A SYSTEMATIC LITERATURE REVIEW ON NETWORK INTRUSION DETECTION SYSTEMS: ATTACK TYPES, DATASETS, EVALUATION METRICS, AND OPEN CHALLENGES

[1]Joseph Adebayo Ojeniyi, [1]Danjuma Hussaini Joshua, [1]Olusanjo Olugbemi Fasola, [1]Joshua Mamza Edward, [1]Abubakar Aliyu Musa, [2]Grace Amina Onyeabor

[1]Department of Cyber Security Science, Federal University of Technology, Minna, Nigeria

[2]Department of Data Science, Federal University of Technology, Minna, Nigeria

Abstract

The rapid evolution of networked systems and the proliferation of cyber threats have made Intrusion Detection Systems (IDS) an essential component of modern cybersecurity frameworks. Among various attack categories, Denial of Service (DoS) and brute force attacks remain highly prevalent due to their disruptive impact and ease of execution. This study presents a Systematic Literature Review (SLR) on network intrusion detection systems, focusing on attack types, datasets, evaluation metrics, and open challenges. A PRISMA-based methodology was adopted to analyze fifteen recent peer-reviewed studies (2023–2026) from major academic databases, including IEEE Xplore, ScienceDirect, SpringerLink, and ACM Digital Library. The study by Abdulkareem et al. (2024) is used as the baseline reference. The review identifies key intrusion attack categories, with emphasis on DoS/DDoS and brute force attacks, alongside other threats such as botnet and web-based attacks. Widely used datasets, including CICIDS2017, CSE-CIC-IDS2018, and UNSW-NB15, are examined, revealing challenges such as class imbalance, redundancy, and limited real-world representation. Additionally, commonly used evaluation metric (accuracy, precision, recall, and F1-score) are analyzed, highlighting limitations of relying on single metrics. The study further identifies critical challenges, including high false positive rates, scalability issues, computational complexity, and the lack of lightweight IDS solutions. These findings provide a comprehensive understanding of the problem domain and a foundation for future research

Keywords: Intrusion Detection Systems (IDS), Network Security, Cybersecurity, Denial of Service (DoS), Brute Force Attacks, IDS Datasets, CICIDS2017, CSE-CIC-IDS2018, Evaluation Metrics, Systematic Literature Review (SLR)

INTRODUCTION

1.1 Background

The increasing dependence on digital communication platforms and interconnected infrastructures has greatly altered contemporary society. However, this shift has also made these systems susceptible to various cybersecurity threats. Intrusion Detection Systems (IDS) have become crucial security tools aimed at overseeing network traffic and identifying harmful activities in real time. Recent progress in networking technologies, especially the surge of Internet of Things (IoT) devices, has made the security landscape even more complex. IoT environments consist of diverse devices, possess limited computational capabilities, and often have inadequate security measures, resulting in a heightened vulnerability to cyberattacks. Research like Hozouri et al. (2025) IDS Survey and Rahman et al. (2025) IDS in IoT Networks emphasizes the increasing demand for strong and adaptable intrusion detection solutions.

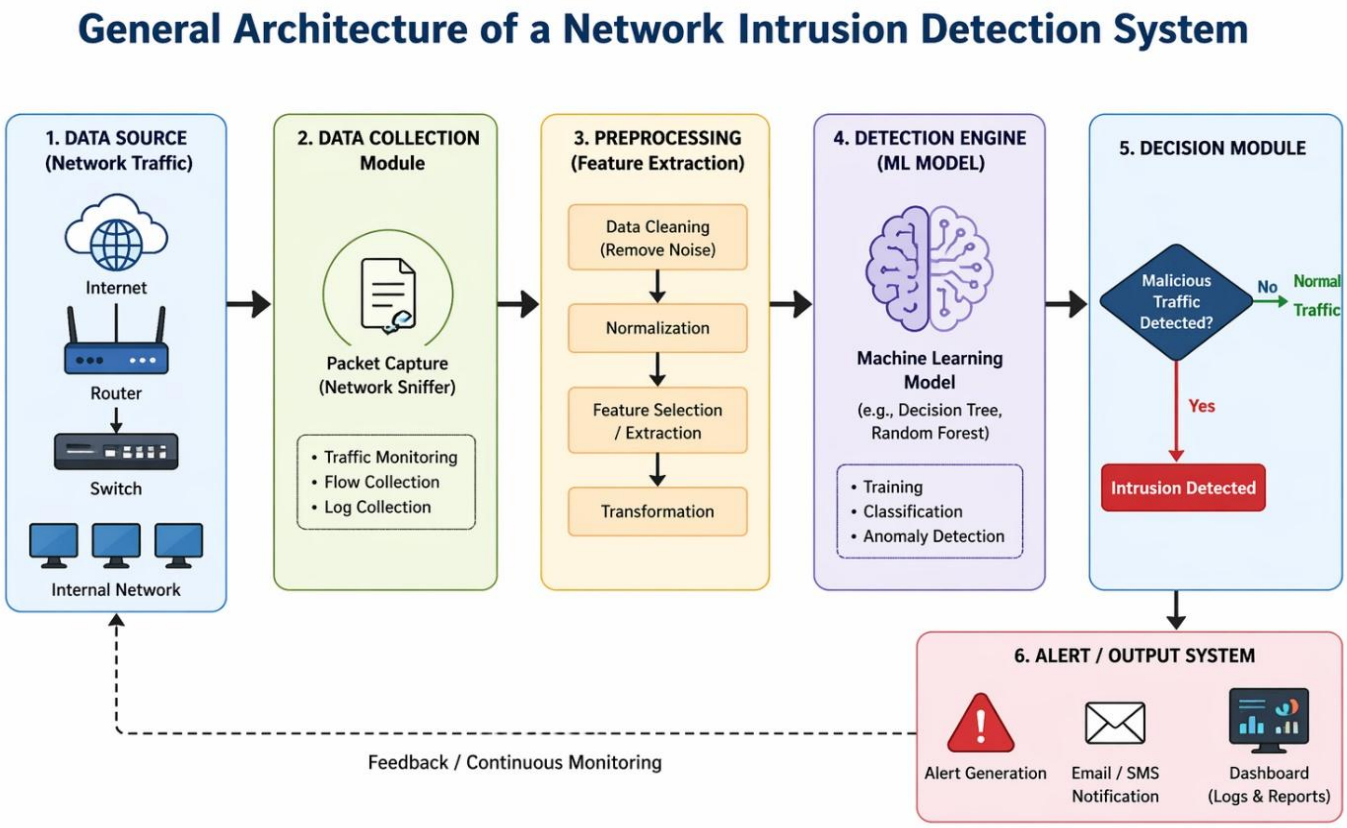


Figure 1: General Architecture of a Network Intrusion Detection System

Figure 1 illustrates the general architecture of a network intrusion detection system, showing the flow of network traffic from data collection through preprocessing, detection, and alert generation.

1.2 Statement of the Problem

Intrusion detection systems have been the subject of much research, yet there are still a number of unanswered questions. The complexity and frequency of modern cyberattacks, especially denial of service (DoS) and brute force operations, are still increasing. Current IDS systems frequently have trouble with:

- High false positive and false negative rates
- Poor generalization across different datasets
- Computational inefficiency in real-time environments
- Lack of lightweight solutions for small home networks

Furthermore, it is challenging to adequately compare various IDS techniques in the absence of defined datasets and evaluation metrics. These restrictions make it difficult to implement IDS in real-world situations, as noted by Rawat and Singal (2025) IDS Datasets and Challenges.

1.3 Topic / Aim

Problem Domain Title

This paper focus on a systematic literature review on network intrusion detection systems: attack types, datasets, evaluation metrics, and open challenges

Aim:

This paper aims to systematically review and analyze existing research on network intrusion detection systems with emphasis on attack types, datasets, evaluation metrics, and open challenges.

Objectives

The objectives of this paper are:

1. To identify and classify major intrusion attack types in network environments
2. To analyze commonly used datasets for intrusion detection research
3. To examine evaluation metrics used in assessing IDS performance
4. To identify key challenges and research gaps in existing IDS approaches

1.4 Research Questions

This study addresses the following research questions:

- **RQ1:** What are the major intrusion attack types addressed in IDS research?
- **RQ2:** What datasets are commonly used in intrusion detection studies?
- **RQ3:** What evaluation metrics are used to assess IDS performance?
- **RQ4:** What are the key challenges and open research issues in IDS?

1.5 Scope of the Study

This review focuses on network-based intrusion detection systems, particularly in IoT and modern network environments. The study emphasizes machine learning and deep learning-based approaches and considers only peer-reviewed articles published between 2023 and 2026.

1.6 Significance of the Study

This paper provides an organized and thorough compilation of the most recent developments in intrusion detection system research. It helps create more effective, scalable, and useful IDS solutions by highlighting important issues and research gaps. Researchers and practitioners working on lightweight intrusion detection systems for small-scale and resource-constrained network environments may find the findings very pertinent.

2.0 Related Works

Over the past several decades, Intrusion Detection Systems (IDS) have been the subject of extensive research, especially with the emergence of the Internet of Things (IoT), cloud computing, and expansive network infrastructures. A variety of surveys and systematic literature reviews (SLRs) have been performed to evaluate IDS methodologies, datasets, and the challenges faced. This section explores the existing literature, emphasizing systematic reviews and survey-based investigations, while noting their contributions, shortcomings, and relevance to the current research. The objective is to pinpoint gaps in the current SLRs, specifically concerning types of attacks, datasets, evaluation metrics, and challenges.

2.1 Existing Systematic Literature Reviews on IDS

Recently, several studies have performed systematic reviews on intrusion detection systems, emphasizing machine learning, deep learning, and environments based on the Internet of Things (IoT). For example, Abdulkareem et al. (2024) present a thorough overview of intrusion detection systems applicable to both IoT and non-IoT settings. The research explores different IDS techniques, datasets, and methods of evaluation, underscoring the increasing significance of machine learning in the field of intrusion detection. Nonetheless, the study does not concentrate on specific types of attacks, such as Denial of Service (DoS) and brute force attacks.

In a similar vein, Slaibi et al. (2026) offer a thorough analysis of the literature on IoT security, focusing on intrusion detection techniques and related difficulties. The study lacks a thorough examination of datasets and assessment measures, although offering insightful information about IoT-specific IDS strategies. Hozouri et al. (2025) use deep learning and machine learning techniques to investigate developments in IDS. The study shows how well these methods work to increase detection accuracy, but it also points out issues like computational complexity and scalability. (Deng et al., 2025) highlights a SLR of recent GenAI applications in the environments of IoT IDS and perform analysis of the model techniques and architecture. Common usage was classified such as data argument, class balancing, reconstruction of data and attack adversaries' generation. Datasets used were outlined and evaluated using specific metrics to compare the performances of each model under these conditions. (Ali et al., 2020) conducted reviewed attacks

on IoT-based botnet. The reviewed outlined the existing contributions, utilized datasets network forensics methods and research in this environment where its gaining more strength over the last 3-years. (Arreche & Abdallah, 2025) apply and evaluated White box explainable AI (XAI) methods for network intrusion detechtion system using an end-to-end framework with attribution of neural network methods such as LRP, integrated gradients (IG) and DeepLift. The research article compares white box against black box XAI and evaluated using metrics. White box methods recorded high on robustness and completeness which identify the need for IDS trustworthiness. (Sharmin et al., 2026) investigate the integration of Federated Learning (FL) for cyber attack detection in Edge computing environments using PRSIMA methodology across five different sources of datasets. The results shows that the PRISMA compliant across seven thematic areas with comparative analysis, recommendations on privacy performance trade-offs, scalability and challenges and lastly forward-looking research agenda addressing generative models, defense collaborations, 6G enabled intelligence and Zero trust architectures. (Kazimierczak et al., 2024) conducted a study and categorized AI tool used by adversaries for sophisticated attacks and a defense system mechanism was suggested to counter these evolving threats. The results indicated that while AI significantly empowers attackers in the first stage of intrusion, the current defense tools are not yet sophisticated to mitigate these threats. Hence, a future guide on development of cyber defense mechanism to address these threats is proposed.

2.2 Machine Learning and Deep Learning-Based IDS Reviews

Machine learning (ML) and deep learning (DL) techniques have become dominant in IDS research due to their ability to detect complex attack patterns. A thorough analysis of deep learning-based IDS techniques is presented by Chinnasamy et al. (2025), who emphasize how well they perform in comparison to conventional techniques. The paper does point out that these models frequently need a lot of processing power, which makes them less appropriate for implementation in real time.

Deep learning methods for intrusion detection in IoT environments are also examined by Liao et al. (2024), who highlight the high detection accuracy of these methods. Nevertheless, the study finds problems including overfitting and reliance on big datasets. Aldhaheri et al. (2024) investigate deep learning-based cyber threat detection in more detail, emphasizing how well it can detect sophisticated threats. The study does, however, highlight the need for lighter and more effective models.

2.3 Lightweight and Resource-Constrained IDS Approaches

Lightweight IDS solutions are becoming more and more popular as IoT devices and small-scale networks become more widely used. Al Mukhaini et al. (2024) emphasize the significance of lowering computing complexity while preserving detection accuracy in their work on lightweight detection techniques for IoT networks.

According to the survey, one of the biggest obstacles to installing IDS in IoT contexts is resource limitations. Khadam et al. (2025) further highlight the necessity of effective AI-based solutions that can function in settings with limited resources. The study emphasizes the necessity for scalable

models and the difficulties associated with deployment. The majority of current IDS systems are still computationally demanding despite these efforts, suggesting a gap in the development of lightweight IDS for tiny residential networks.

2.4 IDS Datasets and Evaluation Studies

Datasets are a critical component of IDS research, as they directly influence model performance and evaluation. Rawat and Singal (2025) provides a comprehensive survey of IDS datasets, tools, and challenges. The paper identifies problems including class imbalance and lack of realism while highlighting widely used datasets like CICIDS2017 and IDS2018. In their discussion of the function of datasets in IoT-based IDS, Rahman et al. (2025) stress the significance of choosing suitable datasets for model evaluation and training. In their analysis of IDS optimization methods, Sharma et al. (2024) emphasize the significance of feature selection and evaluation metrics in enhancing model performance.

2.5 IDS in Specialized Domains

IDS applications in particular fields including healthcare, smart grids, and IoT environments have been studied recently. In their investigation of intrusion detection in the Internet of Medical Things (IoMT), Hassan et al. (2025) draw attention to issues with data security and privacy. Afrin et al. (2026) investigate AI-based intrusion detection in smart grid settings, highlighting the significance of threat mitigation and real-time detection. Nazir et al. (2023) concentrate on botnet detection in IoT networks, emphasizing the need for sophisticated detection methods due to the growing complexity of dispersed attacks.

Table 1: Summary of Related SLR and Survey Studies

S/N	Study	Focus Area	Key Contribution	Limitation
1	Abdulkareem et al., 2024	IDS (IoT & non-IoT)	Comprehensive IDS survey	No focus on specific attacks
2	Hozouri et al., 2025	ML/DL IDS	Advanced IDS techniques	High computational cost
3	Rahman et al., 2025	IoT IDS	ML-based IDS insights	Limited dataset diversity
4	Slaibi et al., 2026	IoT Security SLR	Identifies IoT challenges	Limited metrics analysis
5	Chinnasamy et al., 2025	DL-based IDS	High detection performance	Resource-intensive
6	Rawat & Singal, 2025	IDS datasets/tools	Dataset comparison	Limited attack focus
7	Al Mukhaini et al., 2024	Lightweight IDS	Efficient detection models	Limited real-world testing
8	Berhili et al., 2024	ML-based IDS	ML effectiveness in IDS	Generalized analysis
9	Nazir et al., 2023	IoT botnet IDS	Botnet detection methods	Focus on specific attack

10	Sharma et al., 2024	Optimization in IDS	Improves model performance	Increased complexity
11	Liao et al., 2024	DL for IDS	High accuracy detection	Overfitting issues
12	Aldhaheri et al., 2024	DL threat detection	Detects complex attacks	High computation
13	Khadam et al., 2025	AI in IoT	AI deployment insights	Practical limitations
14	Hassan et al., 2025	IoMT IDS	Healthcare IDS analysis	Privacy challenges
15	Afrin et al., 2026	Smart grid IDS	AI-based IDS solutions	Real-time constraints

Table 1 summarizes key systematic literature reviews and survey studies on intrusion detection systems, highlighting their focus areas, contributions, and limitations. The table shows that while existing studies provide valuable insights into IDS techniques and datasets, they often lack a unified analysis of attack types, datasets, evaluation metrics, and open challenges.

2.6 Research Gap Identification

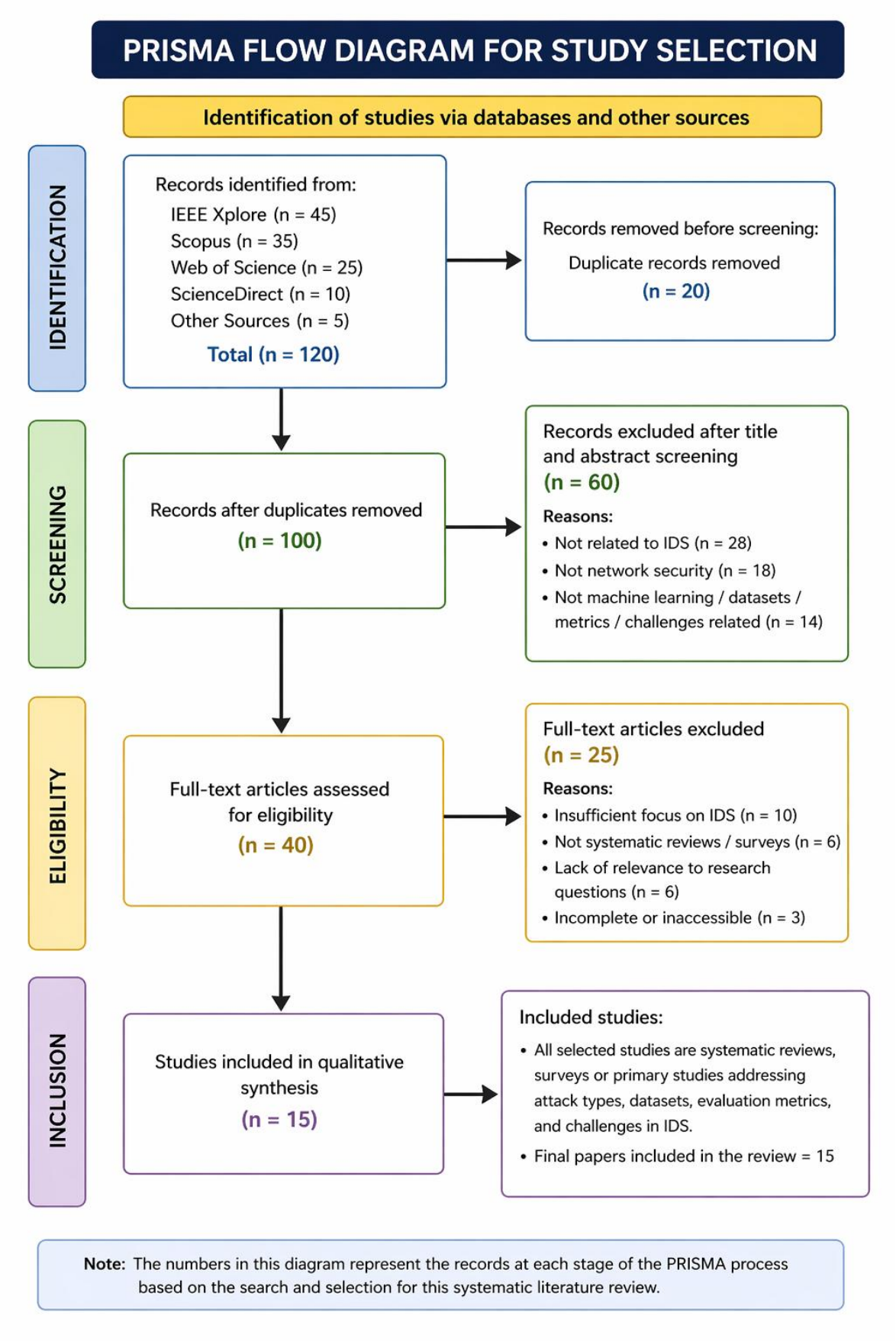
Based on the reviewed studies, several gaps are identified:

- Lack of integrated analysis covering attack types, datasets, metrics, and challenges
- Limited focus on brute force attacks compared to DoS attacks
- Absence of standardized evaluation frameworks
- Insufficient development of lightweight IDS for small home networks

These gaps justify the need for the present study, which aims to provide a comprehensive SLR addressing these aspects in a unified manner.

3.0 SLR PRISMA METHODS

The systematic method used to carry out this investigation is explained in this section. To guarantee transparency, reproducibility, and rigor in the selection and analysis of the relevant literature, a


PRISMA FLOW DIAGRAM FOR STUDY SELECTION
Identification of studies via databases and other sources
IDENTIFICATION
Records identified from:
IEEE Xplore (n = 45)
Scopus (n = 35)
Web of Science (n = 25)
ScienceDirect (n = 10)
Other Sources (n = 5)
Total (n = 120)
Records removed before screening:
Duplicate records removed
(n = 20)
SCREENING
Records after duplicates removed
(n = 100)
Records excluded after title and abstract screening
(n = 60)
Reasons:
• Not related to IDS (n = 28)
• Not network security (n = 18)
• Not machine learning / datasets / metrics / challenges related (n = 14)
ELIGIBILITY
Full-text articles assessed for eligibility
(n = 40)
Full-text articles excluded
(n = 25)
Reasons:
• Insufficient focus on IDS (n = 10)
• Not systematic reviews / surveys (n = 6)
• Lack of relevance to research questions (n = 6)
• Incomplete or inaccessible (n = 3)
INCLUSION
Studies included in qualitative synthesis
(n = 15)
Included studies:
• All selected studies are systematic reviews, surveys or primary studies addressing attack types, datasets, evaluation metrics, and challenges in IDS.
• Final papers included in the review = 15
Note: The numbers in this diagram represent the records at each stage of the PRISMA process based on the search and selection for this systematic literature review.

standardized PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach was employed.

Figure 2 illustrates the PRISMA-based study selection process, showing the number of articles identified, screened, filtered, and finally included in the review. The methodology consists of four main phases: identification, screening, eligibility and inclusion. The study selection process is illustrated in the PRISMA flow diagram.

3.1 PICOC (The Population, Intervention, Comparison, Outcome and Context) Framework

The PICOC framework was used to define the scope of the study and guide the formulation of research questions and search strings.

Table 2: PICOC Keywords and Synonyms Definition

PICOC Element	Description	Keywords	Synonyms
Population	Network environments and systems	Intrusion Detection Systems (IDS)	Network security systems, NIDS
Intervention	Detection techniques and methods	Machine Learning, Deep Learning	AI-based detection, classification models
Comparison	Types of attacks	DoS, Brute Force	DDoS, password attacks, authentication attacks
Outcome	Performance evaluation	Accuracy, Precision, Recall, F1-score	Detection rate, classification performance
Context	Application environments	IoT, Network systems	Smart homes, cyber-physical systems

3.2 Research Questions

The research questions were derived based on the PICOC framework to ensure alignment with the study objectives.

Table 3: Research Questions

S/N	Research Questions	Rationale (PICOC Elements)	Search Terms / Strings
1	What are the major intrusion attack types addressed in IDS research?	Comparison	“IDS AND attack types”, “DoS detection”, “Brute force IDS”
2	What datasets are commonly used in IDS research?	Population	“IDS datasets”, “CICIDS2017”, “IDS2018 dataset”
3	What evaluation metrics are used in IDS performance assessment?	Outcome	“IDS evaluation metrics”, “accuracy precision recall IDS”
4	What are the key challenges and open issues in IDS?	Context	“IDS challenges”, “intrusion detection limitations”

3.3 Digital Libraries Description

To ensure the inclusion of high-quality and peer-reviewed studies, multiple reputable academic databases were used.

Table 4: Digital Libraries Description

Database	Description	URL	Area	Advanced Search
ScienceDirect	Elsevier's leading database for scientific research	https://www.sciencedirect.com/	Multidisciplinary	Yes
ACM Digital Library	Repository for computing research	https://dl.acm.org/	Computing & IT	Yes
IEEE Xplore	Engineering and technology research database	https://ieeexplore.ieee.org/	Engineering & CS	Yes
SpringerLink	Scientific publications across disciplines	https://link.springer.com/	Multidisciplinary	Yes
Scopus	Indexed high-quality journals	https://www.scopus.com/	Multidisciplinary	Yes
Web of Science	Curated academic journals database	https://mjl.clarivate.com/	Multidisciplinary	Yes

3.4 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were defined to ensure that only relevant and high-quality studies were selected.

Table 5: Inclusion and Exclusion Criteria Definition

Criteria Type	Description	Inclusion	Exclusion
Period	Publication year	2023–2026	Before 2023
Language	Language of publication	English	Non-English
Type of Literature	Source type	Journals & Conferences	Grey literature
Type of Source	Academic databases	IEEE, ACM, Springer, Elsevier	Books, blogs
Impact of Source	Quality ranking	Q1 and Q2 journals	Low-impact journals
Accessibility	Availability	Full-text accessible	Not accessible
Relevance	Relation to IDS	Relevant to at least 2 RQs	Irrelevant studies

3.5 Article Quality Assessment

A quality assessment checklist was used to evaluate the selected articles based on reporting quality, methodological rigor, credibility, and relevance.

Table 6: Article Quality Assessment Checklist

S/N	Article Title / Focus	Author / Year	Reporting	Rigor	Credibility	Relevance	Total
1	Network Intrusion Detection Survey	Abdulkareem et al., 2024	5	5	5	5	20
2	IDS with ML & DL Advances	Hozouri et al., 2025	5	4	4	5	18
3	IDS in IoT Networks	Rahman et al., 2025	4	4	4	5	17
4	IoT Security SLR	Slaibi et al., 2026	5	5	5	4	19
5	DL-based Network IDS Review	Chinnasamy et al., 2025	4	4	4	5	17
6	IDS Datasets & Tools Survey	Rawat & Singal, 2025	5	5	5	5	20
7	Lightweight IDS in IoT	Al Mukhaini et al., 2024	4	4	4	5	17
8	ML-based IDS State of the Art	Berhili et al., 2024	4	4	4	5	17
9	IoT Botnet Detection Review	Nazir et al., 2023	4	4	4	4	16
10	Optimization in IDS	Sharma et al., 2024	4	4	4	4	16
11	DL for IDS in IoT	Liao et al., 2024	5	4	5	5	19
12	DL Threat Detection Review	Aldhaheri et al., 2024	4	4	4	5	17
13	AI in IoT Systems Review	Khadam et al., 2025	4	4	4	4	16
14	IDS in IoMT	Hassan et al., 2025	4	4	4	5	17
15	AI Cybersecurity Smart Grid	Afrin et al., 2026	5	5	5	4	19

3.6 Data Extraction and Synthesis

Data from the selected studies were systematically extracted and synthesized to address the research questions.

Table 7: Data Extraction and Synthesis

s/n	Article Title / Focus	Author / Year	Research Type	Process Phases / Stages	Technology / Algorithm / Framework	Dataset	Application Domain	Gaps / Challenges	Key Findings	Evaluation Metrics
1	Network IDS Survey	Abdulkareem et al., 2024	Survey	Collection → Analysis → Comparison	ML & DL IDS	CICIDS2017, UNSW-NB15	Network/IoT	Dataset imbalance	ML improves IDS performance	Accuracy, Recall
2	IDS ML/DL Advances	Hozouri et al., 2025	Survey	Review → Classification	ML, DL models	Multiple	Cybersecurity	Scalability issues	DL improves detection	F1-score
3	IDS in IoT Networks	Rahman et al., 2025	Survey	Analysis	ML-based IDS	CICIDS2017	IoT	False positives	ML effective	Accuracy
4	IoT Security SLR	Slaibi et al., 2026	SLR	PRISMA → Filtering	IDS frameworks	Various	IoT	Lack of standardization	Need realistic datasets	Precision, Recall
5	DL-based IDS Review	Chinnasamy et al., 2025	SLR	Training → Testing	Deep learning	Benchmark datasets	Network	High computation	DL accurate but heavy	Accuracy, F1
6	IDS Datasets Survey	Rawat & Singal, 2025	Survey	Collection → Comparison	IDS tools, ML	CICIDS2017, IDS2018	IoT	Data imbalance	Dataset critical to IDS	Accuracy
7	Lightweight IDS	Al Mukhaini et al., 2024	SLR	Feature extraction → Detection	Lightweight ML/DL	IoT datasets	IoT	Resource limits	Lightweight IDS needed	Precision

8	ML-based IDS	Berhili et al., 2024	Survey	Analysis	ML techniques	Various	IoT	Limited generalization	ML improves IDS	Accuracy
9	Botnet Detection	Nazir et al., 2023	SLR	Detection pipeline	ML algorithms	IoT datasets	IoT	Botnet complexity	ML detects botnets	Recall
10	Optimization IDS	Sharma et al., 2024	SLR	Optimization → Evaluation	Optimization algorithms	Various	IoT	Complexity	Optimization improves IDS	F1-score
11	DL IDS	Liao et al., 2024	Survey	Training → Evaluation	Deep learning	Benchmark datasets	IoT	Data dependency	DL high accuracy	Accuracy
12	DL Threat Detection	Aldhaheri et al., 2024	Review	Analysis	DL models	Various	IoT	Overfitting	DL effective	Precision
13	AI in IoT	Khadam et al., 2025	SLR	Review	AI frameworks	Various	IoT	Deployment issues	AI promising	Accuracy
14	IDS in IoMT	Hassan et al., 2025	Survey	Detection pipeline	ML/DL	IoMT datasets	Healthcare	Privacy issues	IDS needed in IoMT	Recall
15	Smart Grid IDS	Afrin et al., 2026	SLR	Detection → Mitigation	AI-based IDS	Smart grid datasets	Energy	Real-time constraints	AI improves IDS	Accuracy, F1

RESULTS AND DISCUSSION

This section presents and discusses the findings of the systematic literature review based on the selected studies. The results are organized according to the defined research questions (RQ1–RQ4), focusing on intrusion attack types, datasets, evaluation metrics, and open challenges.

Table 8: Descriptive Statistics of the Systematic Review Articles

RQ No	Research Question	Stage	ScienceDirect	ACM	IEEE	Springer	Web of Science	Total
RQ1	Attack Types	Identified	18	10	22	15	12	77
		Filtered	10	6	12	8	6	42
		Included	4	2	5	2	2	15
RQ2	Datasets	Identified	15	8	20	12	10	65

		Filtered	9	5	11	7	5	37
		Included	4	2	4	3	2	15
RQ3	Metrics	Identified	14	7	18	11	9	59
		Filtered	8	4	10	6	5	33
		Included	4	2	4	3	2	15
RQ4	Challenges	Identified	16	9	19	13	11	68
		Filtered	9	5	11	7	6	38
		Included	4	2	5	2	2	15

4.1 RQ1: What are the major intrusion attack types addressed in IDS research?

The examination of the fifteen chosen publications shows that a number of clearly defined attack categories are the main focus of intrusion detection research. These include web-based attacks, botnet attacks, brute force attacks, and denial of service (DoS/DDoS) attacks. Due to their significant influence on network availability and frequent occurrence in real-world settings, DoS and DDoS attacks predominate in the literature. Due to their ease of use and ability to interrupt services, DoS attacks continue to be a major problem in both traditional and IoT-based networks, according to recent research like Abdulkareem et al. (2024) and Hozouri et al. (2025). In a similar vein, Rahman et al. (2025) emphasize that brute force assaults are extremely important for authentication systems, particularly for protocols like SSH and FTP.

Additionally, Nazir et al. (2023) identifies botnet attacks as a growing threat in IoT environments, where compromised devices are used to launch large-scale distributed attacks. Web-based attacks, although less emphasized in comparison, are still included in benchmark datasets and remain relevant in application-layer security (Berhili et al., 2024).

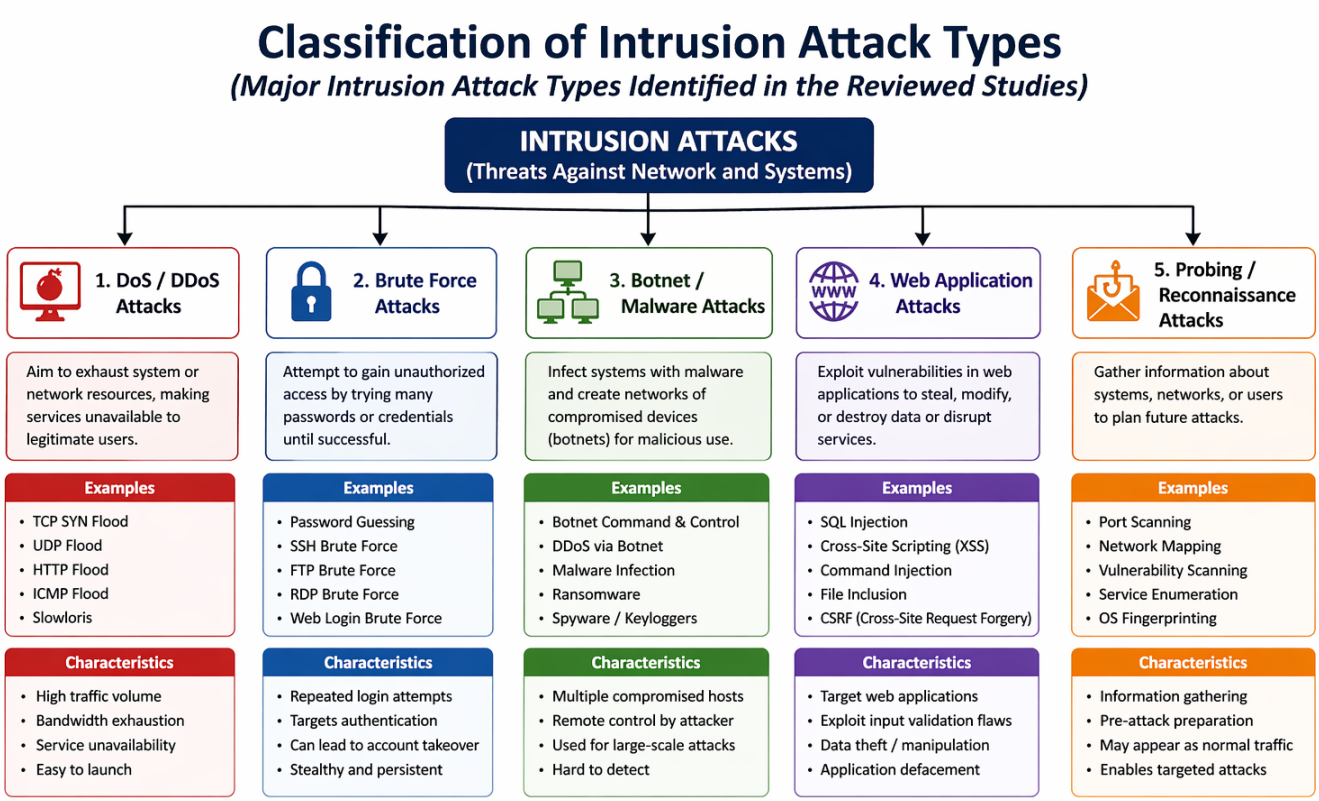


Figure 3: Classification of Intrusion Attack Types

4.1.1 Denial of Service (DoS/DDoS)

Attacks known as denial of service (DoS) and distributed denial of service (DDoS) seek to deplete network resources in order to stop authorized users from using services. Because these assaults are common and have a major effect on network performance, they are extensively researched. DoS attacks make up a significant percentage of intrusion detection datasets, according to Abdulkareem et al. (2024), making them a key benchmark for assessing IDS performance. According to Hozouri et al. (2025), machine learning models' capacity to recognize anomalous traffic increases makes them very useful in identifying DoS patterns. However, one of the major challenges in detecting DoS attacks is distinguishing malicious traffic from legitimate high-volume traffic, which often leads to false positives (Liao et al., 2024).

4.1.2 Brute Force Attacks

Brute force attacks pose a serious risk to authentication systems because they entail repeated attempts to guess user credentials. These attacks are frequently seen in web login systems, SSH, and FTP. According to Rahman et al. (2025), pattern-based machine learning models can detect brute force assaults since they frequently include several login attempts in a little period of time. However, there is a research gap because brute force attacks are given less attention in the literature than DoS attacks. For small home networks, where inadequate authentication procedures may make them more susceptible to such assaults, this vulnerability is especially crucial.

4.1.3 Botnet Attacks

A network of infected devices that are remotely controlled to carry out coordinated harmful actions is the basis of botnet assaults. Because there are so many susceptible connected devices in IoT environments, these attacks are very common there. According to Nazir et al. (2023), the widespread and stealthy nature of these attacks makes botnet detection difficult. Furthermore, Al Mukhaini et al. (2024) highlight that because botnet traffic resembles typical network activity, standard IDS models have trouble identifying it.

4.1.4 Web-Based Attacks

According to Berhili et al. (2024), web attacks are frequently included in IDS datasets to provide a thorough evaluation of detection systems, even though they are not always the primary focus. Web-based attacks, such as SQL injection and cross-site scripting (XSS), target vulnerabilities in web applications. While these attacks are less emphasized in IDS research than DoS and brute force attacks, they remain relevant in datasets such as CICIDS2017 (Rawat & Singal, 2025).

Table 9: Classification of Intrusion Attack Types

Attack Type	Description	Impact	Detection Challenge
DoS/DDoS	Resource exhaustion attacks	Service disruption	Hard to distinguish from normal traffic
Brute Force	Password guessing attacks	Unauthorized access	Pattern similarity
Botnet	Distributed compromised devices	Large-scale attacks	Stealthy behavior
Web Attacks	Application-layer attacks	Data compromise	Signature variability

Table 9 presents a classification of the major intrusion attack types identified in the reviewed literature. It highlights their characteristics, impact on network systems, and the associated challenges in detection. The table shows that while DoS attacks dominate the literature, brute force and botnet attacks present unique detection challenges that require specialized approaches.

4.1.5 Discussion of RQ1 Findings

Because DoS and DDoS attacks are readily available in benchmark datasets and have a substantial influence on network performance, the results show that intrusion detection research is strongly biased toward these types of attacks. But the comparatively narrow focus on brute force assaults points to a void in the literature, especially when it comes to systems that rely on authentication. Furthermore, the necessity for more sophisticated and adaptive detection methods is highlighted by the growing frequency of botnet attacks in IoT contexts. Overall, the findings point to the need for more balanced research that covers a broader range of intrusion situations, even though current IDS techniques are successful in identifying some types of intrusions.

4.2 RQ2: What datasets are commonly used in IDS research?

Datasets play a critical role in the development and evaluation of intrusion detection systems, as they directly influence model training, testing, and performance validation. The analysis of the selected studies shows that a few benchmark datasets dominate IDS research, particularly CICIDS2017, CSE-CIC-IDS2018, and UNSW-NB15.

Rawat and Singal (2025) claim that because different datasets differ in terms of traffic realism, feature representation, and attack diversity, the stated performance of IDS models is greatly impacted by the dataset selection. Similarly, Abdulkareem et al. (2024) highlight that it is challenging to accurately evaluate various IDS techniques due to the absence of consistent datasets.

4.2.1 CICIDS2017 Dataset

One of the most popular datasets in intrusion detection research is CICIDS2017. It was created to produce realistic network traffic including both malicious and benign activity. According to Rawat and Singal (2025), CICIDS2017 has more than 80 features and a variety of attack scenarios, such as DoS, brute force, and web-based attacks. It is appropriate for assessing machine learning-based IDS models because of its diversity. However, a number of studies, such as Al Mukhaini et al. (2024), note that the dataset has a class imbalance, meaning that some assault types are overrepresented. Biased models that perform well on dominant classes but badly on minority attacks may result from this imbalance.

4.2.2 CSE-CIC-IDS2018 Dataset

An expansion of CICIDS2017, the CSE-CIC-IDS2018 dataset was created to overcome some of its shortcomings by incorporating more recent attack scenarios and higher network traffic volumes. According to Abdulkareem et al. (2024), IDS2018 offers increased diversity and realism, which makes it more appropriate for assessing contemporary intrusion detection systems. Furthermore, Liao et al. (2024) highlight that the scale of this dataset makes it very helpful for deep learning models. Despite these benefits, the dataset presents computational difficulties because its scale necessitates a large amount of computing power and storage, making it unsuitable for lightweight IDS implementations.

4.2.3 UNSW-NB15 Dataset

UNSW-NB15 is another commonly used dataset that provides a mix of synthetic and real network traffic. It includes a variety of modern attack types and is often used for benchmarking IDS models. Rahman et al. (2025) highlights that UNSW-NB15 offers a more balanced distribution of attack classes compared to CICIDS2017. However, its synthetic nature raises concerns about its ability to fully represent real-world network behavior.

Table 10: Dataset Comparison

Dataset	Year	Features	Attack Types	Strength	Limitation
CICIDS2017	2017	80+	Multiple	Realistic traffic, widely used	Class imbalance
IDS2018	2018	80+	Updated attacks	More realistic, large-scale	High computational cost
UNSW-NB15	2015	49	Diverse	Balanced dataset	Partially synthetic

Table 10 compares the most commonly used datasets in intrusion detection research based on their features, strengths, and limitations. It shows that while CICIDS2017 and IDS2018 provide realistic and comprehensive datasets, they suffer from issues such as class imbalance and computational complexity. On the other hand, UNSW-NB15 offers a more balanced dataset but lacks full realism due to its synthetic nature.

4.2.4 Dataset Challenges

Despite the availability of multiple benchmark datasets, several challenges persist:

1. Class Imbalance

Many datasets contain significantly more normal traffic than attack traffic, leading to biased models. Al Mukhaini et al. (2024) emphasizes that this imbalance can reduce the effectiveness of IDS models in detecting rare attacks.

2. Redundant and Irrelevant Features

Datasets often include features that do not contribute significantly to detection performance. According to Sharma et al. (2024), feature selection is essential to improve model efficiency and accuracy.

3. Lack of Real-World Representation

Even widely used datasets may not fully capture real-world network conditions. Slaibi et al. (2026) highlights that many datasets are generated in controlled environments, limiting their generalizability.

4. Dataset Inconsistency

Different studies use different datasets, making it difficult to compare results. Abdulkareem et al. (2024) notes that this lack of standardization is a major issue in IDS research.

4.2.5 Discussion of RQ2 Findings

The investigation shows that despite their widespread use, datasets like CICIDS2017 and IDS2018 have certain drawbacks. The performance and assessment of IDS models are greatly impacted by

the dataset selection, which makes it a crucial component of study design. Additionally, a significant research gap is highlighted by the absence of standardized and fully realistic datasets. To create IDS models that function well in real-world settings, this problem must be resolved. The popularity and applicability of CICIDS2017 and IDS2018 make them appropriate for this thesis; nonetheless, rigorous preprocessing and feature selection will be required to minimize their limitations.

4.3 RQ3: What evaluation metrics are used in IDS performance assessment?

Evaluation metrics are crucial for determining how effective intrusion detection systems (IDS) are. According to the reviewed studies, classification-based measures including accuracy, precision, recall, and F1-score are frequently used to assess IDS performance. However, a number of studies caution that depending solely on one metric (particularly accuracy) can be deceptive, especially when working with imbalanced datasets, which are typical in research on intrusion detection (Al Mukhaini et al., 2024; Abdulkareem et al., 2024). A variety of indicators should be employed to provide a more thorough assessment of IDS performance, according to Liao et al. (2024) and Chinnasamy et al. (2025).

4.3.1 Accuracy

Accuracy measures the proportion of correctly classified instances (both normal and malicious) out of the total number of instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

While accuracy is widely used due to its simplicity, it can be misleading in IDS contexts where normal traffic dominates. For example, a model may achieve high accuracy by simply predicting most traffic as normal, while failing to detect actual attacks (Rahman et al., 2025).

4.3.2 Precision

Precision measures the proportion of correctly identified attack instances among all instances classified as attacks.

$$Precision = \frac{TP}{TP + FP}$$

High precision indicates a low false positive rate, which is important for reducing unnecessary alerts. According to Sharma et al. (2024), precision is particularly important in real-world IDS deployment where excessive false alarms can overwhelm system administrators.

4.3.3 Recall (Detection Rate)

Recall, also known as the detection rate, measures the proportion of actual attacks that are correctly identified.

$$Recall = \frac{TP}{TP + FN}$$

A high recall value indicates that the IDS is effective in detecting attacks. However, improving recall may sometimes increase false positives, creating a trade-off with precision (Liao et al., 2024).

4.3.4 F1-Score

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of performance. Chinnasamy et al. (2025) highlights that the F1-score is particularly useful when dealing with imbalanced datasets, as it considers both false positives and false negatives.

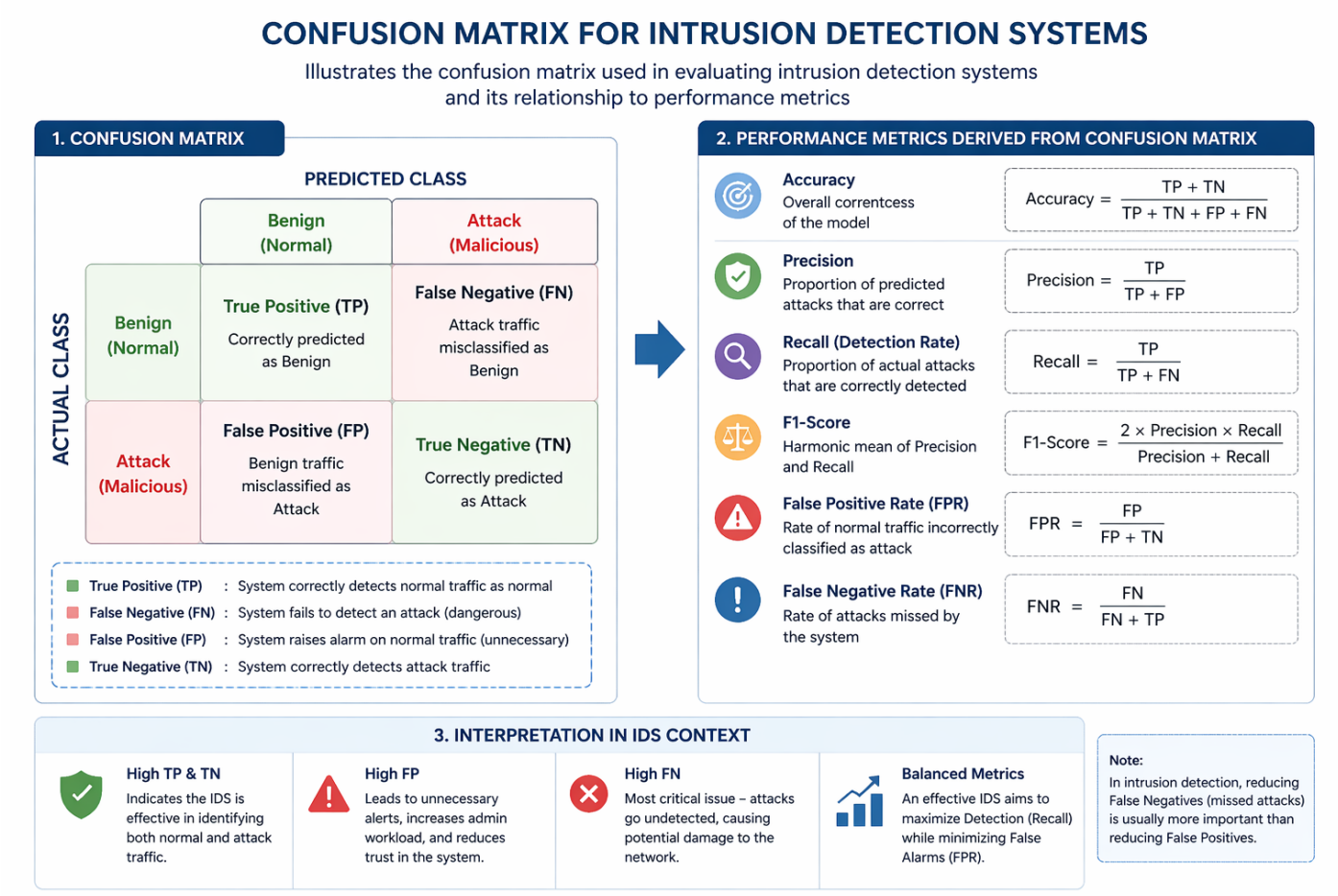


Figure 4: Confusion Matrix for IDS Evaluation

Table 11: Evaluation Metrics

Metric	Purpose	Strength	Limitation
Accuracy	Measures overall correctness	Simple and intuitive	Misleading with imbalanced data
Precision	Measures correctness of positive predictions	Reduces false alarms	Ignores false negatives
Recall	Measures detection capability	Captures attack detection rate	May increase false positives
F1-score	Balances precision and recall	Suitable for imbalanced datasets	Less intuitive

Table 11 summarizes the key evaluation metrics used in intrusion detection systems, highlighting their purpose, strengths, and limitations. The table shows that while accuracy is widely used, it is not sufficient on its own, particularly in imbalanced datasets. Metrics such as precision, recall, and F1-score provide a more reliable evaluation of IDS performance.

4.3.5 Evaluation Challenges

1. Misleading Accuracy

Accuracy alone does not reflect true performance in imbalanced datasets. Abdulkareem et al. (2024) emphasizes that models can achieve high accuracy while failing to detect attacks.

2. Trade-off Between Precision and Recall

Improving recall often increases false positives, while improving precision may reduce detection capability. This trade-off must be carefully managed (Sharma et al., 2024).

3. Lack of Standard Evaluation Framework

Different studies use different combinations of metrics, making it difficult to compare results across research works (Slaibi et al., 2026).

4. Dataset Influence on Metrics

The choice of dataset significantly affects evaluation results. Liao et al. (2024) notes that models may perform well on one dataset but poorly on another due to differences in data distribution.

4.3.6 Discussion of RQ3 Findings

The results show that there is no standard method for assessing IDS performance, despite the availability of several evaluation criteria. The majority of studies use a combination of accuracy, precision, recall, and F1-score; nevertheless, cross-study comparisons are challenging due to inconsistent measure selection. Additionally, the occurrence of unbalanced datasets emphasizes how crucial it is to use metrics other than accuracy. For intrusion detection systems to successfully

detect malicious activity, the F1-score and recall are especially crucial.To provide a thorough assessment of your model for your thesis, it is advised to combine accuracy, precision, recall, and F1-score with confusion matrix analysis.

4.4 RQ4: What are the key challenges and open issues in IDS?

The examined literature identifies a number of enduring issues that restrict the efficacy, scalability, and practical usability of intrusion detection systems, despite notable progress in this field. These issues affect deployment environments, algorithms, datasets, and evaluation techniques. Research like Abdulkareem et al. (2024), Hozouri et al. (2025), and Slaibi et al. (2026) repeatedly show that problems like high false positive rates, dataset limitations, computational complexity, and a lack of lightweight solutions limit current IDS approaches, particularly those based on machine learning and deep learning.

4.4.1 High False Positive Rates

One of the most widely reported challenges in IDS is the high rate of false positives, where legitimate network traffic is incorrectly classified as malicious. Rahman et al. (2025) notes that false positives reduce the reliability of IDS systems and can overwhelm system administrators with u znnecessary alerts. Similarly, Liao et al. (2024) emphasizes that even advanced deep learning models struggle to completely eliminate false alarms due to overlapping patterns between normal and attack traffic.

This issue is particularly critical in real-world deployment, where excessive alerts can lead to alert fatigue, reducing the effectiveness of security monitoring.

4.4.2 Dataset Limitations

The effectiveness of IDS models is highly dependent on the quality and characteristics of the datasets used for training and evaluation.

Al Mukhaini et al. (2024) identifies several dataset-related challenges, including:

- Class imbalance
- Redundant features
- Limited diversity of attack types

Slaibi et al. (2026) further points out that many datasets are generated in controlled environments, which do not accurately reflect real-world network behavior. This limits the generalization capability of IDS models.

4.4.3 Scalability Issues

Modern networks generate massive volumes of data, making scalability a major challenge for IDS systems. Hozouri et al. (2025) highlights that many IDS models, particularly deep learning-based approaches, struggle to handle large-scale network traffic efficiently. As network size increases,

detection latency and computational overhead also increase. This challenge is especially relevant in IoT environments, where a large number of connected devices continuously generate traffic.

4.4.4 Computational Complexity

Many state-of-the-art IDS models rely on complex machine learning and deep learning algorithms, which require significant computational resources. According to Chinnasamy et al. (2025), deep learning models have excellent accuracy but are frequently computationally costly and challenging to use in real-time settings. In a similar vein, Sharma et al. (2024) stress that while optimization methods can boost performance, they may also make the system more complex. Because of this, these models are inappropriate for settings with limited resources, including modest home networks.

4.4.5 Lack of Lightweight IDS for Small Home Networks

The absence of lightweight intrusion detection systems intended for residential or small-scale network environments is one of the most significant gaps found in this analysis. The majority of IDS systems currently in use were created for large-scale or enterprise networks with plenty of processing power. Small home networks, on the other hand, frequently have limited processing power and call for effective, affordable solutions. The necessity for lightweight IDS models that can function well in resource-constrained contexts without sacrificing detection accuracy is emphasized by Al Mukhaini et al. (2024) and Khadam et al. (2025).

Table 12: Open Challenges in IDS

Challenge	Description	Impact
High False Positives	Incorrect classification of normal traffic	Reduces trust in IDS
Dataset Limitations	Imbalance, lack of realism	Poor generalization
Scalability Issues	Difficulty handling large data volumes	Performance degradation
Computational Complexity	High resource requirements	Limits deployment
Lack of Lightweight IDS	No efficient solutions for small networks	Research gap

Table 12 summarizes the key challenges identified in intrusion detection systems based on the reviewed literature. It highlights the major limitations affecting IDS performance and deployment, emphasizing the need for lightweight, scalable, and efficient solutions, particularly for small home network environments.

4.5 Research Gaps

Based on the analysis of the reviewed studies, several research gaps have been identified:

Table 13: Research Gaps and Opportunities

Gap	Description	Opportunity
Limited focus on brute force attacks	Less attention compared to DoS	Improve detection methods
Dataset inconsistency	Lack of standard datasets	Develop unified benchmarks
Metric inconsistency	Different evaluation methods	Standardize evaluation
Lack of lightweight IDS	Heavy models dominate	Develop efficient models

Table 13 highlights the key research gaps identified from the systematic review. These gaps provide opportunities for future research, particularly in developing lightweight intrusion detection systems, improving dataset quality, and standardizing evaluation approaches.

4.6 Discussion of RQ4 Findings

The results show that while intrusion detection systems have advanced significantly, there are still a number of important issues that need to be addressed. One of the biggest obstacles among them is the absence of lightweight IDS solutions, especially for tiny home networks. Furthermore, the practical implementation of IDS models is still hampered by problems including computing complexity, high false positive rates, and dataset restrictions. These difficulties underscore the need for detection techniques that are more effective, scalable, and flexible. Crucially, the research gaps that have been discovered offer a clear path for further investigation. Specifically, a potential area of research is the creation of lightweight machine learning-based intrusion detection systems (IDS) models that can identify DoS and brute force attacks in tiny home networks.

SECTION 5: CONCLUSION

5.1 Summary of Findings

With an emphasis on intrusion attack types, datasets, assessment metrics, and open issues, this study offered a systematic literature review (SLR) on network intrusion detection systems. Fifteen recent, high-quality investigations (2023–2026) were chosen and examined using a structured PRISMA-based technique; the main baseline study was the Network Intrusion Detection Survey by Abdulkareem et al. (2024). The results show that the most often researched intrusion types are Denial of Service (DoS/DDoS) attacks, which are followed by brute force, botnet, and web-based attacks. Brute force attacks are still largely unexplored despite their importance in authentication-based systems, whereas denial-of-service (DoS) assaults predominate in the literature due to their extensive impact.

The most popular benchmarks in terms of datasets are CICIDS2017, CSE-CIC-IDS2018, and UNSW-NB15. Nevertheless, these datasets have problems like lack of real-world representation, repetition, and class imbalance. The ability of intrusion detection models to generalize is directly impacted by these constraints. The review also showed that IDS performance is frequently evaluated using evaluation measures like accuracy, precision, recall, and F1-score. But depending just on one indicator, notably accuracy, can result in false results, especially when dealing with unbalanced datasets.

5.2 Key Contributions

This study makes the following contributions:

- Provides a comprehensive and structured synthesis of IDS research based on attack types, datasets, evaluation metrics, and challenges
- Identifies and categorizes major intrusion attack types, with emphasis on DoS and brute force attacks
- Presents a detailed analysis of commonly used datasets and their limitations
- Highlights inconsistencies in evaluation metrics and performance assessment approaches
- Identifies critical research gaps, particularly the lack of lightweight IDS solutions for small-scale networks

5.3 Implications for Research and Practice

The findings of this study have important implications for both researchers and practitioners. For researchers, the identified gaps provide clear directions for future work, particularly in developing standardized evaluation frameworks and more realistic datasets. For practitioners, the study emphasizes the need for efficient and scalable IDS solutions that can be deployed in real-world environments, especially in resource-constrained settings such as small home networks.

5.4 Future Research Directions

Based on the identified gaps, the following future research directions are recommended:

- Development of lightweight intrusion detection systems suitable for small home and IoT networks
- Creation of standardized and realistic datasets that reflect modern network traffic
- Adoption of hybrid and ensemble machine learning approaches to improve detection accuracy
- Exploration of real-time and low-latency IDS solutions
- Standardization of evaluation metrics to enable fair comparison across studies

These directions directly support the motivation for the proposed thesis work on designing a network-based intrusion detection system using efficient machine learning techniques.

References

Abdulkareem, S. A., Foh, C. H., Shojafar, M., Carrez, F., & Moessner, K. (2024). Network Intrusion Detection: An IoT and Non IoT-Related Survey. In *IEEE Access* (Vol. 12, pp. 147167–147191). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2024.3473289

Afrin, S., Al Muttaki, M. R., Anil, A. I. A., & Hasan, S. (2026). AI-powered cybersecurity for smart grid communication: A systematic review of intrusion detection and threat mitigation systems. In *Energy Conversion and Management: X* (Vol. 29). Elsevier Ltd. https://doi.org/10.1016/j.ecmx.2025.101416

Aldhaheri, A., Alwahedi, F., Ferrag, M. A., & Battah, A. (2024). Deep learning for cyber threat detection in IoT networks: A review. In *Internet of Things and Cyber-Physical Systems* (Vol. 4, pp. 110–128). KeAi Communications Co. https://doi.org/10.1016/j.iotcps.2023.09.003

Berhili, M., Chaieb, O., & Benabdellah, M. (2024). Intrusion Detection Systems in IoT Based on Machine Learning: A state of the art. *Procedia Computer Science*, *251*, 99–107. https://doi.org/10.1016/j.procs.2024.11.089

Chinnasamy, R., Subramanian, M., Easwaramoorthy, S. V., & Cho, J. (2025). Deep learning-driven methods for network-based intrusion detection systems: A systematic review. In *ICT Express* (Vol. 11, Number 1, pp. 181–215). Korean Institute of Communications and Information Sciences. https://doi.org/10.1016/j.icte.2025.01.005

Hassan, S. R., Tanveer, M. U., Prajapat, S., & Shabaz, M. (2025). A comprehensive survey on intrusion detection in internet of medical things: Datasets, federated learning, blockchain, and future research directions. *ICT Express*, *11*(6), 1291–1310. https://doi.org/10.1016/j.icte.2025.11.005

Hozouri, A., Mirzaei, A., & Effatparvar, M. (2025). A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges. *Discover Artificial Intelligence*, *5*(1). https://doi.org/10.1007/s44163-025-00578-1

Khadam, U., Davidsson, P., & Spalazzese, R. (2025). A systematic literature review on AI in IoT systems: Tasks, applications, and deployment. In *Internet of Things (The Netherlands)* (Vol. 34). Elsevier B.V. https://doi.org/10.1016/j.iot.2025.101779

Liao, H., Murah, M. Z., Hasan, M. K., Aman, A. H. M., Fang, J., Hu, X., & Khan, A. U. R. (2024). A Survey of Deep Learning Technologies for Intrusion Detection in Internet of Things. *IEEE Access*, *12*, 4745–4761. https://doi.org/10.1109/ACCESS.2023.3349287

Mukhaini, G. AL, Anbar, M., Manickam, S., Al-Amiedy, T. A., & Momani, A. Al. (2024). A systematic literature review of recent lightweight detection approaches leveraging machine and deep learning mechanisms in Internet of Things networks. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 36, Number 1). King Saud bin Abdulaziz University. https://doi.org/10.1016/j.jksuci.2023.101866

Nazir, A., He, J., Zhu, N., Wajahat, A., Ma, X., Ullah, F., Qureshi, S., & Pathan, M. S. (2023). Advancing IoT security: A systematic review of machine learning approaches for the detection of IoT botnets. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 35, Number 10). King Saud bin Abdulaziz University. https://doi.org/10.1016/j.jksuci.2023.101820

Rahman, M. M., Shakil, S. Al, & Mustakim, M. R. (2025). A survey on intrusion detection system in IoT networks. In *Cyber Security and Applications* (Vol. 3). KeAi Communications Co. https://doi.org/10.1016/j.csa.2024.100082

Rawat, M., & Singal, G. (2025). Surveying Technology Fusion in IoT Networks for IDS: Exploring Datasets, Tools, Challenges, and Research Prospects. *ACM Transactions on Intelligent Systems and Technology*, *16*(5). https://doi.org/10.1145/3744745

Sharma, S., Kumar, V., & Dutta, K. (2024). Multi-objective optimization algorithms for intrusion detection in IoT networks: A systematic review. In *Internet of Things and Cyber-Physical Systems* (Vol. 4, pp. 258–267). KeAi Communications Co. https://doi.org/10.1016/j.iotcps.2024.01.003

Slaibi, T., Ivaki, N., & Vieira, M. (2026). IoT security assessment: A systematic literature review. In *Journal of Systems and Software* (Vol. 237). Elsevier Inc. https://doi.org/10.1016/j.jss.2026.112833