. By analysing existing research on lightweight models, the review demonstrates the potential of hybrid approaches that combine machine learning (ML) and deep learning (DL) techniques to enhance detection while minimising computational costs. The findings emphasise the need for scalable, efficient, and adaptive IDS solutions tailored to low-power environments, enabling real-time cybersecurity for IoT and other resource-constrained networks.

KEYWORDS

Intrusion Detection System, Artificial Intelligence, Machine Learning, Deep Learning

1 INTRODUCTION

The rapid expansion of the Internet of Things (IoT) has led to a vast number of computationally-limited devices constantly transmitting data across networks [37]. The ever-changing topology of IoT networks due to nodes joining and leaving the network makes them vulnerable to security threats [8]. The efficacy of security defenses is largely impacted by IoT device limits and the devices' lack of defense mechanisms. This highlights the need for developing resource-efficient security protocols.

Network Intrusion Detection Systems (IDSs) have been traditionally deployed to monitor computer systems and detect malicious activity [70]. IDSs are categorised into two main types: Signature-based IDS and Anomaly-based IDS. Signature-based techniques detect intrusion by comparing known patterns, while anomaly-based techniques monitor the current device's activities to identify any disturbances [61]. Traditional IDSs may not be able to adapt to the growing sophistication of modern cyberthreats.

The success of machine learning across various fields has led to its adoption in enhancing IDS performance [37]. IDSs can employ machine learning algorithms to evaluate network traffic and identify unusual behavior, learning from large datasets to find patterns which may be too complicated for conventional IDS systems [38]. Machine learning techniques used in IDSs can be categorised into three types: supervised, unsupervised, and semi-supervised learning. AI-powered IDSs have demonstrated superior accuracy and adaptability in detecting sophisticated network intrusions compared to traditional methods [64]. However, the increased complexity of these models requires more computing resources, which many resource-constrained networks cannot handle.

As the number of computationally-limited devices continues to grow, securing resource-constrained networks becomes increasingly difficult [5]. Traditional IDS solutions struggle with evolving cyber-threats, while AI-powered IDSs, despite their advantages, require significant computational resources that resource-constrained networks, like IoT environments, often lack [30]. This creates a crucial gap in cybersecurity: the need for a lightweight AI-driven IDSs that balances efficiency and security [5]. This literature review aims to explore optimised IDS models that leverage machine learning techniques while remaining feasible for resource-constrained networks.

First, traditional IDS approaches, including Signature-Based and Anomaly-Based methods, will be examined. AI-IDS methods will be discussed, broken into machine learning (ML) and deep learning (DL) techniques. The review will examine research focused on the optimisation of IDS models designed to address computational constraints related to memory usage, processing power, storage, and bandwidth. Established evaluation metrics and datasets will be identified followed by an evaluation of AI-IDS models leading to the proposal of the Lightweight AI-IDS (LAIDS) project. Finally, the review will summarise the key findings and outline the proposed LAIDS framework.

2 TRADITIONAL INTRUSION DETECTION SYSTEMS

Intrusion Detection Systems monitor network traffic and system activity, generating alerts when they detect potential security threats. There are two types of IDSs, Network IDSs (NIDSs) and Host IDSs (HIDSs). NIDSs are strategically placed to analyse the traffic of the entire network in order to detect any malicious activity. HIDSs analyse traffic at the host level, monitoring system logs and activities for signs of intrusion. IDS detection methods fall into two main categories: signature-based and anomaly-based detection. Additionally, hybrid detection, which combines both methods, will be discussed along with the method's respective advantages and disadvantages.

2.1 Signature-based IDS

Signature-based detection identifies intrusions by matching network activity with a database of predefined attack patterns. This detection type can detect an attack that has one of the signatures

listed in the IDSs. Every attack has specific signatures such as the source IP address or the payload of the packet [64]. The approach has a high accuracy in detecting known attacks and results in a low false positive rate [56]. However, signature-based detection is ineffective against unknown or advanced threats [30]. The signature database also requires constant updates to keep the system in line with new security threats [30].

2.2 Anomaly-based IDS

Anomaly-based detection defines a baseline state of the network and deviations from this are reported as potential attacks. Baselines are formed using statistical analysis and aim to be similar to common network traffic, such as services used by each host, the volume of activity during the day or specified times of activity [64]. This means that this approach is capable of detecting unknown or advanced attacks. Thus is more adaptable to evolving network use compared to signature-based [30]. However, establishing an accurate baseline is challenging, often resulting in high false positive and false negative rates [56].

2.3 Hybrid IDS

Hybrid detection approaches combine both signature-based and anomaly-based techniques. Signature-based detection ensures the system accurately detects known threats and in combination with anomaly-based detection can discover unknown attacks [1]. Hybrid detection also mitigates the high false positive and false negative rates associated with anomaly detection [30].

While hybrid detection enhances traditional IDSs by combining the strengths of both approaches, it still faces challenges, such as computational overhead [31]. There is a need for systems which can learn from their environment without computational complexity or excessive human input [31]. Lightweight AI-driven IDSs can address these issues faced by traditional methods to secure resource-constrained networks.

3 AI-BASED INTRUSION DETECTION SYSTEMS

Machine learning, a subfield of Artificial Intelligence (AI), develops algorithms that identify patterns in datasets to enable predictive analysis. This capability is particularly valuable in IDS applications [49]. There are three main types of machine learning methods: supervised, unsupervised and semi-supervised. Deep learning is a specialised subset of machine learning that can exhibit characteristics of supervised, unsupervised, or semi-supervised learning, depending on its implementation [13]. However, it does not belong exclusively to one category.

3.1 Supervised learning

Supervised learning models use labelled datasets to learn mappings between inputs and outputs, allowing for accurate classification and prediction [55]. Supervised learning primarily consists of classification models, which are widely used in IDS, and regression models.

Classification models allocate data to specific categories based on its features. These models learn from labelled inputs and classify new data based on learned patterns. A well-trained classification model can effectively detect attacks by distinguishing between normal and abnormal network traffic patterns [64]. Decision tree, k-nearest neighbour (KNN), neural network, support vector machine (SVM) and random forest are commonly used classification algorithms [64].

While regression models predict continuous outcomes, they are less commonly used in IDSs than classification models, which focus on detecting attack patterns. These models analyse patterns and relationships in existing datasets to make predictions on unseen data. Linear regression, logistic regression, decision tree, random forest, and support vector machine are commonly used regression algorithms [64]. Logistic regression models are not limited to the context of intrusion detection and can be used in spam detection [43].

Support Vector Machine (SVM)
SVM is widely used in IDSs due to its strong classification performance. Lightweight adaptations, such as linear SVM with optimised feature selection, have been explored for resource-limited networks [61]. SVM identifies an optimal hyperplane that best separates data points into distinct categories. SVM used for feature selection enhances classification accuracy in network attack detection [62].

Azimjonov et al. [6] proposed a lightweight IDS for IoT networks using a fine-tuned linear SVM (LSVM) model and four feature selection methods, reducing computation overhead. The study demonstrates that selecting relevant feature subsets significantly improved IDS performance. The models trained on the optimised feature subsets outperformed their full-feature counterparts, highlighting the impact of feature selection on IDS accuracy and efficiency [6]. The findings suggest that feature selection plays a critical role in balancing IDS accuracy and computational cost, making LSVMs a viable choice for resource-constrained networks [6]. Jan et al. [24] designed an optimised SVM-based IDS for detecting DDoS attacks in IoT environments which demonstrated its suitability for real-time detection with minimal resource usage. The system focused on two key factors: the packet arrival rate as a classification feature and a lightweight SVM-based classifier. The experimental results demonstrated that the IDS effectively detected intrusions [24]. A comparative analysis further validated SVM's advantages over other machine learning-based classifiers, including K-nearest neighbour and decision trees [24]. The study emphasised SVM's suitability for resource-constrained IoT networks, where lightweight and accurate intrusion detection is critical [24].

K-Nearest Neighbour (KNN)
KNN is a simple yet effective model that classifies data points based on the majority class of their nearest neighbors and is used for both classification and regression tasks [61]. Pan et al. [41] proposed a lightweight intrusion detection model for Wireless Sensor Networks (WSNs) that integrates KNN with the Sine Cosine Algorithm to enhance security, energy efficiency, and real-time performance. Due to their low computational demands, KNN and the Sine Cosine Algorithm (SCA) offer an efficient and lightweight solution for intrusion detection in constrained environments such as IoT and

WSNs [41]. Similarly, Meng et al. [33], introduced a lightweight false alarm filter based on the KNN classifier to enhance the reliability of intrusion detection. The model used a rating mechanism to classify incoming alarms into labelled clusters. The approach effectively reduced unwanted alerts while maintaining affordable CPU usage, demonstrating KNN's practicality for real-time intrusion detection with minimal computational overhead [33].

Random forest algorithm
Random forest is a robust classification model for intrusion detection, though its computational intensity makes it less ideal for lightweight IDSs without optimizations such as feature reduction or depth constraints [61]. Htun et al. [21] shows that random forest performs well in intrusion detection, however lengthy computation time makes it undesirable for real-time detection. Samunnisa et al. [54] proposed an IDS capable of detecting both known and zero-day attacks by combining KNN clustering with a random forest classifier. The hybrid approach demonstrated strong performance, achieving an accuracy of 98.27 percent [54]. The integration of random forest with KNN clustering helped enhance the model's ability to detect novel threats while maintaining high classification accuracy [54]. While random forest has shown strong intrusion detection performance, research on lightweight adaptations remains limited, highlighting a gap in optimising the model for resource-constrained environments.

One of the key advantages of supervised learning is its ability to leverage past experiences to refine outputs [64]. Supervised learning can be optimised using previous results to enhance performance [64]. This makes it useful to solve many computational problems. However, the model requires high-quality input data during training to produce sufficient results. The training time of the model is computationally expensive and it can be challenging to classify big data [64]. Addressing these computational challenges is crucial when developing lightweight AI-based IDSs, ensuring effective threat detection without overwhelming system resources.

3.2 Unsupervised learning

Unsupervised learning models analyse unlabelled data autonomously, detecting hidden patterns without human intervention, making them suitable for novel attack detection in IDSs [55]. These models iteratively refine their understanding, extracting meaningful structures from raw data, which aids in adaptive intrusion detection. The algorithm usually groups different data into categories based on similarities or differences. Two key problem types within unsupervised learning, particularly relevant to this review, are clustering and dimensionality reduction.

Clustering organises unlabelled data based on shared features, grouping similar instances for easier intrusion analysis [19]. There are subcategories within clustering algorithms: exclusive, overlapping, hierarchical and probabilistic. Exclusive clustering assigns each data point to only one cluster. Overlapping clustering allows data points to belong to multiple clusters, capturing complex relationships in intrusion patterns. Hierarchical clustering merges clusters iteratively until a single cluster remains. Probabilistic clustering sorts data points based on the probability of belonging to a given cluster. In IDSs, clustering aids in intrusion signature reduction, high-quality signature generation, and efficient grouping of similar attacks [25]. K-means, Fuzzy c-means, and Gaussian Mixture are some of the common clustering algorithms used [64].

Dimensionality reduction enhances IDS efficiency by reducing feature complexity, lowering computational costs while preserving key intrusion patterns. As datasets grow larger, it has become common practice to apply dimensionality reduction before running a machine learning algorithm [64]. One method of dimensionality reduction is feature elimination, which removes redundant features to streamline prediction models. Another is feature extraction. This method works by creating the same number of features which already exist in the dataset. These new, independent features are a mixture of old features, ordered by importance. The least important new features are then removed. This can be useful in lowering computational costs and latency [27]. Ismail et al. [23] presents an IoT IDS that utilises a Recursive Feature Elimination wrapper with a Decision Tree classifier to remove redundant features which optimises detection speed and resource efficiency. Experimental results demonstrated high accuracy, precision, and recall, highlighting the potential of hybrid IDS models that integrate feature elimination for reduced computational complexity [23].

Unsupervised learning is valuable for IDSs as it detects novel attacks without relying on labelled training data, making it adaptable to emerging threats [19]. Bhattacharjee et al. [10] compared the fuzzy c-means and k-means clustering algorithms for IDS applications. The study found that fuzzy clustering provides a more practical and efficient solution for intrusion detection, as it allows data points to belong to multiple clusters, capturing the uncertainty in attack classification. Horng et al. [20] proposed a hierarchical clustering-based IDS that reduced dataset size while preserving structure, significantly decreasing SVM training time and enhancing real-time detection. Evaluations on the KDD Cup99 dataset demonstrated an accuracy of 95.72 percent with a low false alarm rate, proving the effectiveness of clustering in preprocessing large-scale intrusion data [20]. Alazzam et al. [2] proposed a lightweight IDS utilising k-means clustering and one-class SVM reduced dataset size. The model maintained a high accuracy of 99.3 percent and demonstrated an efficient intrusion detection approach. Ravale et al. [48] proposed a hybrid IDS combining k-means clustering and the RBF kernel of SVM, improving detection accuracy while minimising false alarms and computational load. The hybrid model indicates the potential of clustering methods in the development of lightweight solutions.

The absence of human intervention in unsupervised learning broadens its applicability, making it ideal for large-scale IDS deployments [61]. It identifies differences in large datasets faster than manual analysis [64]. This means that unsupervised learning is beneficial in detecting unknown anomalies within large network traffic if used to improve IDSs. Unsupervised learning has proven to detect network attacks with high accuracy rates making it an appealing approach for IDSs [61]. However, the high amount of data required to train

the model means that a lot of time and computational resources are used [64]. Unsupervised learning, while adaptable, carries a higher risk of inaccuracies compared to supervised methods, necessitating careful validation [64]. Due to potential inaccuracies, manual verification is often required post-training, making unsupervised IDS models more resource-intensive despite their adaptability.

3.3 Semi-supervised learning

Semi-supervised learning combines supervised and unsupervised learning where only part of the dataset is labelled [30]. Semi-supervised learning leverages techniques from both supervised and unsupervised methods, improving algorithm accuracy [64]. It is more time-efficient as it does not necessitate a fully labelled dataset. However, it is still constrained by the disadvantages that unsupervised and supervised learning is subject to [64]. Bahrololum et al. [7] proposed an IDS based on hybrid supervised and unsupervised neural networks (NNs). Unsupervised NNs were used for training normal packets and supervised NNs were used for clustering and classification of network attacks. The model achieved a true positive rate of 73.67 percent and a false positive rate of 26.53 percent.

Semi-supervised learning is particularly beneficial for IDSs as it leverages small amounts of labelled data which can improve performance, in comparison to unsupervised learning, with less time and costs required [25]. This can be a useful approach in the implementation of a lightweight AI-IDS for securing resource-constrained networks.

3.4 Deep learning

Deep learning (DL) models are widely used in predictive modeling due to their capacity to automatically learn intricate patterns from vast datasets, which outperform traditional statistical methods [13]. DL relies on artificial neural networks with multiple layers, which are designed to learn from data and make decisions in a manner inspired by the human brain [63]. Among the deep learning techniques applied in IDSs, convolutional neural networks (CNNs), recurrent neural networks (RNNs), autoencoders (AEs), and deep belief networks (DBNs) will be discussed. These algorithms enhance IDSs by detecting intricate patterns and anomalies in network traffic that conventional rule-based methods often fail to detect [38].

Convolutional neural networks (CNNs)
CNNs are a specialised type of deep learning algorithm predominantly used for object recognition and image classification. CNNs consist of multiple layers, each responsible for learning and extracting relevant features to enhance data classification [61]. Due to their ability to analyse high-dimensional data and recognise complex patterns, CNNs are particularly well-suited for network traffic analysis [49]. CNN algorithms can effectively distinguish between normal and abnormal data, addressing a key limitation faced by traditional anomaly-based IDS [12]. While CNNs efficiently process large IoT datasets, their high computational cost can strain available resources [49]. To address this challenge, hybrid CNN architectures, where CNN layers are combined with other models, have been proposed to reduce computational load while maintaining optimal performance [28].

Xiao et al. [67] developed a CNN-based IDS that optimised feature selection post-training, improving detection speed and accuracy. Chowdhury et al. [12] proposed a few-shot deep learning strategy, where a model trained on a small number of labelled examples combined CNNs with SVM and 1-NN classifiers. This approach achieved high accuracy rates on both the KDD99 and NSL-KDD datasets. Du et al. [14] proposed a CNN-LSTM hybrid model to improve IoT security by leveraging CNN's feature selection capabilities and LSTM's ability to capture temporal dependencies. The hybrid model demonstrated strong classification performance on the KDDCup99 dataset [14]. Similarly, Garg et al. [17] introduced an IoT-focused IDS using a combination of CNN and gray wolf optimization (GWO), which minimised feature dimensions while achieving improved detection performance on the DARPA98, KDD99, and synthetic datasets. Pham et al. [42] designed a lightweight CNN-based IDS to transform raw network traffic into image representations. The IDS utilised packet length and arrival time as key features [42]. This approach enabled the CNN model to classify malicious traffic effectively, achieving 95 percent accuracy on the CSE-CICIDS2018 dataset while requiring minimal training data [42]. Given their increasing adoption in AI-IDS development, CNNs especially in hybrid models present a promising lightweight solution for effective intrusion detection [70].

Recurrent neural networks (RNNs)
While CNNs excel in feature extraction, another class of deep learning models, RNNs, are particularly suited for processing sequential data, making them valuable for IDSs [61]. RNNs process data sequentially, leveraging past inputs to improve real-time predictions, making them particularly useful for analysing time-series network traffic. Subsets of RNNs, such as Long Short-Term Memory (LSTM) networks and Bi-Directional LSTM, are effective in IDSs due to the need to model sequential network traffic flows [49]. LSTM networks mitigate the vanishing gradient problem in traditional RNNs, allowing them to retain long-term dependencies crucial for IDSs [19]. This is particularly beneficial for long sequences of data required to detect time-based attack patterns [51]. Bi-LSTM models process data in forward and backward directions which enhance the detection of complex attack patterns like multi-stage attacks and have proven successful in environments requiring real-time intrusion detection [50].

LSTMs have been widely adopted due to their ability to retain long-term dependencies, enhancing intrusion detection accuracy [61]. Ibrahim et al. [22] proposed an LSTM-RNN model that, when compared to KNN, CNN, and decision tree models, demonstrated improved accuracy in attack detection, required no human intervention, and maintained an efficient run-time. That being said, the computational complexity of RNNs makes them less practical for intrusion detection in resource-constrained environments. To address this, several studies have explored lightweight RNN-based models. A hybrid CNN-RNN approach leveraged CNNs for feature extraction while using RNNs for sequence modeling, achieving high detection accuracy with reduced computational overhead [46]. Altunay et al. [4] proposed a hybrid CNN and LSTM-based IDS for

IoT networks. The model was compared with studies that followed a binary and multi-class classification and attained the highest accuracy for intrusion detection. A Bi-LSTM model applied to IoT networks improved IDS accuracy and response time by capturing bidirectional dependencies in network traffic [51]. These studies show there is potential for RNN-based IDS solutions which use feature reduction techniques to improve real-time performance. However, there is still limited research in optimising RNNs for resource-constrained networks.

Autoencoders (AEs)
In addition to the success of CNNs and RNNs, unsupervised deep learning architectures such as autoencoders and deep belief networks introduce additional advantages, particularly in anomaly detection. AEs are a form of artificial neural networks that learn to reconstruct its input data by efficiently compressing and then decompressing it. AEs consist of an encoder, used to compress input data, and a decoder, used to reconstruct the original input from the compressed representation. Autoencoders have proven effective for IDSs, particularly in detecting previously unknown attacks [60]. By training on normal traffic data, AEs can classify inputs as attacks when reconstruction errors exceed a predefined threshold [60]. However, AE-based IDS models may fail to detect attacks that closely resemble normal network traffic, leading to false negatives. AEs can allow for the development of an efficient IDS with an inexpensive computational cost [58]. Although there exists lightweight implementations of AE models, the models require optimisation to be used for resource-constrained networks [57].

Meidan et al. [32] proposed a deep autoencoder-based technique for detecting IoT botnet attacks, where each IoT device was independently trained on normal traffic to recognise anomalies based on reconstruction errors [32]. The approach effectively minimised false positive rates, achieving an FPR of zero [32]. However, training separate models for each IoT device posed scalability challenges [32]. Farahnakian et al. [15] proposed a denoising autoencoder (DAE) model for intrusion detection in IoT systems, leveraging unsupervised learning for feature extraction and anomaly detection [15]. Using the NSL-KDD and CICIDS 2017 datasets, the DAE model achieved 99.991 percent accuracy with CICIDS 2017 and 99.4 percent accuracy with NSL-KDD, highlighting its effectiveness in real-time intrusion prevention. Another study introduced an optimised common feature selection and deep autoencoder model for lightweight intrusion detection in IoT environments, utilising feature selection, data compression, and pruning techniques [40]. The model, deployed on a Raspberry Pi4 with the TFLite interpreter, demonstrated an overall accuracy of 99 percent and 97 percent across two datasets while significantly reducing computational cost and memory usage [40]. These findings emphasise the potential of autoencoder-based models in developing efficient and scalable IDS solutions tailored for both traditional network security and IoT environments.

Deep belief networks (DBNs)
Similar to autoencoders, DBNs provide a robust approach to intrusion detection. These generative deep neural networks consist of multiple layers of Restricted Boltzmann Machines (RBMs) stacked together. RBMs are a generative stochastic neural network which learns a probability distribution over its input using a two-layered architecture, a hidden and a visible layer. By stacking RBMs, DBNs enable the hidden layer of one RBM to serve as the visible layer of the next, allowing them to generate new data samples resembling the training set. This layered structure makes DBNs particularly useful for feature extraction and classification in IDSs [69]. Key features of DBNs, which make it effective for IDSs, are its ability to preprocess data to filter out noise, perform probabilistic reconstruction of inputs and represent feature detectors with its layered structure [9]. This enhances DBNs classification capabilities in detecting anomalies and threats [9].

Marir et al. [29] proposed a large-scale IDS model using a combination of DBN and SVM methodology. DBNs were used for feature reduction, with the extracted features processed by an ensemble SVM, achieving superior detection rates for abnormal behavior compared to other classifiers [29]. The model also outperformed other classifiers for both normal and attack classes [29]. Alom et al. [3] proposed a DBN IDS model that identified any unknown attack in the NSL-KDD dataset which achieved a 97.5 percent detection accuracy. Wei et al. [66] proposed a DBN-based IDS using a joint optimised method and was found to have a high detection speed and accuracy. However due to the model's complexity, the training duration was long time [66]. Nivaashini et al. [39] designed a federated DBN IDS model which integrated RNNs, CNNs, AEs, and deep Boltzmann machines. The IDS model was found to enhance detection accuracy with 99 percent outperformed conventional DBNs classification performance by 3 percent [39]. These findings underscore the potential of hybrid deep learning models in developing efficient and lightweight IDS solutions, balancing accuracy with computational feasibility.

4 BENCHMARKING INTRUSION DETECTION SYSTEMS

4.1 Standard Datasets

An important factor in building an IDS is the selection of the dataset [59]. The selected dataset plays a crucial role in both training and evaluating IDS models. Due to privacy concerns, datasets containing real network packet captures are rarely available for IDS development [25]. Thus, several publicly labelled network traffic datasets are available for evaluating IDSs such as KDD Cup99, NSL-KDD, UNSW-NB15, Kyoto 2006, CICIDS2017, and CSE-CICIDS-2018.

4.1.1 KDDcup99. KDD Cup99 was one of the first large-scale publicly available IDS datasets and became a standard benchmark for evaluating intrusion detection models [65]. One of KDD Cup99's limitations is its outdated records, which do not include modern malware threats [25]. The dataset consists of 4 900 000 attack records from processed tcpdump data from the 1998 DARPA intrusion detection challenge dataset. Each sample has 41 features and is labelled as Normal or Attack. The attack samples are separated into four categories: Denial of Service (DoS), User to Root (U2R), Remote to Local (R2L), and Probe. It consists of three subsets: (1) the full

dataset, (2) a 10 percent subset, and (3) a test set with 311,029 samples. Due to its limitations, researchers have increasingly shifted to newer datasets to avoid biases in IDS training [59].

4.1.2 NSL-KDD. NSL-KDD was developed to address the shortcomings of KDD Cup99. The dataset removed redundancies and transformed the structure of the KDD Cup 99 dataset [25]. It retains KDD Cup99's four attack categories. The dataset reflects the characteristics of real-time network traffic, which is useful for the training of IDSs [65]. The NSL-KDD contains two files, a training set, and a testing set. The training set contains 21 different attacks and has 126 620 instances. The testing set contains 37 different attacks and has 22 850 instances. The size of the dataset means it is practical to use the whole dataset without the need for random sampling [25]. NSL-KDD reduces the risk of model bias and is effective for misuse detection [65]. NSL-KDD has been widely used in IDS research, producing consistent and comparable results across studies [25].

4.1.3 UNSW-NB15. UNSW-NB15 is a more recent dataset than NSL-KDD, so it is more representative of real network traffic [41]. The UNSW-NB15 dataset was developed by the Australian Centre for Cyber Security to provide a modern benchmark for IDS research, containing labelled network traffic derived from real-world PCAP files [19]. It aims to generate network traffic composed of normal activities and attack behaviours. This dataset contains over 2 million records, each with 49 features, and is labelled with attack categories and binary classification (attack or normal) [63]. The dataset has nine types of attacks: DoS, Reconnaissance, Worms, Fuzzers, Backdoors, Exploits, Analysis, Generic, and Shellcode [64]. It is designed to evaluate anomaly detection systems [19]. A major limitation of UNSW-NB15 is that normal traffic makes up approximately 87 percent of the dataset, which may result in biased IDS models that struggle with minority-class attack detection [63].

4.1.4 CICIDS2017. The CICIDS2017 dataset was created by the Canadian Institute for Cybersecurity in 2017. A key strength of CICIDS2017 is its realistic simulation of modern cyber threats, as it was generated from real-time network traffic collected over several days [18]. The dataset includes detailed metadata, such as timestamps, source and destination ports, IP addresses, and attack types, making it valuable for deep packet inspection and behavioral analysis [64]. The dataset includes 86 features. Common attacks were classified as DoS attacks, brute force SSH attack, brute force FTP attack, web attack, botnet attack, DDoS attack, heartbleed attack and infiltration attack [34].

4.1.5 Kyoto 2006. The Kyoto 2006 dataset was developed by Kyoto University to improve intrusion detection by collecting diverse network traffic data from various security monitoring sources. It was created by deploying honeypots, darknet sensors, email servers, web crawlers and other network security measures outside and inside the university to collect various types of traffic [64]. Kyoto 2006 builds on KDD Cup99 by incorporating 14 existing statistical features and introducing 10 additional features focused on attack behavior and traffic characteristics [19]. Kyoto 2006 categorises traffic into three labels: known, normal and unknown attack. The most recent dataset includes network traffic from 2006 to 2015 [64].

4.2 Evaluation Metrics

Evaluation metrics describe the performance of classification models and play a crucial role in determining the effectiveness of an IDS [53]. The selection of an appropriate evaluation metric is important when comparing different classification algorithms determining top performing models [19].

Confusion Matrix
A confusion matrix is a fundamental tool used to evaluate the performance of IDS classification models [47]. The matrix compares the real target values with those predicted by the AI model. It consists of four components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). A TP is an attack correctly identified as an attack and a TN describes normal network traffic correctly classified as normal. A FP is normal traffic which has been incorrectly classified as an attack and a FN is an attack incorrectly classified as normal. A well-performing IDS should have minimal false positives to reduce false alerts in the system which could result in network disturbances [64]. It is also important to have minimal false negatives to avoid undetected attacks on the network [68].

Accuracy
Accuracy is a widely used metric that measures the overall correctness of an IDS by evaluating the proportion of correctly classified instances over the total number of cases [47].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} [68] \tag{1}$$

Precision
Precision measures the proportion of true positive classifications relative to all instances classified as positive. It helps assess the IDS's ability to correctly identify actual threats [47].

$$\text{Precision} = \frac{TP}{TP + FP} [47] \tag{2}$$

Detection Rate
Also referred to as the Detection Rate, recall measures how well an IDS captures actual attacks by computing the proportion of detected intrusions out of all real intrusions in the dataset [47].

$$\text{Detection Rate} = \frac{TP}{TP + FN} [68] \tag{3}$$

False Alarm Rate
The False Positive Rate evaluates the proportion of incorrectly classified normal traffic, which impacts the reliability of an IDS [47].

$$\text{False Positive Rate/False Alarm rate} = \frac{FP}{FP + TN} [68] \tag{4}$$

By utilising these evaluation metrics, different IDS models can be effectively compared and the most suitable classification algorithms can be identified.

5 DISCUSSION

The field of IDSs has evolved significantly with the integration of AI [61]. While traditional IDS methods remain effective to some extent, they struggle to adapt to increasingly sophisticated cyber threats. The literature indicates that AI-driven IDSs have emerged as powerful alternatives, offering improved detection accuracy and adaptability, with the potential to optimise computational efficiency. The discussion explores the superiority of AI-based IDSs over traditional methods, evaluates different AI models, examines emerging trends such as hybrid models and feature extraction, and highlights the need for Lightweight AI-based IDS (LAIDS) in addressing cybersecurity concerns in resource-constrained environments.

5.1 Observations

The literature suggests a unanimous opinion that AI-based IDSs outperform traditional IDS methods in detecting and mitigating cyber threats [52]. AI-IDSs formalise unknown patterns, help write or improve signature-based rules for new vulnerabilities and attack patterns [26]. In comparison, traditional IDS techniques struggle with evolving attack patterns and accurate detections [30, 56]. AI-powered IDSs, particularly those using machine learning and deep learning, address the limitations of traditional techniques by identifying patterns in large datasets, allowing them to adapt to new threats with minimal manual intervention. Literature suggests that AI-driven IDSs achieve higher accuracy compared to traditional approaches [40, 54].

Unlike traditional versus AI-based IDSs, Deep Learning (DL) and Machine Learning (ML) may not have as contrasting differences. While machine learning models such as SVM, KNN, and Random Forest have been optimised for lightweight applications, they rely heavily on feature engineering and pre-processing [6, 36, 61]. In contrast, DL-based approaches such as CNNs and AEs have shown impressive performance and accuracy by leveraging automated feature extraction and hierarchical learning, reducing the need for extensive pre-processing [14, 35, 61, 67]. The CNN-LSTM hybrid model, for example, effectively captures both spatial and temporal features, enhancing intrusion detection in dynamic IoT environments [14]. Additionally, AEs, particularly denoising and optimised deep AEs, have demonstrated strong anomaly detection capabilities with high accuracy and reduced computational costs [15, 40]. However, DL tends to require more computational resources and training time compared to ML approaches, making their deployment in resource-constrained environments more challenging [35]. The literature suggests a growing trend toward lightweight adaptations of deep learning models, such as optimised feature selection techniques [14, 15, 40]. These lightweight DL adaptations are aimed to be more viable for real-time intrusion detection in IoT and WSNs.

Hybrid models, which combine multiple architectures to enhance detection capabilities, represent an emerging trend in IDS research. These approaches leverage the strengths of different models to mitigate the weaknesses of individual techniques. CNNs are frequently employed in lightweight hybrid IDS implementations, highlighting their potential for resource-constrained networks [12, 14, 17]. SVM, a widely used ML approach, has been paired with clustering algorithms in several studies focused on developing lightweight IDS models [2, 20, 48]. Another common strategy in the literature is feature extraction, which enhances IDS performance by reducing data dimensionality and computational overhead [15, 27]. Effective feature extraction techniques, such as AEs, optimise data preprocessing, making AI-IDS more efficient [40]. The combination of feature extraction with a hybrid model may present a way to further optimise IDSs for resource-constrained networks.

Different IDS models involve tradeoffs between performance, accuracy, and resource consumption. As mentioned, DL models, although highly accurate, require significant processing power and memory which may make them less suitable for IoT and resource-constrained networks. In contrast, ML models balance efficiency and accuracy more evenly, however require extensive feature engineering. Advancements in specialised hardware, such as neuromorphic computing chips and Field-Programmable Gate Array-based accelerators, could reduce the energy and computational costs associated with deep learning, making them more viable for edge deployment in the future [44]. As these hardware solutions become more accessible, they may enable deep learning-based IDS to operate efficiently even in resource-constrained environments, shifting the tradeoff landscape [44]. At present, hybrid approaches offer a means to mitigate tradeoffs, particularly in resource-constrained network environments.

5.2 The LAIDS Project

As technology evolves, so do cyber-attacks and adversarial techniques, necessitating continuous improvements in security measures [26]. The increased use of IoT devices and WSNs has heightened the need for Lightweight AI-IDS (LAIDS). Resource-constrained devices like these lack the capabilities required to run resource-hungry IDS solutions, and thus creates a challenge of securing such networks from advanced attacks. In a South African context, there is widespread mobile phone ownership, however, the digital divide means that many people access resource-constrained networks [11]. This suggests that traditional high-resource IDSs are impractical due to both resource constraints and limited accessibility. The LAIDS project could be a viable alternative which addresses these issues by integrating AI-based security mechanisms with a focus on minimising computational demands.

Based on the models presented in the review, a promising approach for building LAIDS involves leveraging hybrid models that integrate ML and DL techniques. A viable lightweight IDS model could involve a two-stage hybrid approach: an initial filtering stage using a lightweight machine learning classifier, such as SVM, to filter out normal traffic, followed by a deep learning-based anomaly detection system, such as CNNs or LSTMs, for threat analysis. SVMs were recognised for their efficient application in the literature [2, 6, 24, 48]. Both CNNs and LSTMs were used in lightweight

hybrid models and thus, could be an effective approach [4, 14]. Feature extraction techniques such as AEs can be used to reduce computational overhead by reducing redundant data. This allows DL implementations to be more feasible in resource-constrained environments. Similar to the work done by Otokwala et al. [40], the LAIDS project can potentially be deployed on a Raspberry Pi device to determine the model's performance and computational overheads. Optimised architectures such as TinyML may further improve efficiency by reducing the memory and processing demands of DL models [16]. The proposed hybrid models aim to balance performance, accuracy, and resource consumption, making them well-suited for resource-constrained networks. Although the proposed models strive to balance various factors, they will still face resource-performance tradeoffs.

Minimising energy consumption and processing overhead will be a key challenge for the LAIDS project. Beyond algorithmic efficiency, real-world deployment of LAIDS must address security risks such as adversarial attacks, the feasibility of deployment on low-power devices, and real-time responsiveness [45]. Ensuring the security, scalability, and privacy of IDS solutions will be critical considerations, as these factors directly impact the effectiveness of AI-driven IDSs [45].

6 CONCLUSIONS

The evolution of Intrusion Detection Systems (IDSs) has demonstrated the growing importance of AI-driven approaches in cybersecurity. While traditional IDS methods remain relevant, their limitations in adapting to evolving cyber threats have accelerated the shift toward machine learning and deep learning solutions. The tradeoff between computational efficiency and detection accuracy highlights the potential of hybrid models, which combine the strengths of both techniques.

As IoT and wireless sensor networks expand, along with existing resource-constrained networks, the demand for Lightweight AI-IDS (LAIDS) has become increasingly critical. The literature suggests that hybrid models, particularly those incorporating feature extraction techniques such as autoencoders, can enhance performance while minimising computational overhead. The LAIDS project can focus on refining these hybrid models, leveraging techniques such as SVMs, CNNs, and LSTMs, and exploring lightweight architectures like TinyML to further improve efficiency. Additionally, practical implementations on low-power devices, such as Raspberry Pi, can provide valuable insights into real-world feasibility. As cyber threats continue to evolve, developing adaptive, lightweight, and high-performance IDS solutions will remain a crucial area of cybersecurity research.

Model	Overall Accuracy	Weighted Precision	Weighted Recall	Weighted F1	False Alarm Rate	False Negative Rate
Baseline CNN	94.47%	97.00%	94.00%	96.00%	6.34%	1.53%
CNN-LSTM	99.58%	≈100%	≈100%	≈100%	0.27%	1.1%
DBN	91.07%	96.00%	91.00%	93.00%	10.40%	0.85%

Table 1: Key Metrics across Models.

REFERENCES

[1] Moorthy Agoramoorthy, Ahamed Ali, D Sujatha, Michael Raj TF, and G Ramesh. 2023. An Analysis of Signature-Based Components in Hybrid Intrusion Detection Systems. In *2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS)*. IEEE, 1–5.

[2] Hadeel Alazzam, Ahmad Sharieh, and Khair Eddin Sabri. 2022. A lightweight intelligent network intrusion detection system using OCSVM and Pigeon inspired optimizer. *Applied Intelligence* 52, 4 (2022), 3527–3544.

[3] Md Zahangir Alom, VenkataRamesh Bontupalli, and Tarek M Taha. 2015. Intrusion detection using deep belief networks. In *2015 National Aerospace and Electronics Conference (NAECON)*. IEEE, 339–344.

[4] Hakan Can Altunay and Zafer Albayrak. 2023. A hybrid CNN+ LSTM-based intrusion detection system for industrial IoT networks. *Engineering Science and Technology, an International Journal* 38 (2023), 101322.

[5] Zainab Alwaisi, Tanesh Kumar, Erkki Harjula, and Simone Soderi. 2024. Securing constrained IoT systems: A lightweight machine learning approach for anomaly detection and prevention. *Internet of Things* 28 (2024), 101398.

[6] Jahongir Azimjonov and Taehong Kim. 2024. Designing accurate lightweight intrusion detection systems for IoT networks using fine-tuned linear SVM and feature selectors. *Computers & Security* 137 (2024), 103598.

[7] Marjan Bahrololum, Elham Salahi, and Mahmoud Khaleghi. 2009. Anomaly intrusion detection design using hybrid of unsupervised and supervised neural network. *International Journal of Computer Networks & Communications (IJCNC)* 1, 2 (2009), 26–33.

[8] Shahid Allah Bakhsh, Muhammad Almas Khan, Fawad Ahmed, Mohammed S Alshehri, Hisham Ali, and Jawad Ahmad. 2023. Enhancing IoT Network security through deep learning-powered Intrusion Detection System. *Internet of Things* 24 (2023), 100936.

[9] Nagaraj Balakrishnan, Arunkumar Rajendran, Danilo Pelusi, and Vijayakumar Ponnusamy. 2021. Deep Belief Network enhanced intrusion detection system to prevent security breach in the Internet of Things. *Internet of things* 14 (2021), 100112.

[10] Partha Savathi Bhattacharjee, Abul Kashim Md Fujail, and Shahin Ara Begum. 2017. A comparison of intrusion detection by K-means and fuzzy C-means clustering algorithm over the NSL-KDD dataset. In *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 1–6.

[11] Elirea Bornman. 2016. Information society and digital divide in South Africa: results of longitudinal surveys. *Information, Communication & Society* 19, 2 (2016), 264–278.

[12] Md Moin Uddin Chowdhury, Frederick Hammond, Glenn Konowicz, Chunsheng Xin, Hongyi Wu, and Jiang Li. 2017. A few-shot deep learning approach for improved intrusion detection. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*. IEEE, 456–462.

[13] Bo Dong and Xue Wang. 2016. Comparison deep learning method to traditional methods using for network intrusion detection. In *2016 8th IEEE international conference on communication software and networks (ICCSN)*. IEEE, 581–585.

[14] Jiawei Du, Kai Yang, Yanjing Hu, and Lingjie Jiang. 2023. NIDS-CNNLSTM: Network intrusion detection classification model based on deep learning. *IEEE Access* 11 (2023), 24808–24821.

[15] Fahimeh Farahnakian and Jukka Heikkonen. 2018. A deep auto-encoder based approach for intrusion detection system. In *2018 20th international conference on Advanced communication technology (ICACT)*. IEEE, 178–183.

[16] Pietro Fusco, Gennaro Pio Rimoli, and Massimo Ficco. 2024. Tinyids-an iot intrusion detection system by tiny machine learning. In *International Conference on Computational Science and Its Applications*. Springer, 71–82.

[17] Sahil Garg, Kuljeet Kaur, Neeraj Kumar, Georges Kaddoum, Albert Y Zomaya, and Rajiv Ranjan. 2019. A hybrid deep learning-based model for anomaly detection in cloud datacenter networks. *IEEE Transactions on Network and Service Management* 16, 3 (2019), 924–935.

[18] Amirhossein Gharib, Iman Sharafaldin, Arash Habibi Lashkari, and Ali A Ghorbani. 2016. An evaluation framework for intrusion detection dataset. In *2016 International conference on information science and security (ICISS)*. IEEE, 1–6.

[19] Mohammed Sayeeduddin Habeeb and T Ranga Babu. 2022. Network intrusion detection system: a survey on artificial intelligence-based techniques. *Expert Systems* 39, 9 (2022), e13066.

[20] Shi-Jinn Horng, Ming-Yang Su, Yuan-Hsin Chen, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai, and Citra Dwi Perkasa. 2011. A novel intrusion detection system based on hierarchical clustering and support vector machines. *Expert systems with Applications* 38, 1 (2011), 306–313.

[21] Phyu Thi Htun and Kyaw Thet Khaing. 2012. Anomaly intrusion detection system using random forests and k-nearest neighbor. *Probe* 41102, 4107 (2012), 2377.

[22] Mariam Ibrahim and Ruba Elhafiz. 2023. Modeling an intrusion detection using recurrent neural networks. *Journal of Engineering Research* 11, 1 (2023), 100013.

[23] Mahmoud G Ismail, Mohamed Abd El Ghany, and Mohammed A-M Salem. 2022. Enhanced Recursive Feature Elimination for IoT Intrusion Detection Systems. In *2022 International Conference on Microelectronics (ICM)*. IEEE, 193–196.

[24] Sana Ullah Jan, Saeed Ahmed, Vladimir Shakhov, and Insoo Koo. 2019. Toward a lightweight intrusion detection system for the internet of things. *IEEE access* 7 (2019), 42450–42471.

[25] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. 2019. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity* 2, 1 (2019), 1–22.

[26] Aechan Kim, Mohyun Park, and Dong Hoon Lee. 2020. AI-IDS: Application of deep learning to real-time Web intrusion detection. *Ieee Access* 8 (2020), 70245–70261.

[27] Jing Li, Mohd Shahizan Othman, Hewan Chen, and Lizawati Mi Yusuf. 2024. Optimizing IoT intrusion detection system: feature selection versus feature extraction in machine learning. *Journal of Big Data* 11, 1 (2024), 36.

[28] Mohammed A Mahdi. 2024. Secure and Efficient IoT Networks: An AI and ML-based Intrusion Detection System. In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)*. IEEE, 1–6.

[29] Naila Marir, Huiqiang Wang, Guangsheng Feng, Bingyang Li, and Meijuan Jia. 2018. Distributed abnormal behavior detection approach based on deep belief network and ensemble SVM using spark. *IEEE Access* 6 (2018), 59657–59671.

[30] Michal Markevych and Maurice Dawson. 2023. A review of enhancing intrusion detection systems for cybersecurity using artificial intelligence (ai). In *International conference knowledge-based organization*, Vol. 29. 30–37.

[31] Elijah M Maseno, Zenghui Wang, and Hongyan Xing. 2022. A systematic review on hybrid intrusion detection system. *Security and Communication Networks* 2022, 1 (2022), 9663052.

[32] Yair Meidan, Michael Bohadana, Yael Mathov, Yisroel Mirsky, Asaf Shabtai, Dominik Breitenbacher, and Yuval Elovici. 2018. N-baiot—network-based detection of iot botnet attacks using deep autoencoders. *IEEE Pervasive Computing* 17, 3 (2018), 12–22.

[33] Weizhi Meng, Wenjuan Li, and Lam-For Kwok. 2015. Design of intelligent KNN-based alarm filter using knowledge-based alert verification in intrusion detection. *Security and Communication Networks* 8, 18 (2015), 3883–3895.

[34] Nour Moustafa and Jill Slay. 2016. The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Information Security Journal: A Global Perspective* 25, 1-3 (2016), 18–31.

[35] Salman Muneer, Umer Farooq, Atifa Athar, Muhammad Ahsan Raza, Taher M Ghazal, and Shadman Sakib. 2024. A critical review of artificial intelligence based approaches in intrusion detection: A comprehensive analysis. *Journal of Engineering* 2024, 1 (2024), 3909173.

[36] Vu-Duc Ngo, Tuan-Cuong Vuong, Thien Van Luong, and Hung Tran. 2024. Machine learning-based intrusion detection: feature selection versus feature extraction. *Cluster Computing* 27, 3 (2024), 2365–2379.

[37] X.-H. Nguyen, X.-D. Nguyen, H.-H. Huynh, and K.-H. Le. 2022. Realguard: a Lightweight Network Intrusion Detection System for IoT Gateways. *Sensors* 22, 2 (2022), 432. https://doi.org/10.3390/s22020432

[38] Debanjana Niogi, Dr Devender Kumar, Deep Ranjan, and Jasveer Singh. 2023. Recent Advances and Future Directions in Ai-Based Intrusion Detection Systems for Network Security. *Available at SSRN 4485452* (2023).

[39] Mathappan Nivaashini, E Suganya, S Sountharrajan, M Prabu, and Durga Prasad Bavirisetti. 2024. FEDDBN-IDS: federated deep belief network-based wireless network intrusion detection system. *EURASIP Journal on Information Security* 2024, 1 (2024), 8.

[40] Uneneibotejit Otokwala, Andrei Petrovski, and Harsha Kalutarage. 2024. Optimized common features selection and deep-autoencoder (OCFSDA) for lightweight intrusion detection in Internet of things. *International Journal of Information Security* 23, 4 (2024), 2559–2581.

[41] Jeng-Shyang Pan, Fang Fan, Shu-Chuan Chu, Hui-Qi Zhao, and Gao-Yuan Liu. 2021. A lightweight intelligent intrusion detection model for wireless sensor networks. *Security and communication Networks* 2021, 1 (2021), 5540895.

[42] Vinh Pham, Eunil Seo, and Tai-Myoung Chung. 2020. Lightweight Convolutional Neural Network Based Intrusion Detection System. *J. Commun.* 15, 11 (2020), 808–817.

[43] A Ponmalar and V Dhanakoti. 2022. An intrusion detection approach using ensemble support vector machine based chaos game optimization algorithm in big data platform. *Applied Soft Computing* 116 (2022), 108295.

[44] Davy Preuveneers, Ilias Tsingenopoulos, and Wouter Joosen. 2020. Resource usage and performance trade-offs for machine learning models in smart environments. *Sensors* 20, 4 (2020), 1176.

[45] Sumit Pundir, Mohammad Wazid, Devesh Pratap Singh, Ashok Kumar Das, Joel JPC Rodrigues, and Youngho Park. 2019. Intrusion detection protocols in wireless sensor networks integrated to Internet of Things deployment: Survey and future challenges. *IEEE Access* 8 (2019), 3343–3363.

[46] Emad Ul Haq Qazi, Muhammad Hamza Faheem, and Tanveer Zia. 2023. HDLNIDS: hybrid deep-learning-based network intrusion detection system. *Applied Sciences* 13, 8 (2023), 4921.

[47] Md Mahbubur Rahman, Shaharia Al Shakil, and Mizanur Rahman Mustakim. 2025. A survey on intrusion detection system in IoT networks. *Cyber Security and Applications* 3 (2025), 100082.

[48] Ujwala Ravale, Nilesh Marathe, and Puja Padiya. 2015. Feature selection based hybrid anomaly intrusion detection system using K means and RBF kernel function. *Procedia Computer Science* 45 (2015), 428–435.

[49] B Ravinder Reddy. 2024. Network Intrusion Detection using Machine Learning. (2024).

[50] Vignesh Reddy, R Sunitha, M Anusha, S Chaitra, and Abhilasha P Kumar. 2024. Artificial Intelligence Based Intrusion Detection Systems. In *2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC)*. IEEE, 1–6.

[51] Bipraneel Roy and Hon Cheung. 2018. A deep learning approach for intrusion detection in internet of things using bi-directional long short-term memory recurrent neural network. In *2018 28th international telecommunication networks and applications conference (ITNAC)*. IEEE, 1–6.

[52] A.H. Salem, S.M. Azzam, O.E. Emam, and A.A. Abohany. 2024. Advancing cybersecurity: A Comprehensive Review of AI-driven Detection Techniques. *Journal of Big Data* 11, 1 (2024). https://doi.org/10.1186/s40537-024-00957-y

[53] Azar Abid Salih and Adnan Mohsin Abdulazeez. 2021. Evaluation of classification algorithms for intrusion detection system: A review. *Journal of Soft Computing and Data Mining* 2, 1 (2021), 31–40.

[54] K Samunnisa, G Sunil Vijaya Kumar, and K Madhavi. 2023. Intrusion detection system in distributed cloud computing: Hybrid clustering and classification methods. *Measurement: Sensors* 25 (2023), 100612.

[55] Pilar Schummer, Alberto del Rio, Javier Serrano, David Jimenez, Guillermo Sánchez, and Álvaro Llorente. 2024. Machine Learning-Based Network Anomaly Detection: Design, Implementation, and Evaluation. *AI* 5, 4 (2024), 2967–2983.

[56] Alhassan Seiba, Gaddafi Abdul-Salaam, Yaw Missah, and Mohammad Hossein Anisi. 2023. HYBRID NETWORK INTRUSION DETECTION SYSTEMS: A SYSTEMATIC REVIEW. *Scientific and practical cyber security journal* (2023).

[57] BS Sharmila and Rohini Nagapadma. 2023. Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset. *Cybersecurity* 6, 1 (2023), 41.

[58] Amardeep Singh and Julian Jang-Jaccard. 2022. Autoencoder-based unsupervised intrusion detection using multi-scale convolutional recurrent networks. *arXiv preprint arXiv:2204.03779* (2022).

[59] Insoo Sohn. 2021. Deep belief network based intrusion detection techniques: A survey. *Expert Systems with Applications* 167 (2021), 114170.

[60] Youngrok Song, Sangwon Hyun, and Yun-Gyung Cheong. 2021. Analysis of autoencoders for network intrusion detection. *Sensors* 21, 13 (2021), 4294.

[61] T. Sowmya and E.A. Mary Anita. 2023. A Comprehensive Review of AI Based Intrusion Detection System. *ScienceDirect* 28 (2023), 100827–100827. https://doi.org/10.1016/j.measen.2023.100827

[62] Huaglory Tianfield. 2017. Data mining based cyber-attack detection. *System simulation technology* 13, 2 (2017).

[63] Miracle Udurume, Vladimir Shakhov, and Insoo Koo. 2024. Comparative Evaluation of Network-Based Intrusion Detection: Deep Learning vs Traditional Machine Learning Approach. In *2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 520–525.

[64] Patrick Vanin, Thomas Newe, Lubna Luxmi Dhirani, Eoin O'Connell, Donna O'Shea, Brian Lee, and Muzaffar Rao. 2022. A study of network intrusion detection systems using artificial intelligence/machine learning. *Applied Sciences* 12, 22 (2022), 11752.

[65] Zhendong Wang, Yong Zeng, Yaodi Liu, and Dahai Li. 2021. Deep belief network integrating improved kernel-based extreme learning machine for network intrusion detection. *IEEE Access* 9 (2021), 16062–16091.

[66] Peng Wei, Yufeng Li, Zhen Zhang, Tao Hu, Ziyong Li, and Diyang Liu. 2019. An optimization method for intrusion detection classification model based on deep belief network. *Ieee Access* 7 (2019), 87593–87605.

[67] Yihan Xiao, Cheng Xing, Taining Zhang, and Zhongkai Zhao. 2019. An intrusion detection model based on feature reduction and convolutional neural networks. *IEEE Access* 7 (2019), 42210–42219.

[68] Chuanlong Yin, Yuefei Zhu, Jinlong Fei, and Xinzheng He. 2017. A deep learning approach for intrusion detection using recurrent neural networks. *Ieee Access* 5 (2017), 21954–21961.

[69] Guangzhen Zhao, Cuixiao Zhang, and Lijuan Zheng. 2017. Intrusion detection using deep belief network and probabilistic neural network. In *2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing (EUC)*, Vol. 1. IEEE, 639–642.

[70] Ruijie Zhao, Guan Gui, Zhi Xue, Jie Yin, Tomoaki Ohtsuki, Bamidele Adebisi, and Haris Gacanin. 2021. A novel intrusion detection method based on lightweight neural network for internet of things. *IEEE Internet of Things Journal* 9, 12 (2021), 9960–9972.

Received 28 March 2025; revised TBC; accepted TBC