Review Article

Melad Mohammed Issa, Mohammad Aljanabi*, and Hassan M. Muhialdeen

Systematic literature review on intrusion detection systems: Research trends, algorithms, methods, datasets, and limitations

https://doi.org/10.1515/jisys-2023-0248
received November 11, 2023; accepted December 29, 2023

Abstract: Machine learning (ML) and deep learning (DL) techniques have demonstrated significant potential in the development of effective intrusion detection systems. This study presents a systematic review of the utilization of ML, DL, optimization algorithms, and datasets in intrusion detection research from 2018 to 2023. We devised a comprehensive search strategy to identify relevant studies from scientific databases. After screening 393 papers meeting the inclusion criteria, we extracted and analyzed key information using bibliometric analysis techniques. The findings reveal increasing publication trends in this research domain and identify frequently used algorithms, with convolutional neural networks, support vector machines, decision trees, and genetic algorithms emerging as the top methods. The review also discusses the challenges and limitations of current techniques, providing a structured synthesis of the state-of-the-art to guide future intrusion detection research.

Keywords: SLR, IDS, cybersecurity

1 Introduction

Machine learning (ML) and deep learning (DL) techniques are transforming intrusion detection systems (IDS) [1–3], enabling enhanced security, adaptability, and scalability. Network intrusions through malicious attacks can disrupt services and operations, necessitating intelligent detection systems capable of identifying known and unknown threats. This has motivated extensive research on applying advanced ML and DL algorithms for IDS over the past decade. However, the rapid growth of studies presents challenges in synthesizing state-of-the-art advancements in a structured manner. This study provides a systematic literature review of ML- and DL-based intrusion detection techniques published between 2018 and 2023. A comprehensive search strategy is devised to survey recent studies from major scientific databases. After screening and analyzing 393 qualified papers, we extracted key information to understand publication trends, frequently adopted algorithms, datasets, limitations, and future challenges. Bibliometric analysis techniques help visualize research themes, prominent authors, and frequently studied algorithms like convolutional neural networks (CNNs), support vector machines (SVMs), XGBoost, and genetic algorithms (GAs) [1–5]. The structured taxonomy of the review categorizes techniques into four broad categories: ML, DL, optimization algorithms and datasets. Detailed

*** Corresponding author: Mohammad Aljanabi,** Department of Computer, College of Education, Al-Iraqia University, Baghdad, 10053, Iraq; Department of Computer, Imam Ja'afar Al-Sadiq University, Baghdad, 10052, Iraq, e-mail: mohammad.aljanabi@ijsu.edu.iq
Melad Mohammed Issa: Department of Computer Engineering, Al-Iraqia University, Baghdad, 10053, Iraq, e-mail: melad.eng@aliraqia.edu.iq
Hassan M. Muhialdeen: Department of Computer Engineering, Al-Iraqia University, Baghdad, 10053, Iraq, e-mail: muhialdeen.hassan@aliraqia.edu.iq

sub-categorizations are presented to summarize advancements within each domain. The findings reveal SVM, CNN, decision trees (DTs), and GAs as leading techniques for attaining high classification performance [6–10]. Comparative analysis provides insights into the relative effectiveness and limitations of different algorithms and datasets. This systematic review consolidates scattered developments into an organized synthesis to benefit researchers and practitioners working on ML/DL-driven IDS. By clarifying the state-of-the-art, it can guide selection of appropriate algorithms and datasets while also highlighting open challenges for advancing intelligent anomaly detection. The taxonomy of techniques provides a starting point for new researchers to swiftly comprehend key concepts, methods, and terminology in this rapidly progressing field. The systematic literature review of IDS adds significant value to the field by addressing gaps, challenges, and establishing its significance in advancing the domain of cybersecurity. Here is how the research achieves these objectives: (1) Addressing gaps: The review consolidates scattered developments into an organized synthesis, providing a structured taxonomy of techniques and categorizations within the domain of IDS. By identifying key concepts, methods, and terminology, the review serves as a starting point for new researchers to swiftly comprehend the rapidly progressing field of trustworthy ML. The study offers insights into the relative effectiveness and limitations of different algorithms and datasets, addressing the need for comparative analysis in the field of IDS. (2) Adding value: The comprehensive science mapping analysis helps organize the outcomes of previous investigations, summarizes key issues, and identifies potential research gaps, contributing to the existing body of knowledge. The review provides valuable insights into the conceptual framework of trustworthy ML, benefiting practitioners, policymakers, and academics, thus adding value to the understanding of the current state of research in this area. By visualizing research themes, prominent authors, and frequently studied algorithms, the review offers a comprehensive overview of the research landscape in IDS, thereby adding value to researchers and practitioners working in this domain. (3) Establishing significance: The study establishes significance by offering insights into the annual scientific production and advancements in reliable ML in intrusion detection over the past 10 years, highlighting the continuous evolution and significance of the field. Through the utilization of bibliometric analysis techniques and the extraction of key information from a large number of qualified papers, the review establishes the significance of its findings in understanding publication trends, frequently adopted algorithms, datasets, limitations, and future challenges in the field of IDS. The identification of the most popular and crucial keywords from previous research using a word cloud adds significance by highlighting the broad and diverse nature of the field of trustworthy ML, covering a wide range of subjects and applications. In summary, the systematic literature review of IDS adds value by addressing gaps, providing valuable insights, and establishing its significance in advancing the field of cybersecurity through comprehensive analysis and synthesis of research findings. The reviewed literature has identified several challenges and limitations of current intrusion detection techniques. These include: (1) Limited availability of labeled datasets: The availability of labeled datasets is a significant challenge in intrusion detection research, as it limits the ability to train and evaluate ML models effectively. (2) Imbalanced datasets: Imbalanced datasets, where the number of samples in one class is significantly higher than the other, can lead to biased models and reduced performance. (3) Adversarial attacks: Adversarial attacks, where attackers intentionally manipulate data to evade detection, pose a significant challenge to IDS. (4) Interpretability and explainability: The interpretability and explainability of ML models are crucial in intrusion detection, as it is essential to understand how the models make decisions and identify potential vulnerabilities. (5) Scalability: The scalability of IDS is a significant challenge, particularly in large-scale networks, where the volume of data can be overwhelming. These challenges and limitations can significantly influence the overall effectiveness and reliability of IDS. For instance, limited availability of labeled datasets can lead to poorly trained models, while imbalanced datasets can result in biased models that perform poorly on underrepresented classes. Adversarial attacks can evade detection and compromise the security of the system, while the lack of interpretability and explainability can limit the ability to identify and address potential vulnerabilities. Finally, scalability challenges can limit the ability to deploy IDS in large-scale networks, reducing their overall effectiveness and reliability. Overall, addressing these challenges and limitations is crucial for enhancing the effectiveness and reliability of IDS, and future research should focus on developing solutions to overcome these obstacles.

2 Methodology

In this analytical section (Figure 1), we followed the recommended reporting guidelines for a systematic review and meta-analysis technique. The procedure entailed the utilization of several bibliographic citation databases encompassing a broad spectrum of medical, scientific, and social science periodicals across various domains. Specifically, we considered three prominent digital databases: Scopus, IEEE Xplore, and Web of Science when searching for the target papers.

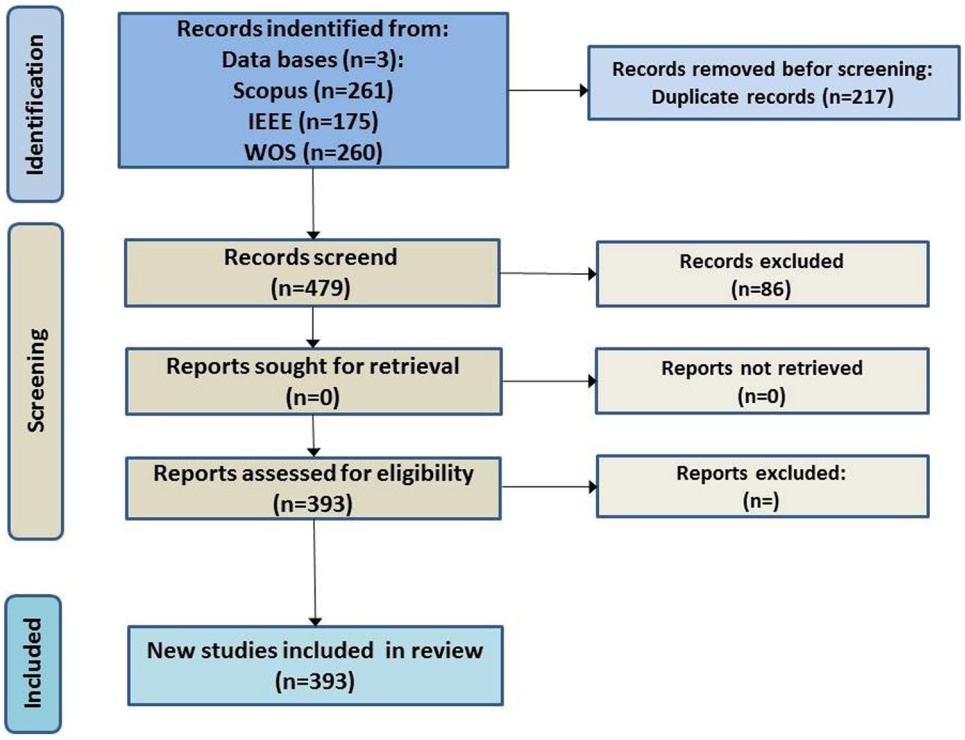

Figure 1: PRISMA protocol.

Scopus is renowned for its reliable resources across a wide array of disciplines, encompassing engineering, technology, science, medicine, and health. IEEE is a comprehensive repository of technical and scientific literature, offering full-text articles and abstracts across various publications in the fields of computer science, electronics, and electrical engineering. On the other hand, the Web of Science (WoS) database is a cross-disciplinary resource that incorporates research papers from diverse fields, including science, technology, art, and social science. These databases collectively offer extensive coverage of research in all scientific and technological domains, delivering valuable insights to researchers.

2.1 Search strategy

The three databases under consideration (Scopus, IEEE, and WoS) each underwent a thorough bibliographic search for academic papers written in English. All scientific articles production from 2018 to 2023 were included in this search.

This search used a Boolean query to link the keywords trustworthy, using two operator (AND) and (OR) as follows: “Intrusion detection system OR IDS” (AND) “machine learning OR deep learning OR classification algorithms” (AND) “optimisation algorithm OR optimization algorithm”).

2.2 Inclusion and exclusion criteria

The most crucial aspect of this conducted systematic review of the literature is the criteria taken into consideration for the inclusion/selection of studies (Figure 1). And for this, the following parameters were taken into account:

The papers had to be written in English, published in a journal or conference proceedings, and they had to take into account one or more reliable elements.

- Each of the components required to be considerably connected to reliable ML or DL in order to be integrated into various ML techniques/methods for the intrusion detection domain.
- Highly relevant articles that were published from 2018 to 2023.

On the other hand, papers beyond the purview of this investigation were omitted based on the following exclusion criteria:

- Papers that were not from peer reviewed publication forums.
- Papers that were not written in English.

2.3 Study selection

This study adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement for conducting a systematic literature review. The methodology involved several phases, with the initial phase focusing on the removal of duplicate papers. To achieve this, the titles and abstracts of the contributions were screened using the Mendeley program. All authors participated in this process, leading to the exclusion of numerous unrelated works. Disagreements and discrepancies among the authors were resolved by the corresponding author. The third step entailed a detailed examination of the full texts, during which articles failing to meet the previously established inclusion criteria (as outlined in Section 2.2) were eliminated. Only studies meeting the stipulated requirements were incorporated into this research. The initial search yielded 696 entries, of which 261 originated from Scopus, 175 from IEEE, and 260 from WoS. After removing approximately 217 duplicates and carefully scrutinizing the remaining entries, 393 studies remained. According to the inclusion criteria, studies were deemed relevant and subsequently included in the final collection of publications. The analysis of these gathered articles is subject to various bibliometric techniques, which are discussed in Section 3. The bibliometric analysis conducted in the systematic literature review involved the extraction and analysis of key information from the 393 papers meeting the inclusion criteria. The analysis aimed to understand publication trends, frequently adopted algorithms, datasets, limitations, and future challenges in the field of IDS. Key information extracted from the papers included: (1) Publication trends in the field of IDS from 2018 to 2023, indicating the number of published papers each year. (2) Frequently adopted ML, DL, and optimization algorithms for intrusion detection. (3) Utilized benchmark datasets for evaluating IDS. (4) Limitations and future challenges identified in the reviewed studies. To assess the significance and impact of the identified studies, bibliometric analysis techniques were used to visualize research themes, prominent authors, and frequently studied algorithms like CNN, SVM, XGBoost, and GA. Additionally, the structured taxonomy of the review techniques were categorized into four broad categories: ML, DL, optimization algorithms, and datasets, with detailed sub-categorizations to summarize advancements within each domain. The bibliometric analysis provided insights into the relative effectiveness and limitations of different algorithms and datasets, consolidating scattered developments into an organized synthesis to benefit researchers and practitioners working on ML- and DL-driven IDS.

3 Comprehensive science mapping analysis

Finding crucial information in previous studies has become increasingly challenging due to the growing volume of contributions and applied research. Keeping pace with this vast stream of theoretical and practical input can be daunting. Managing the vast literature has posed a significant challenge. To help organize the outcomes of previous investigations, summarize key issues, and identify potential research gaps, some academics recommend employing the PRISMA approach. Systematic reviews, in comparison, bolster the study's framework, contribute to the existing body of knowledge, and amalgamate the literature's findings. Nonetheless, systematic reviews still contend with issues of impartiality and reliability, as they depend on the authors' perspective to restructure the results of previous investigations. Various studies have proposed the following measures to enhance transparency when summarizing the findings of earlier studies.

3.1 Annual scientific production

In the previous 10 years, reliable ML in intrusion detection has advanced. In particular, the annual scientific output depicted in Figure 2 provides an explanation for the emergence of earlier theoretical and practical investigations on reliable ML. The annual scientific output for intrusion detection is depicted in Figure 2. The number of papers published in 2018 and 2019 reached approximately 23 papers, it can be seen that the quantity of publications has significantly expanded in recent years. There was an increase in the number of articles published in 2020 and 2021. In 2022, the number of articles grew even further, reaching a notable high of 138 papers. This pattern persisted in 2023, where 95 papers were published. The increase in research output suggests a growing interest and emphasis on the development and improvement of IDS over the years. Furthermore, specific authors have contributed significantly to this field, with some focusing on optimization and feature selection based on intrusion detection, while others have concentrated on IDS for Internet of Things (IoT) based on DL. These authors have achieved high accuracies in their research, indicating the advancement and effectiveness of the techniques employed in IDS. Overall, the observed publication trends demonstrate a substantial growth in research output in the field of IDS from 2018 to 2023, reflecting an increasing focus on enhancing the reliability and effectiveness of intrusion detection techniques. Figure 3 shows authors' production over time, where Motwakel [4–10] published seven papers in 2022 and 2023 and he focused in his research on optimization and feature selection based on intrusion detection. In his research, the highest accuracy he reached was 99.87 using sand paper optimization. As for Al_Qaness [11–15], he published five papers in 2021–2023) and he focused on IDS for IOT based on DL. In his research, he reached the highest

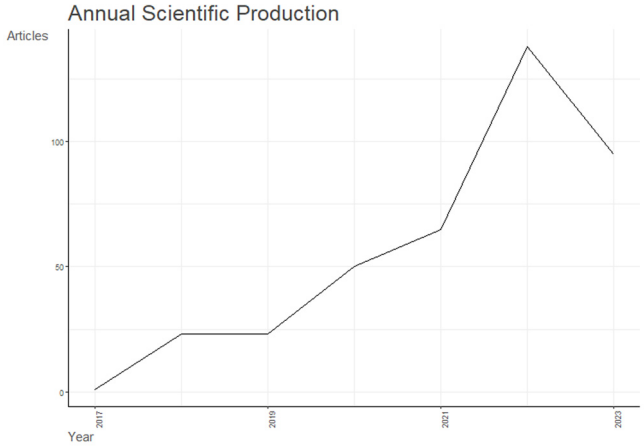

Figure 2: Annual scientific production.

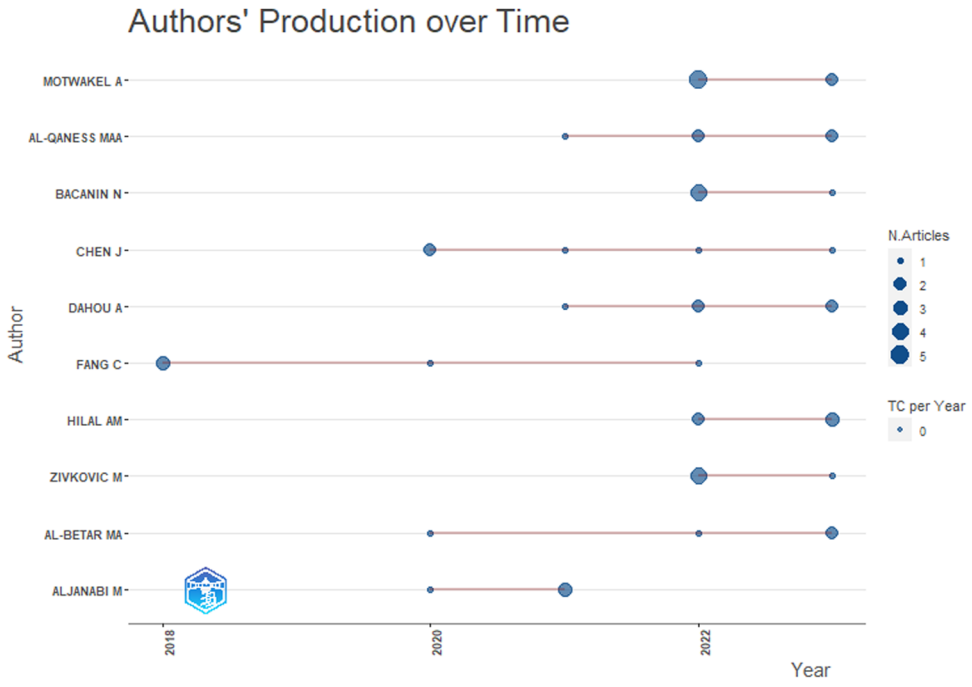

Figure 3: Authors' production over time.

accuracy of 99.997 using swarm intelligence optimization. Bacanin [16–20] has published five papers from 2020 to 2023 and he focused on optimization algorithms and feature selection based on intrusion detection in his research, the highest accuracy he reached was 99.6878 using GOA and MPO. Chen [21,22] published five papers from 2020 to 2023 and he focused on intrusion detection using hybrid algorithms, the highest accuracy he reached was 99.44 using COBYLA optimization. Dahou [11–15] published five papers, one paper in 2021, two papers in 2022, and two papers in 2023 and he focused on IOT IDS using DL and optimization in his research, the highest accuracy he reached was 99.997 using swarm intelligence optimization. Fang [23–27] published five papers in 2018, 2020, and 2022 and he focused on optimization algorithm for feature selection of network intrusion detection in his research, the highest accuracy he reached was 97.89 using WOA and OPS optimization. Hilal [4,28–31] published five in 2022 and 2023, he focused on DL algorithms optimization based on intrusion detection in his research, the highest accuracy he reached was 99.77 using optimization IFSO-FS. Zivkovic [16–20] published five papers, three papers in 2022 and two papers in 2023, he focused on intrusion detection for optimization algorithms and feature selection in his research, the highest accuracy he reached was 99.6878 using GOA and MPO. Dahou [32–35] published four papers in 2020, 2022, and 2023) and he focused on IDS for optimization algorithms in his research, the highest accuracy he reached was 100 using Artificial organism algorithm (AOA) optimization. In addition to the mentioned authors, Al-Janabi [36–39] has contributed significantly by publishing four papers, with one in 2020 and three in 2021. His research was focused on IDS, optimization algorithms, and feature selection. Remarkably, his work achieved the highest accuracy of 100% using NTLBO optimization. Table 1 presents a comprehensive overview of the most influential authors in the field. Each of these authors has demonstrated exceptional achievements by reaching the highest accuracy through the utilization of classification and optimization algorithms.

3.2 Three-field chart

A three-field chart is used to display data with three parameters. In this representation, the left field corresponds to the Research Title (RT), the middle field represents the Journal in which the Research is published or source (SO), and the right field contains the Researcher's Name (RN). Figure 4, is utilized to examine the

Table 1: Highest accuracy studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Motwakel et al. [4]	99.87	IGAN-OKELM	Sequential parameter optimization (SPO)	CICIDS2015	Optimal parameter	Binary and multi-class classification	During the upgrade phase, it is essential to implement the SPO behavior to optimize the location. This approach should be designed to strike a balance between minimizing computational complexity, maintaining accuracy, and reducing the overall cost function.
2.	Al-qaness et al. [11]	99.997	Recurrent neural network (RNN), ANN	Swarm intelligence	CIC2017	Feature selection	Binary and multi-class classification	The developed method exhibits a high time complexity, particularly during the feature extraction phase, and also during the process of selecting relevant features. This results in extended processing times and may hinder efficiency. The Capuchin search algorithm (CSA) suffers from slow convergence, which ultimately diminishes the quality of solutions it produces.
3.	Bacanin et al. [16]	99.6878	XGBoost	GOA, MPO	USNW-NB15	Hyper parameter	Binary	The framework was assessed using a single dataset, underscoring the necessity for additional evaluations on different datasets to validate its overall effectiveness.
4.	Chen [32]	99.44	XGBoost	COBYLA	NSL-KDD	Feature selection	Multi-class classification	Stronger defenses are consistently required to address all types of adversarial scenarios, often necessitating the retraining of models with larger training datasets, both of which are time-consuming processes. Despite these challenges, classifiers typically exhibit poor generalization.
5.	Fang et al. [23]	97.89	SVM	WOA, Particle swarm optimization (PSO)	KDD CUP 99	Feature selection	Binary	The algorithm still faces challenges related to delayed convergence and low accuracy. One of the underlying issues is the optimization process itself. The basic Elephant Herding Optimization method lacks an efficient mutation mechanism. This deficiency often leads individuals to get trapped in local extreme values, resulting in premature convergence.

(Continued)

Table 1: *Continued*

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
								Furthermore, the position of the best individual significantly influences how the algorithm updates the positions of other individuals. When a local extreme value attracts the individual currently deemed the best, it can impact the overall performance of the algorithm.
6.	Hilal et al. [4]	99.77	BiLSTM	IFSO-FS	AOEDBC-DL	Feature selection	Multi-class classification	Using a predetermined number of variables is a required process. Optimization variables are the same as FS characteristics.
7.	Dahou [32]	100	XGB, DT, ET	AOA	NF-BOT-IOT-V2	Feature selection	Binary and multi-class classification	An inappropriate balance between exploration and exploitation throughout the search can cause the AOA to become trapped in an early state of convergence. An additional challenge is the computational time, as we must perform the ML approach at every AOA iteration, which necessitates substantial processing resources.
8.	Al-Janabi and Ismail [36]	100	SVM, Extreme learning machine (ELM), LR	NTLBO	KDD CUP 99	Feature selection	Binary	The authors acknowledge that the proposed method might not be well-suited for real-time intrusion detection, primarily because of the time required for feature selection and model training.

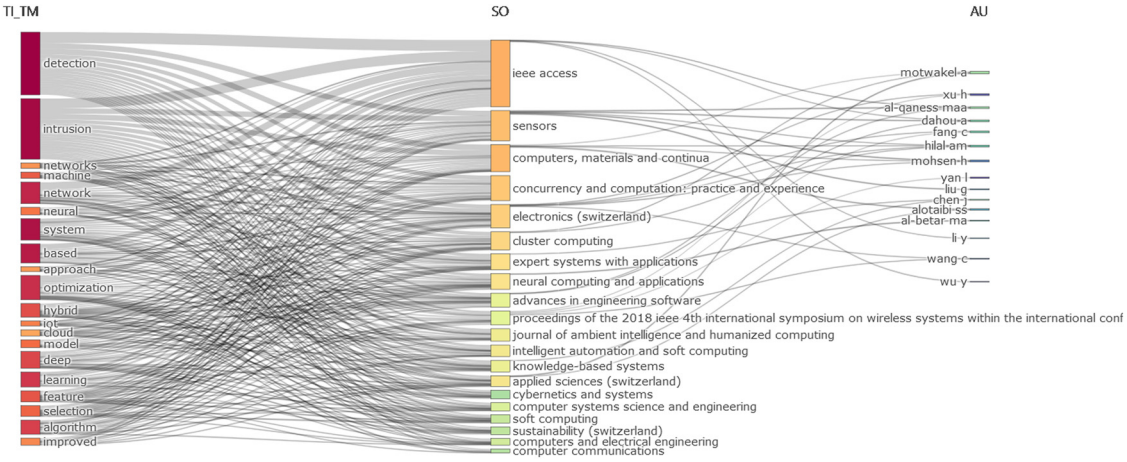

Figure 4: Three-field plot: left (TI), middle (SO), and right (AU).

relationships between these three parameters. According to the study, the RT on the left side is most frequently cited by Scopus, IEEE Xplore (IEEE), and WoS, as observed in the middle field (SO) of Figure 4. Furthermore, among the Research Titles (TI) that focus on the subject of reliable and understandable ML, the Scopus journal stands out as the most prominent. Additionally, as indicated in the corresponding box (TI), when considering all keywords, the journals listed in the middle field (SO) most frequently match the most popular keywords, which include “IEEE Access,” “Sensors,” “Cluster Computing,” “Neural Computing and Application,” and “Soft Computing.”

3.3 Word cloud

This study has effectively identified the most popular and crucial keywords from previous research using a word cloud. In order to provide a comprehensive overview of these keywords and reorganize the information, Figure 5 presents these essential keywords extracted from the results of previous studies. In Figure 5, the keywords are displayed in different sizes, with the size indicating their frequency in the literature. Larger keywords are more prevalent, while smaller keywords occur less frequently. Based on the term frequency illustrated in Figure 5, “ID,” “DL,” and “ML” emerge as some of the most frequently discussed topics in the field of trustworthy ML, with “DL” having the highest frequency. The image also highlights the significance of “IDS” and “Intrusion Detection System” as critical topics in this area. Additionally, other related terms with relatively high frequencies include “classification,” “optimization,” and “feature selection,” emphasizing the importance

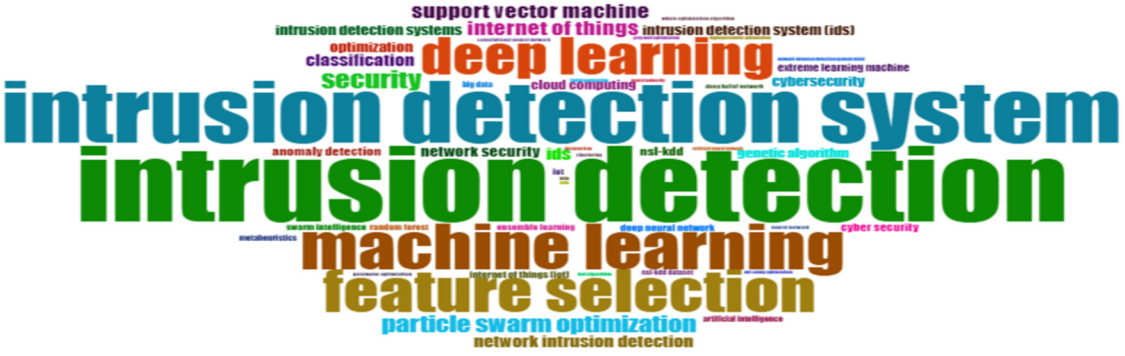

Figure 5: Word cloud.

of considering these factors when designing and implementing ML systems. Figure 5 also showcases various specific applications of ML across a range of algorithms, such as GA, SVM, and PSO. Furthermore, it includes several methodologies related to ML, including “Internet of Things,” “network security,” and “artificial intelligence.” The word cloud of trustworthy ML articles reveals the broad and diverse nature of this field, covering a wide range of subjects.

3.4 Co-occurrence

Another method employed in bibliometric analysis is the co-occurrence network, which involves studying common words in earlier research. This semantic network offers valuable insights into the conceptual framework of a specialized field, benefiting practitioners, policymakers, and academics. Figure 6 presents information specifically related to a co-occurrence network based on the titles of reliable ML articles.

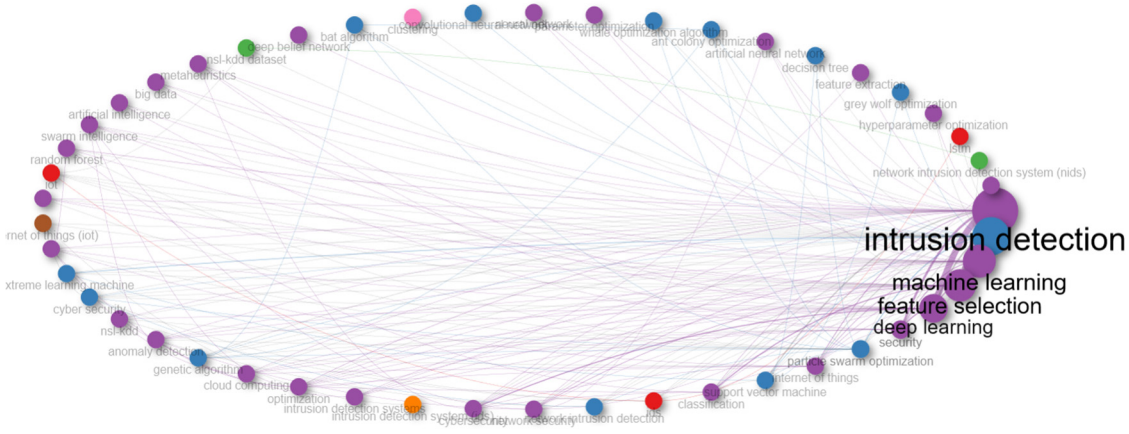

Figure 6: Co-occurrence network.

The network is composed of nodes, representing individual words in the titles, and edges connecting the nodes indicate how frequently these words appear together in a title. Figure 6 displays various nodes, including the clusters to which they belong and their proximity, a metric for gauging how well-connected a node is to other nodes in the network. Evidently, the nodes are divided into eight distinct clusters, with each cluster comprising words associated with a specific theme or idea related to trustworthy ML. For instance, Cluster 1 features terms like “support vector machine,” “DT,” “Extreme learning machine,” “cyber security,” “genetic algorithm,” and “bat algorithm.” These terms suggest a connection to the establishment of dependable ML systems in optimization. Cluster 2, on the other hand, includes words like “neural network,” “artificial intelligence,” “feature extraction,” “random forest,” “anomaly detection,” and “network security,” signifying a focus on IDS. Similarly, other clusters are linked to subjects such as clustering, the IoT, the NSL-KDD dataset, ML, and IDS. The closeness of a node within the network measures its centrality, and nodes with higher closeness values are more central to the topic of trustworthy ML, signifying their importance within the network. This figure offers an overview of the relationships between different concepts and words associated with trustworthy ML, as derived from the titles of papers in the field. This information is valuable for understanding the current state of research in this area and for identifying areas where further investigation is warranted.

4 Findings and analysis: A taxonomy

The final set of papers met both the considered inclusion and exclusion criteria, indicating that ML in intrusion detection has been identified through the conducted procedure (see Section 2.2). Additionally, out of the 393 articles included, these were categorized into three broad categories. After analyzing each category, efforts were made to identify or generate subcategories using a variety of reliable ML algorithms in the context of intrusion detection. Within the first major category of 393 papers, we found

1. ML, comprising 249 papers
2. DL, with 144 papers
3. Optimization algorithms, covering all 393 papers
4. Dataset.

We also have a section displaying the taxonomy of subdivisions in Figure 7, which includes ML, DL, Optimization, and Dataset.

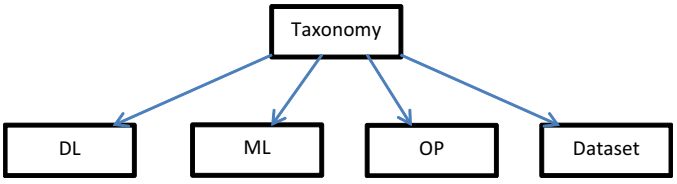

Figure 7: Taxonomy.

These categories are discussed here comprehensively, aiming to offer academics and practitioners valuable insights on trustworthy ML in intrusion detection. This endeavor is reported to enhance the reliability of ML in intrusion detection. Consequently, 249 out of the 393 articles fall under this category within the context of intrusion detection.

4.1 ML

Arthur Samuel originally described ML in 1959 as a branch of research that allows computers to learn without requiring them to first be programmed. This section describes network ID strategies with a focus on ML algorithms used to create security tools. In recent years, ML has gained increasing importance in IDS for computer networks [40–42]. The foundation of this lies in the model for training and prediction, which has the capability to quickly identify both attacks and typical cases [32]. The feature selection process can be considered as data preprocessing for ML algorithms. Intrusion detection can involve two types of classification: two-class, where intrusions are detected based on class labels, and multi-class, which categorizes attacks into different classes. ML techniques like Random forest (RF), SVM, ELM, and Naive Bayes classifiers can be applied in this field, as well as methods such as Self-Organizing Maps, Fuzzy clustering, and K-Means clustering. Figure 8 show the classification of ML algorithms.

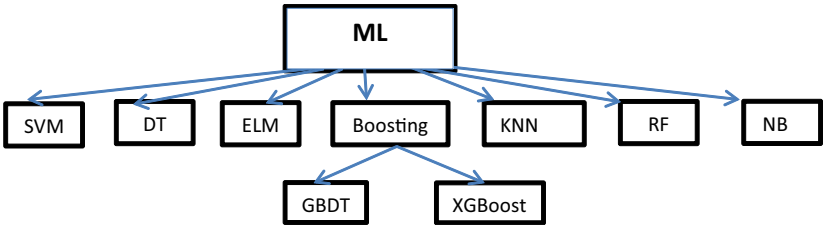

Figure 8: ML algorithms.

In the reviewed literature, several ML and DL algorithms have emerged as the most frequently used in intrusion detection research. Specifically, CNN, SVM, DTs, and GA have been prominently featured in the analyzed studies. Comparatively, these algorithms have demonstrated varying levels of popularity and effectiveness in the context of intrusion detection: (1) CNN: CNN has gained significant popularity and has been widely utilized in intrusion detection research due to its effectiveness in learning hierarchical representations of data, particularly in image-based intrusion detection scenarios. (2) SVM: SVM has also been frequently used and is known for its effectiveness in binary classification tasks, making it a popular choice for intrusion detection applications. (3) DTs: DTs have been commonly employed for their interpretability and ease of understanding, making them popular in certain intrusion detection contexts, especially when explainability is a priority. (4) GA: While GAs have been utilized, they may not be as prevalent as CNN, SVM, and DT in intrusion detection research. However, they offer the advantage of optimization and search capabilities, which can be beneficial in certain scenarios. In terms of effectiveness, the reviewed literature may provide insights into the comparative performance of these algorithms in specific intrusion detection contexts, such as their accuracy, precision, recall, and $F1$ scores. Additionally, the specific datasets and features used in the studies may influence the relative effectiveness of these algorithms. Overall, while CNN, SVM, and DT have emerged as popular and effective choices in intrusion detection research, the comparative effectiveness of these algorithms may vary depending on the specific context, dataset, and evaluation metrics used in the reviewed studies.

4.1.1 SVM

In 1998, Vapnik Chih-Fong Tsai introduced the SVM. The SVM begins by transforming the input vector into a higher-dimensional feature space and subsequently identifies the optimal separating hyperplane. What distinguishes the SVM is its creation of a decision boundary, or separation hyperplane, using support vectors rather than the entire training sample. This property makes it highly robust against outliers. SVM classifiers are tailored for binary classification, meaning they are designed to divide a set of training vectors into two distinct classes. It is worth noting that the support vectors represent the training samples at the decision boundary. Additionally, the SVM incorporates a user-specified parameter known as the penalty factor, allowing users to strike a balance between the number of incorrectly classified samples and the width of the decision boundary [1,4,43,44].

Table 2 highlights the top five authors globally, who utilized SVM algorithms in their research, each achieving the highest accuracy using SVM classification and optimization algorithms. Alqarni [45] achieved the highest accuracy of 100%, followed by Aljanabi and Ismail [36] at 100%. Lavanya and Kannan [46] reached an accuracy of 99.98%, while Dwivedi et al. [47] achieved 99.89%, and Liu et al. [48] reached an accuracy of 99.88%.

4.1.2 DT

Chih-Fong Tsai utilizes a sequence of decisions to categorize a sample, with each decision influencing the subsequent one. These decisions are represented in the form of a tree structure. When classifying a sample, you start at the root node and traverse the tree until you reach an end leaf node, each of which represents a distinct classification category. At each node, the sample's characteristics are considered, and the branch value matches the attributes. Classification and Regression Tree (CART) is a well-known tool for creating DTs. A classification tree employs discrete (symbolic) class labels, while a regression tree deals with continuous (numeric) attributes [48].

Many researchers used DT algorithms in their research. Table 3 highlights the top five authors worldwide, each achieving the highest accuracy using DT classification and optimization algorithms. Dahou [32] achieved the highest accuracy at 100%, followed by Injadat et al. [49] at 99.99%. Mousavi et al. [50] reached an accuracy

Table 2: Highest accuracy of SVM algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Alqarni [45]	100%	SVM	Ant colony optimization (ACO)	KDD Cup 99	Feature selection	Binary and multi-class classification	It took a long time to reach the highest accuracy
2.	Aljanabi and Ismail [36]	100%	SVM	GA, TLBO	KDD Cup 99	Feature subset selection	Binary and multi-class classification	Greater parameter values result in increased accuracy, albeit at the cost of longer computation times. However, the extended time required for researching a new person is considerable.
3.	Lavanya and Kannan [46]	99.98	SVM	Krill herd	NSL-KDD 2015	Parameters	Binary and multi-class classification	Installing non-traditional IDS like VANET-based IDS in a VANET application requires caution to ensure real-time performance is not compromised. The survey explores solutions to VANET-related challenges, including increased false positives, reduced detection rates, higher network overhead, longer detection times, and associated issues. However, it may struggle to identify newer and modified attacks.
4.	Dwivedi et al. [47]	99.89	SVM	Grasshopper	KDD Cup 99	Feature selection	Binary	Despite employing security measures like cryptography and communication protocols, preventing invasions entirely remains a challenge. Detecting when a user's actions disrupt the intended use of computer networks is crucial.
5.	Liu et al. [48]	99.88	SVM, SSA	Swarm intelligence	KDD Cup 99	Feature selection	Multi-class classification	While Simulated Annealing has several advantages for optimizing various problems, it still faces issues with convergence accuracy and escaping local optima.

Table 3: Highest accuracy of DT algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Dahou [32]	100%	XGB, DT, ET	AOA	NF-BOT-IOT-V2	Feature selection	Binary and multi-class classification	Enhancing and validating the suggested method can be achieved through diverse datasets and parallel execution to reduce computation time. Furthermore, incorporating the proposed AOA with other effective elements can bolster the balance between exploration and exploitation, ultimately improving the results.
2.	Injadat et al. [49]	99.99	DT	PSO and Genetic	CICIDS 2017	Optimal parameter	Multi-class classification	Both datasets are initially skewed, containing far fewer attack samples than standard samples. Consequently, the model struggles to detect attack patterns and behaviors.
3.	Mousavi et al. [50]	99.92	DT	ACO	KDD Cup 99	Feature selection	Binary and multi class classification	The method used in this research to select a small training subset for multiclass classification under imbalanced conditions is currently challenging but not optimal. It lacks flexibility and efficiency, and there is a need for a better and more suitable approach to address this issue.
4.	Maza and Touahria [51]	99.83	DT, MOEDAFS	Multi-objective	NSL-KDD	Feature selection	Multi-class classification	Some crowd solutions that can restrict the algorithm's capacity for exploration are filtered away by MOEDAFS.
5.	Mahmood et al. [52]	99.36%	DT, SVM, K-nearest neighbors (KNN)	PSO and genetic	NSL-KDD	Feature selection	Multi-class classification	The experimental results suggest that reducing the number of features to a minimum, even if they are carefully chosen and relevant, does not always lead to higher accuracy. Instead, it is essential to select the right quantity of important and relevant features, which may even be a large number, to enhance the performance of ML models.

of 99.92, Maza and Touahria [51] reached an accuracy of 99.83%, and Mahmood et al. [52] reached an accuracy of 99.36%.

4.1.3 ELM

The ELM approach, introduced by Huang et al., is known for its speed and simplicity as it does not require iterative training. It consists of three layers: the input layer, a single hidden layer, and the output layer. ELM is specifically a single hidden layer feedforward neural network (SLFN) because it employs only one hidden layer. It excels at solving complex nonlinear mapping problems, and its adaptive training sets random input weights and biases for a number of nodes in the hidden layer utilized ELM algorithms in their research [36]. In Table 4, we highlight the top five authors globally, each achieving the highest accuracy using ELM classification and optimization algorithms. Al-Janabi and Ismail [36] achieved the highest accuracy at 100%, ElDahshan et al. [53] achieved the highest accuracy at 100%, followed by. Vaiyapuri et al. [54] reached an accuracy of 99.63%, while Ghasemi et al. [55] achieved 98.73%, and Wang et al. [56] reached 89.1%.

4.1.4 Boosting (Light gradient boosting machine; LGBM, XGBOOST, Gradient boosting decision tree; GBDT)

Boosting is a potent ensemble learning technique widely applied in IDS to enhance the performance of individual weak classifiers. It combines multiple weak classifiers to construct a strong classifier capable of effectively identifying intrusions. Notable boosting algorithms include LGBM, GBDT, and XGBoost. XGBoost, initially proposed by Tianqi Chen has gained widespread acceptance among researchers and developers. This technique applies boosting to machines, utilizing numerous weak learners like shallow DTs (typically of depth 1 or 2). Each learner learns from the errors of the preceding one, and the combination of many weak learners (often hundreds) forms a powerful final model [57]. The authors employed boosting algorithms in their research. Table 5 present the top five authors globally, each achieving the highest accuracy using Boosting classification and optimization algorithms. Dahou [32] reached the highest accuracy at 100%, Kilincer et al. [58] attained the accuracy at 99.98%, followed by Xu and Fan [59] who achieved an accuracy of 99.92%. Bacanin et al. [16] reached an accuracy of 99.65% and Zivkovic et al. [17] reached an accuracy of 99.68%.

4.1.5 KNN and RF

4.1.5.1 KNN

KNN is a supervised classifier where data are divided into K clusters based on the Euclidean distance between data points. The data points with the smallest distance are grouped together due to their shared properties. KNN is simple to use and effective for large datasets.

4.1.5.2 RF

RF is an ensemble method that combines multiple DTs to enhance model effectiveness. Bagging is employed to divide data into subsets, and DTs are built from these subgroups. RF is known for its low classification errors and absence of overfitting issues. Individual trees in the forest are constructed using bootstrap samples from the dataset. The Gini impurity measurement is used to determine the optimal node for splitting, and the model includes a maximum of 25 trees [67].

The authors utilized KNN and RF algorithms in their research. Table 6 presents the top five authors globally, each achieving the highest accuracy using KNN and RF classification and optimization algorithms. Gaber et al. [60] achieved the highest accuracy at 99.99%, followed by Samawi et al. [61] at 99.98%. Mohi-ud-din

Table 4: Highest accuracy of ELM algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Al-Janabi and Ismail [36]	100	SVM, ELM, LR	NTLBO	KDD CUP 99	Feature selection	Binary	The authors acknowledge that the proposed method might not be well-suited for real-time intrusion detection, primarily because of the time required for feature selection and model training.
2.	ElDahshan et al. [53]	100	ELM	Grey wolf optimization (GWO)	CICIDS 2017	Parameter	Binary and multi-class classification	Resolving these problems should prioritize a focus on attack instances over normal instances, as misclassifying attacks among attack instances can cause more significant harm than misclassifying attacks among normal instances.
3.	Vaiyapuri et al. [54]	99.63	ELM	SGOA	FedMCCS, TON_IoT	Feature selection	Multi-class classification	Due to the fact that this approach shares only taught methods of opinions prior to viewing specific local data, it reduces the transmission overhead of devices.
4.	Ghasemi et al. [55]	98.73	ELM	Genetic	-KDD cup 99	Feature selection	Multi-class classification	This technique is unable to accurately identify normal records. In KELM simulations, even though all features are considered, the results for attack labels are disappointing.
5.	Wang et al. [56]	89.1	ELM	*LSA-ELM*	KDD 99	Parameter	Multi-class classification	Initial biases and weights are randomly selected in this algorithm. The only parameter to determine is the total number of hidden nodes in the network. During its operation, the algorithm does not modify the network's input weights and the thresholds of the hidden components, aiming to achieve a specific optimization solution. This approach differs from alternative feedforward neural networks.

Table 5: Highest accuracy of Boosting algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Dahou [32]	100	XGB, DT, ET	AOA	NF-BOT-IOT-V2	Feature selection	Binary and multi-class classification	An inappropriate balance between exploration and exploitation throughout the search can cause the AOA to become trapped in an early state of convergence. An additional challenge is the computational time, as we must perform the ML approach at every AOA iteration, which necessitates substantial processing resources.
2.	Kilincer et al. [58]	99.98%	Boosting	Hyper-parameter	CICIDS 2017	Optimal feature	Multi-class classification	The dataset has an excessive amount of variables and observations, which causes the XGBoost training time to increase.
3.	Xu and Fan [59]	99.92%	XGBoost	PSO	UNSW-NB15	Feature selection	Multi class classification	It significantly underperforms in terms of runtime compared to most IDS models.
4.	Bacanin et al. [16]	99.65%	XGBoost	Artificial bee colony (ABC)	UNSW-NB15	Feature selection	Multi-class classification	One of the upcoming challenges in this domain is to validate the suggested hybrid model on additional intrusion detection datasets. This step is crucial for increasing confidence in the results before applying the model in real-world scenarios.
5.	Zivkovic et al. [17]	99.6878	XGBoost	GOA, MPO	USNW-NB15	Hyper parameter	Binary	The framework was assessed using a single dataset, underscoring the necessity for additional evaluations on different datasets to validate its overall effectiveness.

Table 6: Highest accuracy of RF algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Gaber et al. [60]	99.99%	RF	PSO and Bat	WUSTL-IIoT-2021	Feature selection	Multi-class classification	Limited availability of real-world data for evaluating IIoT system effectiveness. Use of unbalanced datasets in ML-based IDS, potentially resulting in minority attack detection failures. Examination of the suggested feature selection technique's performance using three ML algorithms (RF, KNN, and MLP). Lack of consideration for how different attack types may impact the suggested intrusion detection method's effectiveness.
2.	Samawi et al. [61]	99.98%	RF	SMO	NLS-KDD	Feature selection	Multi-class classification	Using the entire dataset for training, while resulting in a high accuracy of (99.98), is not ideal. Also, it is crucial to develop an IDS capable of identifying new types of intrusions.
3.	Mohi-ud-din et al. [27]	99.95%	RF	CSA-PSO	UNSW-NB15	Feature selection	Multi-class classification	To reach high-quality results, the algorithm CSA takes longer
4.	Bangui and Buhnova [62]	95.6%	RF	ACO	CICIDS2017	Feature selection	Multi class classification	It took a long time to analyze the comprehensive data to enhance its security against various attacks
5.	Mahmood et al. [52]	99.36%	DT, SVM, KNN	PSO and Genetic	NSL-KDD	Feature selection	Multi-class classification	The experimental results suggest that reducing the number of features to a minimum, even if they are carefully chosen and relevant, does not always lead to higher accuracy. Instead, it is essential to select the right quantity of important and relevant features, which may even be a large number, to enhance the performance of ML models.

Table 7: Highest accuracy of RF algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Shitharth et al. [63]	99.99%	NB	PPGO	NLS-KDD	Feature selection	Multi-class classification	Large amounts of noisy data can impair system performance by increasing false positives, misclassifying outcomes, and requiring a lot of time to train the model.
2.	Devi and Singh [64]	99.91%	NB	SMO	KDD Cup 99	Feature selection	Multi-class classification	SMO achieves a high accuracy rate but is not recommended due to the lengthy model-building process. NB, while quick to build, has poor accuracy, making it an unsuitable choice. Depending on our criteria for both speed and accuracy in detection, we can opt for either J48 or RF.
3.	Kunhare et al. [57]	99.32%	NB	PSO	NLS-KDD	Feature selection	Multi-class classification	Research on parameter optimization for the PSO algorithm is in its early stages. Regarding accuracy, it is observed that the detection accuracy varies slightly between iterations, by about 0.5%, until the 27th iteration. This small variation occurs as the search particles alternate between their personal and awareness states.
4.	Samriya et al. [65]	99.5%	NB	ACO	NLS-KDD	Feature selection	Binary and multi-class classification	Detecting anomalies in IoT networks and identifying malware in uncertain and overcast conditions can be time-consuming.
5.	Iwendi et al. [66]	98.81%	NB	Genetic	NSL-KDD	Feature selection	Multi-class classification	While the ROS algorithm produced overfitting and redundant data, the RUS algorithm resulted in the loss of usable data. Data intersection and noise traffic were created via SMOTE interruption, and the amount of complex samples in the training assembly.

et al. [27] reached an accuracy of 99.95%, Bangui and Buhnova [62] reached an accuracy of 95.6%, and Mahmood et al. [52] reached an accuracy of 99.36%.

4.1.6 Naïve Bayes (NB)

NB employs a probabilistic approach based on Bayes theorem and conditional probability calculations. It is referred to as "naïve" due to the simplifying assumption of predictor variable independence, meaning it assumes that all attributes are unrelated to each other. This class of methods includes those offering categorization functions without explicitly producing a tree or set of rules [68]. Many researchers used the NB algorithm in their research. Table 7 presents the top five authors worldwide, each achieving the highest accuracy using NB classification and optimization algorithms. Shitharth et al. [63] achieved the highest accuracy at 99.99%, followed by Devi and Singh [64] at 99.91%. Kunhare et al. [57] reached an accuracy of 99.32%, while Samriya et al. [65] achieved and accuracy of 99.5%, and Iwendi et al. [66] achieved an accuracy of 98.81%.

4.2 DL

DL, a subcategory of ML, consists of multiple hidden layers and finds applications in various domains, including image processing and natural language processing. It excels in understanding the meaning of vast multidimensional data, performing feature selection, classification, and uncovering data correlations, particularly in speech recognition and language processing [69] (Figure 9).

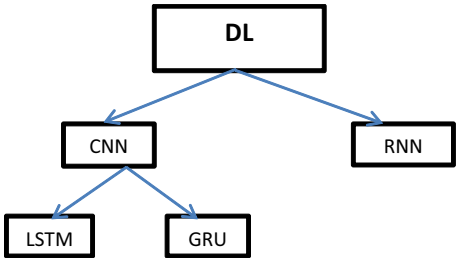

Figure 9: Deep learning algorithms.

The utilization of ML and DL techniques has significantly evolved in the development of IDS from 2018 to 2023. A systematic review of 393 studies revealed that ML and DL techniques have demonstrated significant potential in enhancing the reliability and effectiveness of IDS. The review identified frequently used algorithms, with CNN, SVM, DTs, and GA emerging as the top methods. Tables 2 and 3 and and 4 in the review indicate that SVM, DT, and ELM algorithms exhibit superior performance, particularly with the KDD Cup 1999 and NF-Bot datasets, both for multi-class and binary classification, as assessed by accuracy. In the realm of DL algorithms, Table 8 in the review showcases improved outcomes with the CNN algorithm and the NLS-KDD L dataset compared to Table 9, which demonstrates lower results with the RNN algorithm and the CICIDS2017 dataset, once again, gauged by accuracy. Overall, the utilization of ML and DL techniques has evolved significantly in the development of IDS, with CNN, SVM, DT, and GA emerging as the top methods. These techniques have demonstrated significant potential in enhancing the reliability and effectiveness of IDS.

Table 8: Highest accuracy of CNN algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Vijayalakshmi et al. [70]	99.99%	CNN	Swarm intelligence	KDD Cup 99	Feature selection	Binary and multi-class classification	The established method still has several limitations, including AQO, which may be addressed with further research. Moreover, the IDS system to be employed in the IoT environment will incorporate various swarm intelligence approaches and DL designs.
2.	Fatani et al. [11]	99.99%	CNN	PSO, WOA	NLS-KDD	Feature selection	Multi-class classification	The developed method still exhibits several shortcomings, including AQU.
3.	Prabhakaran and Kulandasamy [67]	99.98%	CNN	CMBA	NLS-KDD	Feature selection	Binary and multi-class classification	The findings show that, for varying file sizes, the suggested approach marginally lengthens both the encryption and decryption times.
4.	Om Kumar et al. [71]	99.95%	CNN	MMBO	CICIDS-2017	Feature selection	Multi-class classification	The performance measurements are quite low because of the constraints of the current methodologies, which include poor performance and time complexity.
5.	Chen et al. [22]	99.84%	CNN	MOEA/D	AWID and CIC-IDS2107	parameters	Binary and multi-class classification	When there is a significant imbalance in the training data, MECNN performs marginally worse.

Table 9: Highest accuracy of RNN algorithm studies

	Research Name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Al-qaness et al. [11]	99.997%	RNN, ANN	Swarm intelligence	CIC2017	Feature selection	Binary and multi-class classification	The developed method exhibits high time complexity, particularly during feature extraction and the selection of relevant features. Additionally, the slow convergence of the CSA results in reduced solution quality.
2.	Murugesh and Murugan [24]	99.72%	RNN	*MSODL-ID*	Kaggle	Feature selection	Binary and multi-class classification	NN training becomes hard due to the requirement of a low learning rate for the internal covariance migration event.
3.	Al Sawafi et al. [72]	99%	RNN	Adam, Adamax	TON-IoT	Feature selection	Binary and multi-class classification	The suggested approach still exhibits several drawbacks that should be addressed in future studies. For instance, the suggested system heavily relies on the ToN-IoT dataset, which may not comprehensively represent the diverse threats encountered in real-world scenarios.
4.	Lateef et al. [73]	98.34%	RNN	CSO	KDD Cup 99	Feature selection	Binary and multi-class classification	To reach high-quality results, the algorithm CSO takes longer time
5.	Keserwani et al. [74]	98.11%	RNN	Genetic	UNSW-NB15	Feature selection	Binary and multi-class classification	In a network this size, overfitting lowers the effectiveness of the classification methods.

4.2.1 CNN

The supervised learning algorithm CNN [47] is built upon the foundation of conventional artificial neural networks. CNN excels in strong feature extraction and efficiently analyzes high-dimensional data using shared convolutional kernels. While multilayer FNN has some drawbacks such as slow learning rates and susceptibility to overfitting, leveraging CNN features like local field perception, weight sharing, and pooling can enhance learning, expression, and neural network performance. Local field perception significantly reduces the number of weight parameters required for training, while weight sharing further reduces the training parameters. Additionally, pooling layers result in smaller-sized and dimension features.

Many researchers employed the CNN algorithm. Table 8 presents the top five authors globally, each achieving the highest accuracy using CNN classification and optimization algorithms. Vijayalakshmi et al. [70] achieved a perfect accuracy of 99.99% followed by Fatani et al. [11] at 99.99%, Prabhakaran and Kulandasamy [67] at 99.98%, and Om Kumar et al. [71] at 99.95%. Chen et al. [22] achieved an accuracy of 99.84%.

4.2.2 RNN

Feedforward neural networks, the predecessors of RNNs [75], possess internal memory to handle input sequences of varying length. An RNN typically comprises an input layer, an output layer, and multiple hidden layers, often referred to as memory units. Each hidden layer depends on the output of preceding input layers and its current input to identify patterns in data. RNNs have found applications in various domains, including speech recognition, handwriting identification, sentiment analysis, and human activity recognition. They excel in handling sequential data due to their effectiveness with contextual information. RNNs have also been employed in intrusion detection and classification; however, they often face vanishing gradient problems

1. LSTM, a variant of RNN, was introduced to address the vanishing gradient issue. It consists of an input gate, an output gate, and a forget gate, allowing it to manage both single and series of input data. LSTMs find applications in areas such as speech recognition, handwriting recognition, and intrusion detection [69].
2. Another RNN variation, GRU, utilizes a gating mechanism to handle sequential data. Unlike LSTM, GRU incorporates two gates, an update gate and a reset gate. Update gates capture long-term dependencies in input sequences, while reset gates focus on short-term dependencies. GRU is suitable for handling input sequences with substantial time steps and is applied in domains like signal processing, music modeling, and natural language processing [69].

Table 9 presents the top five authors globally, each achieving the highest accuracy using RNN classification and optimization algorithms. Al-qaness et al. [11] achieved the highest accuracy at 99.997%, followed by Murugesh and Murugan [24] at 99.72%, Khan [72] at 99%, and Lateef et al. [73] at 98.34%, Keserwani et al. [74] at 98.11%.

4.3 Optimization (OP) algorithms

OP algorithms [76] is often the most efficient and accurate method for solving problems, although its definition can vary by context. In mathematics, it involves exploring the behavior of a problem by adjusting values within a specified range to either minimize or maximize a function. Optimization processes hold a significant role in DL, where various optimization functions are employed to minimize or maximize error functions. These functions have been developed in diverse environments. Figure 10 presents the most important OP algorithms used in research.

OP algorithms have been extensively integrated into intrusion detection research, with several studies focusing on OP algorithms for feature selection and DL algorithms based on intrusion detection. The review identified GA, ACO, and GWO as significant OP algorithms, consistently delivering high results across Tables 10, 12, and 14 . For instance, Fang published five papers between 2018 and 2022, focusing on OP algorithms for

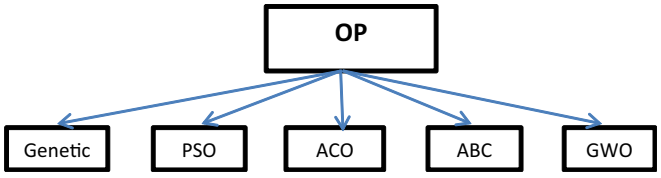

Figure 10: Optimization algorithms.

feature selection of network intrusion detection. The highest accuracy he reached is 97.89 using WOA and OPS optimization. Similarly, Zivkovic published five papers between 2018 and 2023, focusing on intrusion detection for OP algorithms and feature selection. The highest he reached is 99.6878 using GOA, MPO optimization. The integration of OP algorithms has contributed significantly to the overall performance of IDS. For instance, the KDD Cup 1999 dataset, which is one of the most widely utilized in the field of intrusion detection, has been used in comparison to the PSO algorithm, with Table 11 indicating a marginal difference of 0.1. Overall, OP algorithms have been extensively integrated into intrusion detection research, with GA, ACO, and GWO emerging as significant OP algorithms. These algorithms have contributed significantly to the overall performance of IDS, enhancing their reliability and effectiveness.

4.3.1 GAs

One of the most commonly used evolutionary metaheuristic algorithms for IDS design in the literature is GAs [22]. Hogue utilized GAs to develop an IDS capable of effectively identifying various types of network intrusions, and his work has been published. This strategy incorporates an evolutionary information evolution mechanism for processing traffic data. The KDD Cup 99 standard dataset served as the foundation for developing and evaluating this IDS, with the results demonstrating a reasonable detection rate. To provide a comprehensive perspective, this IDS was compared with numerous other techniques. In a similar vein, a piece of work based on GA fuzzy-class association mining was presented by Dwivedi et al. [47]. Many of the rules essential for creating an intrusion detection model are generated using GAs. Instead of generating every possible rule that satisfies the criteria for misuse detection, an association rule mining technique is employed to identify a sufficient number of key rules aligned with the user's goals. In an experimental study using the KDD Cup 99 intrusion detection dataset, Ibrahim Hayat Hassan proposed a method that exhibited a higher detection rate compared to traditional data mining approaches.

Many researchers applied the GA to enhance the performance and achieve higher accuracy. Table 10 showcases the top five authors in the field, each achieving the highest accuracy using classification and GA. Notably, Duo et al. [68] reached a remarkable accuracy of 100%, while Aljanabi and Ismail [36] also achieved a 100% accuracy rate. Additionally, Gorzałczany and Rudzinski [77] reached a perfect accuracy of 100%, and Injadat et al. [49] achieved an accuracy rate of 99.99%. Finally, Mahmood et al. [52] reached an accuracy rate of 99.36%.

4.3.2 PSO

Lavanya and Kannan introduced PSO [46], a technique inspired by the behavior of birds in a flock, which guides particles to explore the optimal global solution. PSO is generally easier to implement than GA due to the absence of evolutionary operators.

The authors employed a PSO algorithm. Table 11 highlights the top five authors globally, each achieving the highest accuracy using PSO for classification and to enhance their work's performance and achieve high accuracy. Notably, Injadat et al. [49] stands out with an accuracy of 99.99%, followed closely by Gaber et al. [60]

Table 10: Highest accuracy of GA algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Duo et al. [68]	100%	CART	Genetic	KDD 99	Hyper-parameter	Multi-class classification	Despite its strong performance on the training dataset, it encountered challenges when applied to the test dataset. This discrepancy suggests that the IDS model may not be suitable for the specific scenarios considered in this article. In the context of real-time Ethernet scenarios on a train, the model must possess the capability to effectively recognize anomalous data. Additionally, in the case of the CART model, having either too many or too few nodes can significantly impact the accuracy of the DT's classification.
2.	Aljanabi and Ismail [36]	100%	SVM	GA, TLBO	KDD Cup 99	Feature selection	Binary and multi-class classification	Increasing these parameters will result in a more precise outcome, albeit at the cost of longer computation times. Researching a new population can be a time-consuming process.
3.	Gorzałczany and Rudzinski [77]	100%	FRBC	Genetic	MQTT-IOT-IDS2020	Feature selection	Binary and multi-class classification	It took a long time to reach the highest accuracy.
4.	Injadat et al. [49]	99.99	DT	PSO AND Genetic	CICIDS 2017	Optimal parameter	Multi-class classification	The model faces challenges in detecting attack patterns and behaviors.
5.	Mahmood et al. [52]	99.36%	DT, SVM, KNN	PSO and Genetic	NSL-KDD	Feature selection	Multi-class classification	The experimental results suggest that reducing the number of features to a minimum, even if they are carefully chosen and relevant, does not always lead to higher accuracy. Instead, it is essential to select the right quantity of important and relevant features, which may even be a large number, to enhance the performance of ML models.

at 99.99%. Fatani et al. [11] reached a commendable 99.99% accuracy, while Al-qaness et al. [12] achieved 99.99% accuracy. Mohi-ud-din et al. [27] achieved an accuracy of 99.95%.

4.3.3 ACO

This algorithm is inspired by the real-world behavior of ants [51], which seek the shortest route between their colony and food sources, and ACO has been developed. ACO emulates the way ants communicate through pheromones within their population to discover the most optimal search space solution. It has been effectively employed to tackle discrete optimization challenges. ACO also offers an intriguing approach to feature selection for IDS, although its current application is somewhat limited.

Many researchers utilized the ACO algorithm to enhance the performance and achieve high accuracy in their work. Table 12 presents the top five authors globally, each achieving the highest accuracy using ACO for classification and OP algorithms. Notably, Alqarni et al. [45] attained a remarkable accuracy of 100%, followed closely by Mousavi et al. [50] at 99.92%. Samriya et al. [21] achieved an accuracy of 99.5%, while Bangui and Buhnova [62] reached 95.6% accuracy. Thakkar and Lohiya [69] reached 90.6% accuracy.

4.3.4 ABC

The inspiration for the ABC algorithm stems from the foraging behavior of bees [78]. Among the available solutions, ABC aims to locate the optimal one. The beehive consists of three types: scout bees, employed bees, and observer bees. These bees collaborate in various tasks, such as work distribution, food source selection, reproduction, scouting for the best food sources, and performing waggle dances to communicate the location of the optimal food sources. Initially, food sources are selected from the available options within the population. Employed bees then undertake random searches to discover superior food sources compared to those initially assigned to them.

Many researchers utilized the ABC algorithm to enhance the performance and achieve high accuracy. Table 13 shows the top five authors globally, each achieving the highest accuracy using ABC for classification and OP algorithms. Notably, Bacanin et al. [18] achieved an impressive accuracy of 99.65%, while Soni et al. [78] reached 97.42% accuracy. Mahboob et al. [79] achieved an accuracy of 97.23%, followed by Kalaivani et al. [80] with 97% accuracy, and Thakkar and Lohiya [69] reached 90.6% accuracy.

4.3.5 GWO

ML models often utilize meta-heuristic algorithms inspired by nature [395950-34]. One such algorithm, GWO, was introduced by Mirjalili et al. in 2014. GWO draws inspiration from the social structure and clever hunting tactics of grey wolves. In the natural world, grey wolves typically travel in packs consisting of 5–12 individuals. The GWO algorithm emulates the hunting behavior and leadership structure of these wolves [80].

Many researchers like Swarna Priya employed the GWO algorithm to enhance their work's performance and achieve high accuracy. Table 14 showcases the top five authors globally, each achieving the highest accuracy using GWO for classification and OP algorithms. Notably, ElDahshan et al. [53] attained a remarkable accuracy of 100%, while Alzubi et al. [33] reached an accuracy of 99.22%. Davahli et al. [81] achieved an accuracy of 99.10% and Swarna Priya et al. [82] reached 99.9% accuracy. Kunhare et al. [83] achieved an accuracy of 97.894%.

Many other researchers employed a range of OP algorithms in their research, including Firefly, Harris Hawk, Multi-verse optimizer, Whale OP algorithm, Cuckoo search, Bat algorithm, AOA, and more [84–91]. The systematic review provides a structured synthesis of the state-of-the-art in intrusion detection research by consolidating and analyzing a comprehensive set of 393 papers meeting the inclusion criteria. Through

Table 11: Highest accuracy of PSO algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Injadat et al. [49]	99.99	DT	PSO AND Genetic	CICIDS 2017	Optimal parameter	Multi-class classification	Both datasets are initially imbalanced, containing significantly fewer attack samples than standard samples. Consequently, the model encounters challenges in identifying attack patterns and behaviors.
2.	Gaber et al. [60]	99.99%	RF	PSO and Bat	WUSTL-IIOT-2021	Feature selection	Multi-class classification	The IoT system's inability to collect sufficient real-world data for evaluating existing solutions. The unbalanced dataset used in developing ML-based IDS, potentially resulting in the failure to detect minority attacks. The study solely assessed the performance of the suggested feature selection technique with three ML algorithms (RF, k-NN, and MLP). The study did not consider the potential impact of different attack types on the effectiveness of the proposed intrusion detection method.
3.	Fatani et al. [11]	99.99%	CNN	PSO, WOA	NLS-KDD	Feature selection	Multi-class classification	The developed method still exhibits several shortcomings, including AQU.
4.	Al-qaness et al. [12]	99.99%	RNN, ANN	Swarm intelligence	CIC2017	Feature selection	Binary and multi class classification	The developed method exhibits high time complexity, particularly during feature extraction and the selection of relevant features. The slow convergence of the CSA results in diminished solution quality.
5.	Mohi-ud-din et al. [27]	99.95%	RF	CSA-PSO	UNSW-NB15	Feature selection	Multi-class classification	To reach high-quality results, the algorithm CSA takes a long time

Table 12: Highest accuracy of ACO algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Alqarni et al. [45]	100%	SVM	ACO	KDD Cup 99	Feature selection	Binary and multi-class classification	It took a long time to reach the highest accuracy
2.	Mousavi et al. [50]	99.92%	DT	ACO	KDD Cup 99	Feature selection	Binary and multi-class classification	The method used in this research to select a small training subset for multiclass classification under imbalanced conditions is currently challenging and suboptimal. It requires greater flexibility, compatibility, and efficiency. Consequently, there is a need to develop a better and more suitable method to address this issue.
3.	Samriya et al. [65]	99.5%	NB	ACO	NLS-KDD	Feature selection	Binary and multi-class classification	Detecting anomalies in IoT networks and identifying malware in uncertain and overcast conditions can be time-consuming.
4.	Bangui and Buhnova [62]	95.6%	RF	ACO	CICIDS2017	Feature selection	Multi-class classification	It took a long time to analyze the comprehensive data to enhance its security against various attacks
5.	Thakkar and Lohiya [69]	90.6%	SVM	ACO, ABC	KDD Cup 99	Feature selection	Multi-class classification	The performance evaluation dataset exhibits class imbalance, necessitating the development of SWEVO-based methods.

Table 13: Highest accuracy of ABC algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	Bacanin et al. [18]	99.65%	XGBoost	ABC	UNSW-NB15	Feature selection	Multi class classification	One of the upcoming challenges in this domain is to validate the suggested hybrid model on additional intrusion detection datasets. This step is crucial for increasing confidence in the results before applying the model in real-world scenarios.
2.	Soni et al. [78]	97.42%	CNN	ABC	NLS-KDD	Feature selection	Binary and multi-class classification	The ABC algorithm took a long time to reach the required results and this may affect its performance
3.	Mahboob et al. [79]	97.23%	KNN, ANN	ABC	NLS-KDD	Feature selection	Binary and multi-class classification	The exploration of the mentioned properties may result in increased processing time and additional hardware overhead. However, it is possible to convert symbolic features into numerical ones.
4.	Kalaivani et al. [80]	97%	ANN	ABC	CICIDS2017	Feature selection	Multi-class classification	The detection algorithm's error rate should be as low as possible given the appropriate limit setting. An ideal limit value was not found for the network's unidentified distribution model.
5.	Thakkar and Lohiya [69]	90.6%	SVM	ABC, ACO	KDD Cup 99	Feature selection	Multi-class classification	The performance evaluation dataset exhibits class imbalance, necessitating the development of SWEVO-based methods to address this issue.

Table 14: Highest accuracy of GWO algorithm studies

	Research name	Highest accuracy	Classification algorithm	Optimization algorithm	Dataset	Optimization field	Classification type	Limitation of the study
1.	ElDahshan et al. [53]	100	ELM	GWO	CICIDS 2017	Parameter	Binary and multi-class classification	Resolving these problems should prioritize a focus on attack instances over normal instances, as misclassifying attacks among attack instances can cause more significant harm than misclassifying attacks among normal instances. (1) The need for accurate detection of various attack instances. (2) The challenge of reducing false alarm rates. (3) The difficulty in efficiently identifying different types of attacks. (4) The issue of dealing with large amounts of data and class imbalance in datasets. (5) The challenge of selecting relevant features for ML-based IDS. (6) The need for effective hyperparameter optimization methods for ML models. (7) The challenge of achieving high detection rates with a small amount of data. (8) The issue of dealing with constantly evolving attack methods and techniques
2.	Alzubi et al. [33]	99.22	SVM	GWO	KDD Cup99	Feature selection	Binary and multi-class classification	MBGWO and bGWO convergence. Due to its limited ability to identify a small number of FS with numerous aims, the bGWO has always been inadequate in addressingideal solutions in a single run, which entails making several runs to achieve a predetermined number of features.
a.	Davahli et al. [81]	99.10	SVM	GWO	AWID	Feature selection	Binary and multi-class classification	The computational costs, such as the time and memory needed for intrusion detection, are very important in wireless networks due to resource limitations. Some initiatives or attempts should be made to further shorten the GA-GWO runtime, such as parallelizing tasks.
4.	Swarna Priya et al. [82]	99.9	SVM	GWO	KDD Cup 99	Feature selection	Multi-class classification	Expanded the amount of data that have to be categorized and examined. High impact features should be chosen, while undesirable elements should be removed.
5.	Kunhare et al. [83]	97.894	DT	GWO	NSL-KDD	Feature selection	Multi-class classification	The proposed work has certain limitations. The stochastic nature of GA results in longer convergence times. Optimization based on biological evolution can be computationally demanding. GWO exhibits a low convergence rate, limited precision in solving, and a limited ability for local searching.

bibliometric analysis and categorization, the review offers key insights and overarching themes derived from the analysis of the reviewed literature:

1. Publication trends: The review identifies increasing publication volumes in the field of intrusion detection, indicating a growing interest and research activity in this domain.
2. Frequently adopted algorithms: The review highlights the dominance of specific ML and DL algorithms, such as SVM, CNN, DTs, and GA, as leading techniques for intrusion detection.
3. Utilized datasets: The review emphasizes the significance of benchmark datasets, including KDD Cup 1999, NSL-KDD, UNSW-NB15, and CICIDS2017, as commonly used resources for evaluating intrusion detection models.
4. Challenges and limitations: The review identifies challenges and limitations, such as limited availability of labeled datasets, imbalanced datasets, adversarial attacks, interpretability, explainability, and scalability, which influence the overall effectiveness and reliability of IDS.
5. Future research directions: The review suggests future research directions, including the exploration of DL methods, addressing computational complexity, enhancing model interpretability, and evaluating diverse new datasets.

By synthesizing these key insights and overarching themes, the review provides a comprehensive overview of the current state-of-the-art in intrusion detection research. It offers valuable guidance for researchers and practitioners, enabling them to understand the prominent trends, challenges, and potential areas for further investigation in the field of IDS.

4.4 Dataset

The datasets encompass fields containing both unprocessed and processed data extracted from underlying network traffic [90]. These data are typically generated through studies aimed at identifying network intrusions. An intentional effort is made to manipulate the data, creating adversarial examples capable of deceiving classifiers and detection systems. When creating adversarial instances that alter the source data in network security applications, caution is essential, as highlighted by [90]. The most prominent datasets utilized in research include KDD Cup99, NSL-KDD, CICIDS 2017, UNSW-NB15, AWID, Kaggle, and TON-IOT. Table 15 provides an overview of the most crucial datasets commonly used in the field of intrusion detection [37–39,91–97].

The most commonly used datasets in IDS are as follows:

1. KDD Cup 1999: This dataset stands as one of the most widely utilized in the field of intrusion detection. It comprises a substantial and diverse collection of network traffic data, encompassing potential attacks.
2. NSL-KDD: An enhanced iteration of the KDD Cup 1999 dataset, it offers a more demanding and realistic environment for testing intrusion detection models.
3. UNSW-NB15: This dataset consists of samples of network traffic extracted from real-world internet environments, providing a formidable challenge for detecting advanced attacks.
4. CICIDS2017: This dataset encompasses a diverse range of data reflecting different attack scenarios, serving as an invaluable resource for evaluating the performance of intrusion detection models.

The reviewed studies have frequently utilized several benchmark datasets for evaluating IDS. The most commonly used datasets include: (1) KDD Cup 1999: This dataset is one of the most widely utilized in the field of intrusion detection. It comprises a substantial and diverse collection of network traffic data, encompassing potential attacks. (2) NSL-KDD: An enhanced iteration of the KDD Cup 1999 dataset, it offers a more demanding and realistic environment for testing intrusion detection models. (3) UNSW-NB15: This dataset consists of samples of network traffic extracted from real-world internet environments, providing a formidable challenge for detecting advanced attacks. (4) CICIDS2017: This dataset encompasses a diverse range of data reflecting different attack scenarios, serving as an invaluable resource for evaluating the performance of intrusion

Table 15: Most commonly used datasets

Dataset	Dataset description	Dataset size	Link of the dataset	Public or private	Limitation
KDD Cup 99 Om Kumar et al. [71]	The KDD Cup 99 dataset was developed by MIT Lincoln Laboratories and encompasses four main attack categories: Denial of Service , Remote to Local, User to Root, and probing. The dataset comprises 41 features and one label, which provides information about the type of attack.	4,900,000 records 41 attributes	https://kdd.ics.uci.edu/datab ases/kddcu p99/kddcu p99.html	Public	While the KDD Cup 99 dataset remains the most popular and widely used public dataset for IDS, it requires data cleaning or preprocessing due to roughly 78% redundant records. Record duplication can introduce bias towards frequent chromosomes, reducing the effectiveness of network intrusion detection.
NSL-KDD Li et al. [89]	The NSL-KDD dataset is an upgrade of the KDD Cup 99 dataset, designed to address some of its underlying issues.	125,973 records 41 attributes	http://scikit-learn.org/stable/modules/preprocessing.html#encoding-categorical-features	Public	Due to the limited availability of publicly accessible datasets for network-based IDS, this updated version of the KDD dataset still has some issues and may not fully represent the current real networks.
UNSW-NB15 Li et al. [89]	The UNSW-NB15 dataset combines raw network packets to include a mixture of real, current normal activity and synthetic attacks	2,450,044 records 49 attributes		Public	Before using this dataset for model development, it is essential to address its two main issues: class imbalance and class overlap. Failure to resolve these problems could potentially hinder IDS in identifying and detecting attacks.
CICIDS 2017 Yousef [75]	The CIC-IDS-2017 dataset, created by the Faculty of Computer Science, encompasses both regular and various attack data within network traffic. The dataset exhibits class imbalance, with an uneven distribution between the dominant and minority classes in the database.	25,00,000 records 78 attributes	https://www.unb.ca/cic/datasets/ids-2017.html	Public	The CICIDS 2017 dataset has a few shortcomings and weaknesses: Large volume of data. Missing values.
AWID Davahli et al. [81]	The AWID dataset is a newly developed Wi-Fi network intrusion benchmark that is useful for evaluating IDSs employed by network IDS research communities. This dataset is particularly relevant for research involving IoT wireless networks connected to Wi-Fi networks.	162,385 records 154 features	https://icsdweb.aegean.gr/awid/	Private	The AWID intrusion dataset encompasses various data types, including discrete, continuous, and symbolic (nominal) data with a wide range of values. These data variations pose a challenge for classifiers, even for high-performing classifiers like SVM, to effectively train on normal and abnormal patterns.

detection models. The utilization of these benchmark datasets has influenced the comparability of research outcomes by providing a standardized basis for evaluating the performance of IDS. Researchers can compare the effectiveness of different algorithms and techniques using these commonly accepted benchmark datasets, thereby facilitating the assessment of the reliability and generalizability of intrusion detection models.

5 Discussion

This section delves into the findings derived from comparing various ML algorithms. Notably, Tables 2–4 indicate that ML models exhibit superior performance when using SVM, DT, and ELM algorithms, particularly with the KDD Cup 1999 and NF-Bot datasets, both for multi-class and binary classification, as assessed by accuracy. In contrast, Table 5 presents slightly lower results when employing the XGBoost algorithm, and the CICIDS2017 dataset is utilized for multi-class classification. Tables 6 and 7 reveal results nearly identical to Table 5 with the RF and NB algorithms. Therefore, SVM, DT, and ELM algorithms outperform RF, NB, and XGBoost, though the margin is relatively small, typically within the range of 0.1–0.2 in terms of accuracy. In the realm of DL algorithms, Table 8 showcases improved outcomes with the CNN algorithm and the NLS-KDD L dataset compared to Table 9, which demonstrates lower results with the RNN algorithm and the CICIDS2017 dataset, once again, gauged by accuracy. Notably, GA, ACO, and GWO stand out as significant OP algorithms, consistently delivering high results across Tables 10 and 12 and and 14. For instance, the KDD CUP 99 and CICIDS2017 datasets are used in comparison to the PSO algorithm. While Table 11 indicates a marginal difference of 0.1, Table 13 reveals a discrepancy of approximately 0.34 when the ABC algorithm is utilized with the UNSW-NB15 dataset. The KDD CUP 99 dataset emerges as the most frequently employed in conjunction with ML and DL algorithms, signifying that these algorithms exhibit great potential for enhancing intrusion detection in both binary and multi-class classification scenarios.

6 Conclusion

This systematic review presents a structured synthesis of research on ML and DL techniques for intrusion detection published over the past 5 years. An analysis of 393 studies reveals a noticeable increase in publication volumes, indicating a growing interest in this field. The mapping of frequently used algorithms highlights SVM, CNN, DTs, and GA as dominant techniques. The most commonly used public datasets include KDD Cup 1999, NSL-KDD, CICIDS2017, and UNSW-NB15. The review methodology integrates findings from multiple studies to provide a holistic overview of the current state-of-the-art. The results can inform future research by identifying promising techniques and gaps for further investigation. For instance, DL methods show potential but require ongoing exploration. Aspects such as computational complexity, model interpretability, and evaluation on diverse new datasets require further attention. Overall, this review provides a valuable reference that captures the current landscape of intelligent intrusion detection techniques and datasets, helping researchers position their work in this evolving research domain and select appropriate methodologies for comparative evaluation. The conclusion of the systematic literature review on IDS presents key findings, insights, and implications derived from the research, emphasizing the significance of the study's outcomes in the broader context of the research area. The conclusion highlights the following key results, insights, and implications: (1) Key Results: Increasing publication trends: The review reveals increasing publication trends in the research domain of IDS, indicating the growing interest and significance of the field. Frequently used algorithms: The study identifies CNN, SVM, DTs, and GA as the top methods frequently used in intrusion detection research, providing insights into the prevalent algorithms in the field. - Commonly used datasets: The review identifies widely utilized datasets such as KDD Cup 1999, NSL-KDD, UNSW-NB15, and CICIDS 2017, emphasizing the importance of diverse and realistic datasets for evaluating intrusion detection models. (2) Insights: ML and DL Techniques: The review underscores the significant potential of ML and DL

techniques in the development of effective IDS, highlighting their transformative impact on IDS in terms of security, adaptability, and scalability. Challenges and limitations: The study discusses the challenges and limitations of current techniques in intrusion detection, providing a structured synthesis of the state-of-the-art to guide future research in the field, thus offering valuable insights for researchers and practitioners. (3) Implications: Future research directions: The findings of the review have implications for guiding future intrusion detection research, particularly in the selection of algorithms, utilization of datasets, and addressing the challenges and limitations identified in the study. Advancements in cybersecurity: The study's outcomes have broader implications for advancing the field of cybersecurity by providing insights into the utilization of ML, DL, OP algorithms, and datasets in intrusion detection research, thus contributing to the enhancement of security measures against network intrusions. In summary, the systematic literature review provides valuable insights into the publication trends, frequently used algorithms, commonly utilized datasets, challenges, and implications for future research in the field of IDS. The study's outcomes have significant implications for advancing the field of cybersecurity and guiding future research endeavors in intrusion detection.

Based on the findings of the systematic review, several areas and methodologies warrant further exploration and improvement in future intrusion detection research. The review provides guidance for future research in the following areas: (1) DL methods: The review suggests ongoing exploration of DL methods for intrusion detection, indicating their potential for enhancing detection capabilities. Future research could focus on leveraging advanced DL architectures and techniques to improve the accuracy and robustness of IDS. (2) Computational complexity: Addressing the computational complexity of intrusion detection models is highlighted as an area for improvement. Future research could explore methods to optimize the computational efficiency of ML and DL algorithms, particularly in large-scale network environments. (3) Model Interpretability: Enhancing the interpretability of intrusion detection models is identified as a crucial area for improvement. Future research could focus on developing methods to improve the transparency and explainability of ML and DL models, enabling better understanding of their decision-making processes. 4. Evaluation on diverse new datasets: The review emphasizes the importance of evaluating intrusion detection models on diverse new datasets. Future research could involve the creation and utilization of novel datasets that capture a wide range of network traffic scenarios, including emerging threats and attack patterns. (5) Addressing adversarial attacks: Given the challenge of adversarial attacks, future research could focus on developing robust intrusion detection techniques that are resilient to adversarial manipulation of data. (6) Scalability: Addressing the scalability of IDS is highlighted as an important area for improvement. Future research could explore methods to ensure the effective deployment of IDS in large-scale network environments. Overall, the review guides future intrusion detection research by identifying promising areas for exploration and improvement, including the continued investigation of DL methods, addressing computational complexity, enhancing model interpretability, evaluating on diverse new datasets, addressing adversarial attacks, and ensuring scalability in real-world.

Acknowledgments: The corresponding auhtor would like to thank Imam Ja'afar Al-Sadiq University for their support.

Funding information: No funding received for this paper.

Author contributions: Melad: collect the data, analysis, and write the results; Mohammad: methodology design, Rstudio results, and write the introduction and discussion section; Hassan: interpretation of results, and conclusion.

Conflict of interest: Authors declare no conflict of interest.

Data availability statement: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] Yaseen MG, Aljanabi M. Recent advances in control theory for complex systems. Babylon J Math. 2023;2023:7–11.

[2] Gopi RS, Dhanesh L, Aljanabi M, Rao TV, Thiruveni M, Mahalakshmi S. Design of Covid19 disease detection for risk identification using deep learning approach. J Adv Res Appl Sci Eng Technol. 2023;32(1):139–54.

[3] Aljanabi M, Mohammed SY. Metaverse: Open possibilities. Iraqi J Computer Sci Math. 2023;4(3):79–86.

[4] Hilal AM, Al-Otaibi S, Mahgoub H, Al-Wesabi FN, Aldehim G, Motwakel A, et al. Deep learning enabled class imbalance with sand piper optimization based intrusion detection for secure cyber physical systems. Clust Comput. 2023;26(3):2085–98. doi: 10.1007/s10586-022-03628-w.

[5] A. Alissa K, S. Alrayes F, Tarmissi K, Yafoz A, Alsini R, Alghushairy O, et al. Planet optimization with deep convolutional neural network for lightweight intrusion detection in resource-constrained IoT networks. Appl Sci (Switz). 2022;12(17):1–15. doi: 10.3390/app12178676.

[6] Mohamed HG, Alotaibi SS, Eltahir MM, Mohsen H, Ahmed Hamza M, Sarwar Zamani A, et al. Feature selection with stacked autoencoder based intrusion detection in drones environment. Computers Mater Continua. 2022;73(3):5441–58. doi: 10.32604/cmc.2022.031887.

[7] Alissa KA, Alotaibi SS, Alrayes FS, Aljebreen M, Alazwari S, Alshahrani H, et al. Crystal structure optimization with deep-autoencoder-based intrusion detection for secure internet of drones environment. Drones. 2022;6(10):297. doi: 10.3390/drones6100297.

[8] Mohamed HG, Alrowais F, Al-Hagery MA, Al Duhayyim M, Hilal AM, Motwakel A. Optimal wavelet neural network-based intrusion detection in internet of things environment. Computers Mater Continua. 2023;75(2):4467–83. doi: 10.32604/cmc.2023.036822.

[9] Ahmed Hamza M, Hassan Abdalla Hashim A, Mohamed HG, Alotaibi SS, Mahgoub H, Mehanna AS, et al. Hyperparameter tuned deep learning enabled intrusion detection on Internet of Everything environment. Computers Mater Continua. 2022;73(3):6579–94. doi: 10.32604/cmc.2022.031303.

[10] Duhayyim MA, Alissa KA, Alrayes FS, Alotaibi SS, Tag El Din EM, Abdelmageed AA, et al. Evolutionary-based deep stacked auto-encoder for intrusion detection in a cloud-based cyber-physical system. Appl Sci (Switz). 2022;12(14):6875. doi: 10.3390/app12146875.

[11] Fatani A, Dahou A, Al-Qaness MAA, Lu S, Elaziz MA. Advanced feature extraction and selection approach using deep learning and aquila optimizer for IoT intrusion detection system. Sensors. 2022;22(1):140. doi: 10.3390/s22010140.

[12] Abd Elaziz M, Al-qaness MAA, Dahou A, Ibrahim RA, El-Latif AAA. Intrusion detection approach for cloud and IoT environments using deep learning and Capuchin Search Algorithm. Adv Eng Softw. 2023;176(September 2022):103402. doi: 10.1016/j.advengsoft.2022.103402.

[13] Dahou A, M AbdElaziz, Chelloug SA, Awadallah MA, Al-Betar MA, Al-Qaness M, et al. Intrusion detection system for IoT based on deep learning and modified reptile search algorithm. Comput Intell Neurosci. 2022;2022:1–15. doi: 10.1155/2022/6473507.

[14] Fatani A, Elaziz MA, Dahou A, Al-Qaness MAA, Lu S. IoT intrusion detection system using deep learning and enhanced transient search optimization. IEEE Access. 2021;9:123448–64. doi: 10.1109/ACCESS.2021.3109081.

[15] Fatani A, Dahou A, M AbdElaziz, Al-Qaness M, Lu S, Alfadhli SA, et al. Enhancing intrusion detection systems for IoT and cloud environments using a growth optimizer algorithm and conventional neural networks. Sensors. 2023;23(9):1–14. doi: 10.3390/s23094430.

[16] Stankovic M, Zivkovic M, Antonijevic M, Tanaskovic M, Bacanin N, Jovanovic D. Feature selection by hybrid artificial bee colony algorithm for intrusion detection. International Conference on Edge Computing and Applications, ICECAA 2022 – Proceedings, no. Icecaa; 2022. p. 500–5. doi: 10.1109/ICECAA55415.2022.9936116.

[17] Zivkovic M, Tair M, Venkatachalam K, Bacanin N, Hubálovský Š, Trojovský P. Novel hybrid firefly algorithm: An application to enhance XGBoost tuning for intrusion detection classification. PeerJ Comput Sci. 2022;8:1–38. doi: 10.7717/peerj-cs.956.

[18] Bacanin N, Petrovic A, Antonijevic M, Zivkovic M, Sarac M, Tuba E, et al. Intrusion detection by XGBoost model tuned by improved social network search algorithm. In International Conference on Modelling and Development of Intelligent Systems. Cham: Springer Nature Switzerland; 2022. p. 104–21.

[19] Jovanovic D, Marjanovic M, Antonijevic M, Zivkovic M, Budimirovic N, Bacanin N. Feature selection by improved sand cat swarm optimizer for intrusion detection. Proceedings - 2022 International Conference on Artificial Intelligence in Everything, AIE 2022; 2022. p. 685–90. doi: 10.1109/AIE57029.2022.00134.

[20] Jovanovic L, Jovanovic D, Antonijevic M, Zivkovic M, Budimirovic N, Strumberger I, et al. The XGBoost tuning by improved firefly algorithm for network intrusion detection. In 2022 24th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC). IEEE; 2022. p. 268–75.

[21] Chen Y, Lin Q, Wei W, Ji J, Wong KC, Coello CAC. Intrusion detection using multi-objective evolutionary convolutional neural network for Internet of Things in Fog computing. Knowl Based Syst. 2022;244:108505. doi: 10.1016/j.knosys.2022.108505.

[22] Chen P, You C, Ding P. Event classification using improved salp swarm algorithm based probabilistic neural network in fiber-optic perimeter intrusion detection system. Optical Fiber Technol. 2020;56(September 2019):102182. doi: 10.1016/j.yofte.2020.102182.

[23] Xu H, Przystupa K, Fang C, Marciniak A, Kochan O, Beshley M. A combination strategy of feature selection based on an integrated optimization algorithm and weighted k-nearest neighbor to improve the performance of network intrusion detection. Electronics (Switzerland). 2020;9(8):1–22. doi: 10.3390/electronics9081206.

[24] Murugesh C, Murugan S. Moth search optimizer with deep learning enabled intrusion detection system in wireless sensor networks. SSRG Int J Electr Electron Eng. 2023;10(4):77–90. doi: 10.14445/23488379/IJEEE-V10I4P108.
[25] Chaudhary DK, Yadav P, Gupta S, Jha K. IOT network feature based intrusion detection techniques - Review. Proceedings of 2022 IEEE International Conference on Current Development in Engineering and Technology, CCET 2022; 2022. p. 1–5. doi: 10.1109/CCET56606.2022.10080392.
[26] Pathania A. A hybrid approach for intrusion detection system using data minining and artificial neural network. 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). 2018, 2021. p. 1508–14. doi: 10.1109/ICAC3N53548.2021.9725482.
[27] Mohi-ud-din G, Zhiqiang L, Jiangbin Z, Sifei W, Zhijun L, Asim M, et al. Intrusion detection using hybrid enhanced CSA-PSO and multivariate WLS random-forest technique. IEEE Trans Netw Serv Manag. 2023;20:1. doi: 10.1109/tnsm.2023.3258901.
[28] Almuqren L, Al-Mutiri F, Maashi M, Mohsen H, Hilal AM, Alsaid MI, et al. Sine-cosine-adopted African vultures optimization with ensemble autoencoder-based intrusion detection for cybersecurity in CPS environment. Sensors. 2023;23(10):1–19. doi: 10.3390/s23104804.
[29] Alohali MA, Al-Wesabi FN, Hilal AM, Goel S, Gupta D, Khanna A. Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment. Cogn Neurodyn. 2022;16(5):1045–57. doi: 10.1007/s11571-022-09780-8.
[30] Alrowais F, Marzouk R, Nour MK, Mohsen H, Hilal AM, Yaseen I, et al. Intelligent intrusion detection using arithmetic optimization enabled density based clustering with deep learning. Electron (Switz). 2022;11(21):1–15. doi: 10.3390/electronics11213541.
[31] Kavitha S, Maheswari NU, Venkatesh R. Intelligent intrusion detection system using enhanced arithmetic optimization algorithm with deep learning model. Tehnicki Vjesn. 2023;30(4):1217–24. doi: 10.17559/TV-20221128071759.
[32] Dahou A, AbdElaziz M, Chelloug SA, Awadallah MA, Al-Betar MA, Al-Qaness M, et al. Intrusion detection system for IoT based on deep learning and modified reptile search algorithm. Comput Intell Neurosci. 2022;2022:1–15.
[33] Alzubi QM, Anbar M, Alqattan ZNM, Al-Betar MA, Abdullah R. Intrusion detection system based on a modified binary grey wolf optimisation. Neural Comput Appl. 2020;32(10):6125–37. doi: 10.1007/s00521-019-04103-1.
[34] Alawad NA, Abed-alguni BH, Al-Betar MA, Jaradat A. Binary improved white shark algorithm for intrusion detection systems. Neural Comput Appl. 2023;35(26):19427–51. doi: 10.1007/s00521-023-08772-x.
[35] Ramasamy M, Eric PV. A novel classification and clustering algorithm for intrusion detection system on convolutional neural network. Bull Electr Eng Inform. 2022;11(5):2845–55. doi: 10.11591/eei.v11i5.4145.
[36] Aljanabi M, Ismail M. Improved intrusion detection algorithm based on TLBO and GA algorithms. Int Arab J Inf Technol. 2021;18(2):170–9. doi: 10.34028/IAJIT/18/2/5.
[37] Aljanabi M, Ismail MA, Mezhuyev V. Improved TLBO-JAYA algorithm for subset feature selection and parameter optimisation in intrusion detection system. Complexity. 2020;2020:1–18. doi: 10.1155/2020/5287684.
[38] Alhayali RAI, Aljanabi M, Ali AH, Mohammed MA, Sutikno T. Optimized machine learning algorithm for intrusion detection. Indonesian J Electr Eng Computer Sci. 2021;24(1):590–9. doi: 10.11591/ijeecs.v24.i1.pp590-599.
[39] Aljanabi M, Ismail MA, Ali AH. Intrusion detection systems, issues, challenges, and needs. Int J Comput Intell Syst. 2021;14(1):560–71. doi: 10.2991/ijcis.d.210105.001.
[40] Mijwil MM, Aljanabi M. A comparative analysis of machine learning algorithms for classification of diabetes utilizing confusion matrix analysis. Baghdad Sci J. 2023.
[41] Aljanabi M. Safeguarding connected health: Leveraging trustworthy AI techniques to harden intrusion detection systems against data poisoning threats in IoMT environments. Babylon J Internet Things. 2023;2023:31–7.
[42] Aljanabi M. Navigating the landscape: A comprehensive bibliometric analysis of decision-making research in civil engineering. Mesopotamian J Civ Eng. 2023;2023:35.
[43] Omran AH, Mohammed SY, Aljanabi M. Detecting data poisoning attacks in federated learning for healthcare applications using deep learning. Iraqi J Computer Sci Math. 2023;4(4):225–37.
[44] Aljanabi M, Yaseen MG, Ali AH, Mohammed MA. Prompt engineering: Guiding the way to effective large language models. Iraqi J Computer Sci Math. 2023;4(4):151–5.
[45] Alqarni AA. Toward support-vector machine-based ant colony optimization algorithms for intrusion detection. Soft Comput. 2023;27(10):6297–305. doi: 10.1007/s00500-023-07906-6.
[46] Lavanya R, Kannan S. Intrusion detection system for energy efficient cluster based vehicular adhoc networks. Intell Autom Soft Comput. 2022;32(1):323–37. doi: 10.32604/iasc.2022.021467.
[47] Dwivedi S, Vardhan M, Tripathi S. Building an efficient intrusion detection system using grasshopper optimization algorithm for anomaly detection. Clust Comput. 2021;24(3):1881–900. doi: 10.1007/s10586-020-03229-5.
[48] Liu Z, Shi R, Lei M, Wu Y. Intrusion detection method based on improved sparrow algorithm and optimized SVM. Proceedings - 2022 4th International Conference on Data Intelligence and Security, ICDIS 2022; 2022. p. 27–30. doi: 10.1109/ICDIS55630.2022.00012.
[49] Injadat M, Moubayed A, Nassif AB, Shami A. Multi-stage optimized machine learning framework for network intrusion detection. IEEE Trans Netw Serv Manag. 2021;18(2):1803–16. doi: 10.1109/TNSM.2020.3014929.
[50] Mousavi SM, Majidnezhad V, Naghipour A. A new intelligent intrusion detector based on ensemble of decision trees. J Ambient Intell Humaniz Comput. 2022;13(7):3347–59. doi: 10.1007/s12652-019-01596-5.
[51] Maza S, Touahria M. Feature selection for intrusion detection using new multi-objective estimation of distribution algorithms. Appl Intell. 2019;49(12):4237–57. doi: 10.1007/s10489-019-01503-7.

[52] Mahmood RAR, Abdi AH, Hussin M. Performance evaluation of intrusion detection system using selected features and machine learning classifiers. Baghdad Sci J. 2021;18(2):884–98. doi: 10.21123/bsj.2021.18.2(Suppl.).0884.
[53] ElDahshan KA, AlHabshy AAA, Hameed BI. Meta-heuristic optimization algorithm-based hierarchical intrusion detection system. Computers. 2022;11(12):170. doi: 10.3390/computers11120170.
[54] Vaiyapuri T, Algamdi S, John R, Sbai Z, Al-Helal M, Alkhayyat A, et al. Metaheuristics with federated learning enabled intrusion detection system in Internet of Things environment. Expert Syst. 2022;40(July 2022):1–16. doi: 10.1111/exsy.13138.
[55] Ghasemi J, Esmaily J, Moradinezhad R. Intrusion detection system using an optimized kernel extreme learning machine and efficient features. Sadhana - Acad Proc Eng Sci. 2020;45(1):1–9. doi: 10.1007/s12046-019-1230-x.
[56] Wang C, Cai W, Ye Z, Yan L, Wu P, Wang Y. Network intrusion detection based on lightning search algorithm optimized extreme learning machine. 13th International Conference on Computer Science and Education, ICCSE 2018, no. Iccse; 2018. p. 562–6. doi: 10.1109/ICCSE.2018.8468727.
[57] Kunhare N, Tiwari R, Dhar J. Particle swarm optimization and feature selection for intrusion detection system. Sadhana - Acad Proc Eng Sci. 2020;45(1):1–14. doi: 10.1007/s12046-020-1308-5.
[58] Kilincer IF, Ertam F, Sengur A. A comprehensive intrusion detection framework using boosting algorithms. Computers Electr Eng. 2022;100(May 2021):107869. doi: 10.1016/j.compeleceng.2022.107869.
[59] Xu W, Fan Y. Intrusion detection systems based on logarithmic autoencoder and XGBoost. Secur Commun Netw. 2022;2022:1–8. doi: 10.1155/2022/9068724.
[60] Gaber T, Awotunde JB, Folorunso SO, Ajagbe SA, Eldesouky E. Industrial Internet of Things intrusion detection method using machine learning and optimization techniques. Wirel Commun Mob Comput. 2023;2023:1–15. doi: 10.1155/2023/3939895.
[61] Samawi VW, Yousif SA, Al-Saidi NMG. Intrusion detection system: An automatic machine learning algorithms using auto-WEKA. 2022 IEEE 13th Control and System Graduate Research Colloquium, ICSGRC 2022 - Conference Proceedings; 2022. July. p. 42–6. doi: 10.1109/ICSGRC55096.2022.9845166.
[62] Bangui H, Buhnova B. Lightweight intrusion detection for edge computing networks using deep forest and bio-inspired algorithms. Computers Electr Eng. 2022(July 2021);100:107901. doi: 10.1016/j.compeleceng.2022.107901.
[63] Shitharth S, Kshirsagar PR, Balachandran PK, Alyoubi KH, Khadidos AO. An innovative perceptual pigeon galvanized optimization (PPGO) based Likelihood Naïve Bayes (LNB) classification approach for network intrusion detection system. IEEE Access. 2022;10:46424–41. doi: 10.1109/ACCESS.2022.3171660.
[64] Devi TJ, Singh KJ. Anomaly-based intrusion detection system in two benchmark datasets using various learning algorithms. vol. 225, Singapore: Springer; 2021. doi: 10.1007/978-981-16-0878-0_19.
[65] Samriya JK, Tiwari R, Cheng X, Singh RK, Shankar A, Kumar M. Network intrusion detection using ACO-DNN model with DVFS based energy optimization in cloud framework. Sustain Comput: Inform Syst. 2022;35(September 2021):100746. doi: 10.1016/j.suscom.2022.100746.
[66] Iwendi C, Anajemba JH, Biamba C, Ngabo D. Security of things intrusion detection system for smart healthcare. Electron (Switz). 2021;10(12):1–27. doi: 10.3390/electronics10121375.
[67] Prabhakaran V, Kulandasamy A. Hybrid semantic deep learning architecture and optimal advanced encryption standard key management scheme for secure cloud storage and intrusion detection. Neural Comput Appl. 2021;33(21):14459–79. doi: 10.1007/s00521-021-06085-5.
[68] Duo R, Nie X, Yang N, Yue C, Wang Y. Anomaly detection and attack classification for train real-time ethernet. IEEE Access. 2021;9:22528–41. doi: 10.1109/ACCESS.2021.3055209.
[69] Thakkar A, Lohiya R. Role of swarm and evolutionary algorithms for intrusion detection system: A survey. Swarm Evol Comput. 2020;53(December 2019):100631. doi: 10.1016/j.swevo.2019.100631.
[70] Vijayalakshmi S, Subha TD, Manimegalai L, Reddy ES, Yaswanth D, Gopinath S. A novel approach for IoT intrusion detection system using modified optimizer and convolutional neural network. 6th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2022 - Proceedings; 2022. p. 180–6. doi: 10.1109/I-SMAC55078.2022.9987314.
[71] Om Kumar CU, Marappan S, Murugeshan B, Beaulah PMR. Intrusion detection model for IoT using recurrent Kernel convolutional neural network. Wirel Pers Commun. 2023;129(2):783–812. doi: 10.1007/s11277-022-10155-9.
[72] Al Sawafi Y, Touzene A, Hedjam R. Hybrid deep learning-based intrusion detection system for RPL IoT networks. J Sens Actuator Netw. 2023;12(2):13491–520. doi: 10.3390/jsan12020021.
[73] Lateef AAA, Al-Janabi STF, Al-Khateeb B. Hybrid intrusion detection system based on deep learning. 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy, ICDABI 2020; 2020. doi: 10.1109/ICDABI51230.2020.9325669.
[74] Keserwani PK, Govil MC, Pilli ES. An effective NIDS framework based on a comprehensive survey of feature optimization and classification techniques. Neural Comput Appl. 2023;35(7):4993–5013. doi: 10.1007/s00521-021-06093-5.
[75] Almaghthawi Y, Ahmad I, Alsaadi FE. Performance analysis of feature subset selection techniques for intrusion detection. Mathematics. 2022;10(24):4745. doi: 10.3390/math10244745.
[76] Karatas G, Demir O, Sahingoz OK. A deep learning based intrusion detection system on GPUs. Proceedings of the 11th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2019; 2019. doi: 10.1109/ECAI46879.2019.9042132.
[77] Gorzałczany MB, Rudzinski F. Intrusion detection in Internet of Things with MQTT protocol - An accurate and interpretable genetic-fuzzy rule-based solution. IEEE Internet Things J. 2022;9(24):24843–55. doi: 10.1109/JIOT.2022.3194837.

[78] Soni M, Singhal M, Jatin, Katarya R. Optimizing deep neural network using enhanced artificial bee colony algorithm for an efficient intrusion detection system. 2022 2nd International Conference on Intelligent Technologies, CONIT 2022; 2022. p. 1–7. doi: 10.1109/CONIT55038.2022.9848014.
[79] Mahboob AS, Shahhoseini HS, Ostadi Moghaddam MR, Yousefi S. A coronavirus herd immunity optimizer for intrusion detection system. 2021 29th Iranian Conference on Electrical Engineering, ICEE 2021; 2021. p. 579–85. doi: 10.1109/ICEE52715.2021.9544165.
[80] Kalaivani S, Vikram A, Gopinath G. An effective swarm optimization based intrusion detection classifier system for cloud computing. 2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019; 2019. p. 185–8. doi: 10.1109/ICACCS.2019.8728450.
[81] Davahli A, Shamsi M, Abaei G. Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks. J Ambient Intell Humaniz Comput. 2020;11(11):5581–609. doi: 10.1007/s12652-020-01919-x.
[82] Swarna Priya RM, Maddikunta PKR, Koppu S, Gadekallu TR, Chowdhary CL, et al. An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Comput Commun. 2020;160(June):139–49. doi: 10.1016/j.comcom.2020.05.048.
[83] Kunhare N, Tiwari R, Dhar J. Intrusion detection system using hybrid classifiers with meta-heuristic algorithms for the optimization and feature selection by genetic algorithm. Computers Electr Eng. 2022;103(September):108383. doi: 10.1016/j.compeleceng.2022.108383.
[84] Aljanabi M, Hayder R, Talib S, Ali AH, Mohammed MA, Sutikno T. Distributed denial of service attack defense system-based auto machine learning algorithm. Bull Electr Eng Inform. 2023;12(1):544–51.
[85] Mijwil M, Aljanabi M. Towards artificial intelligence-based cybersecurity: the practices and ChatGPT generated ways to combat cybercrime. Iraqi J Computer Sci Math. 2023;4(1):65–70.
[86] Mijwil M, Filali Y, Aljanabi M, Bounabi M, Al-Shahwani H. The purpose of cybersecurity governance in the digital transformation of public services and protecting the digital environment. Mesopotamian J Cybersecur. 2023;2023:1–6.
[87] Yaseen MG, Aljanabi M, Ali AH, Abd SA. Current cutting-edge research in computer science. Mesopotamian J Computer Sci. 2022;2022:1–4.
[88] Ali AH, Yaseen MG, Aljanabi M, Abed SA, et al. Transfer learning: A new promising techniques. Mesopotamian J Big Data. 2023;2023:31–2.
[89] Li K, Zhang Y, Wang S. An intrusion detection system based on PSO-GWO hybrid optimized support vector machine. Proceedings of the International Joint Conference on Neural Networks; 2021-July, 2021. p. 1–7. doi: 10.1109/IJCNN52387.2021.9534325.
[90] Alhajjar E, Maxwell P, Bastian N. Adversarial machine learning in network intrusion detection systems. Expert Syst Appl. 2021;186(August):115782. doi: 10.1016/j.eswa.2021.115782.
[91] Khaleel MK, Ismail MA, Yunan U, Kasim S. Review on intrusion detection system based on the goal of the detection system. Int J Integr Eng. 2018.
[92] Mohammed MA, Hasan RA, Ahmed MA, Tapus N, Shanan MA, Khaleel MK, et al. A focal load balancer based algorithm for task assignment in cloud environment. In 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). IEEE; 2018. p. 1–4.
[93] Ali AH, Aljanabi M, Ahmed MA. Fuzzy generalized Hebbian algorithm for large-scale intrusion detection system. Int J Integr Eng. 2020;12(1):81–90.
[94] Al-Janabi M, Ismail MA. Improved intrusion detection algorithm based on TLBO and GA algorithms. Int Arab J Inf Technol. 2021;18(2):170–9.
[95] Abd SN, Alsajri M, Ibraheem HR. Rao-SVM machine learning algorithm for intrusion detection system. Iraqi J Computer Sci Math. 2020;1(1):23–7.
[96] Ali AH, Abdullah MZ, Abdul-wahab SN, Alsajri M. A brief review of big data analytics based on machine learning. Iraqi J Computer Sci Math. 2020;1(2):13–5.
[97] Aljanabi M, Abd-Alwahab SN, Saedudin R, Ebraheem HR, Defni, Hadi R, et al. Cloud computing issues, challenges, and needs: A survey. JOIV: Int J Inform Vis. 2021;5(3):298–305.