Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Deep Learning, Federated Learning, and Meta-Learning for Intrusion Detection in the Internet of Things: A Systematic Review, Taxonomy, and Research Agenda

Wilvens PIERRE LOUIS

piew0001@uqar.ca

Université du Québec à Rimouski

Mehdi Mehdi Adda

Université du Québec à Rimouski

Research Article

Keywords: Internet of Things, intrusion detection, deep learning, federated learning, meta-learning, few-shot learning

Posted Date: September 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10756458/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Deep Learning, Federated Learning, and Meta-Learning for Intrusion Detection in the Internet of Things: A Systematic Review, Taxonomy, and Research Agenda

Corresponding author: Wilvens Pierre Louis · Mehdi Adda

Received: date / Accepted: date

Abstract Operational Technology and Industrial Internet of Things environments expose connected systems to network attacks whose detection is challenging because of distributed data, class imbalance, and edge-computing constraints. This article presents a systematic review of Internet of Things intrusion detection systems based on deep learning, federated learning, and meta-learning. The review follows a process inspired by PRISMA 2020 and by the recommendations of Kitchenham and Charters for systematic reviews in software engineering. It analyzes references published in or indexed by IEEE Xplore, ScienceDirect/Elsevier, and Springer Nature over the harmonized period 2017–2026, complemented by methodological and algorithmic references required for formalization. The synthesis classifies the works along six axes: detection type, learning paradigm, architecture, federated strategy, adaptation mechanism, and evaluation protocol. The results show that CNN, LSTM/GRU, autoencoder, Transformer, and hybrid architectures are frequently used; that FedAvg remains a dominant baseline; and that FedProx, FedNova, SCAFFOLD, federated personalization, and few-shot approaches are used to mitigate part of the effects associated with non-IID data and rare attacks. The critical analysis reveals recurrent limitations: non-comparable protocols, incomplete treatment of non-IID types, limited availability of code and client partitions (22% of the assessed empirical studies provide code), rare reporting of edge measurements (26%), limited evaluation of explainability, and a high risk of publication bias due to the absence of a systematic search in the grey literature.

Keywords Internet of Things · intrusion detection · deep learning · federated learning · meta-learning · few-shot learning

1 Introduction

The Internet of Things (IoT) connects heterogeneous devices in domestic, industrial, medical, and urban environments [1]. This connectivity increases the attack surface: IoT objects often have limited resources, use heterogeneous protocols, and are deployed in environments that are difficult to administer [2, 3]. In critical infrastructures, an undetected intrusion can have major material and financial consequences.

Intrusion detection systems play a central role in this context. An IDS aims to detect malicious or abnormal behavior from network flows, system logs, telemetry, or features extracted from communications. Early reviews of IoT IDSs showed that classical methods, often designed for enterprise networks, are difficult to transfer to IoT because of mobility, heterogeneity, energy consumption, and the variety of attacks [4]. Similar difficulties had already been observed in wireless sensor networks, where security must be ensured under computation, memory, and energy constraints [5].

Wilvens Pierre Louis
E-mail: piew0001@uqar.ca
Present address: Department of Mathematics, Computer Science and Engineering, University of Quebec At Rimouski, 300 Allée des Ursulines, Rimouski, QC, G5L 3A1, Canada

Mehdi Adda
Tel.: +418 833-8800, poste 3197
E-mail: addame01@uqar.ca
Department of Mathematics, Computer Science and Engineering, University of Quebec At Rimouski, 300 Allée des Ursulines, Rimouski, QC, G5L 3A1, Canada

1.1 Learning-based IDSs: from ML to DL, FL, and MetaL

Machine learning has gradually replaced approaches based purely on signatures. Classical ML models, such as Random Forest, SVM, or XGBoost, remain strong baselines when the data are tabular. Deep learning then made it possible to learn representations automatically from sequential or multivariate flows. CNNs extract local patterns, LSTM and GRU networks model temporal dependencies, autoencoders detect anomalies through reconstruction, Transformers exploit attention to model long-range dependencies, and GNNs can represent relationships between devices or flows [6, 7].

However, centralized training of deep models often assumes that sensitive data can be gathered on a central server. This assumption is problematic in IoT and IIoT: data may be personal, medical, industrial, or strategic; transmitted volumes may be large; and distributions vary strongly across sites, devices, and time periods. Federated learning partly addresses these limitations by allowing each client to keep its data locally while participating in the training of a global model [8, 9]. Despite this advantage, FL alone is not sufficient: IoT clients are often non-IID, rare attack classes are poorly represented, some clients may be compromised, and edge constraints limit model size and communication frequency [10–12].

1.2 Motivation and guiding thread

This review examines a precise question: to what extent do IoT IDSs based on DL and FL document the effects of non-IID data, rare attacks, and edge constraints, and which adaptation solutions are actually evaluated? Non-IID data can produce client drift that reduces the stability of global aggregation; rare or zero-day attacks require evaluation protocols that are sensitive to minority classes; and edge devices impose trade-offs between accuracy, memory, latency, and communication. Recent work analyzes FL for IoT IDSs [13–15] and evaluates, in some scenarios, personalization or meta-learning when clients and attacks are heterogeneous [16–19].

1.3 Contributions

The contributions of this review are as follows:

1. to propose a systematic review inspired by PRISMA 2020 and by the recommendations of Kitchenham and Charters [20, 21];
2. to build a unified taxonomy that clearly separates detection type, learning paradigm, architecture, federated strategy, adaptation mechanism, and evaluation protocol;
3. to analyze specifically the role of non-IID data, client drift, and rare attacks in federated IoT IDSs, with an exhaustive classification of heterogeneity types;
4. to synthesize few-shot, MAML, meta-federated, and personalized approaches as adaptation solutions, and to evaluate their connection with the MAML formalism;
5. to assess methodological limitations in existing studies, including data leakage, reproducibility (22% public code), inadequate metrics, and the absence of realistic edge tests (26% report edge measurements);
6. to propose a research agenda centered on personalized, robust, explainable, and edge-deployable IoT IDSs.

1.4 Positioning with respect to existing reviews

Several reviews have already investigated IoT IDSs, federated learning, deep learning, or the security of distributed systems. However, these studies rarely cover the following dimensions jointly: deep learning, federated learning, non-IID data, meta-learning, edge constraints, explainability, a PRISMA-based process, and recent datasets. Table 1 positions this review against existing works and highlights its original contribution.

Table 1: Positioning of this review with respect to existing reviews

Review	DL	FL	Non-IID	MetaL	Edge	XAI	PRISMA	Recent datasets
Zarpelão et al. [4]	Partial	No	No	No	Partial	No	No	No
Mothukuri et al. [10]	Partial	Yes	Partial	No	No	No	No	No
Campos et al. [11]	Partial	Yes	Partial	No	Partial	No	No	Partial
Belenguer et al. [12]	Partial	Yes	Partial	No	Partial	Partial	Partial	Yes
Zhang et al. [22]	Partial	Yes	Partial	No	Partial	Partial	Partial	Yes
Buyuktanir et al. [23]	Partial	Yes	Partial	No	Partial	Partial	Partial	Yes
Our review	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

DL: deep learning; FL: federated learning; Non-IID: statistical heterogeneity and client drift. MetaL: meta-learning, few-shot learning, and federated personalization; Edge: computation, memory, latency, and communication constraints; XAI: explainability and interpretability; Recent datasets: consideration of datasets such as TON_IoT, CICIoT2023, Edge-IIoTset, or recent work from 2023–2026.

Unlike existing reviews, this study does not treat IoT IDSs, FL, or DL separately. It proposes an integrated reading centered on a precise problem: the fragility of federated IoT IDSs when facing non-IID data, rare attacks, and edge constraints. Its main contribution is therefore to connect federated learning, meta-learning, personalization, explainability, and recent evaluation protocols in a single critical taxonomy.

1.5 Research questions

Table 2: Research questions guiding the review

Code	Research question
RQ1	Which deep learning models dominate IoT and IIoT IDSs?
RQ2	How is federated learning applied to IoT IDSs?
RQ3	How do studies handle non-IID heterogeneity, client drift, and class imbalance?
RQ4	What role does meta-learning play in adaptation to rare or emerging attacks?
RQ5	What are the current methodological limitations: data leakage, reproducibility, metrics, and comparison?
RQ6	What challenges remain open for real edge deployment: communication, memory, latency, robustness, and explainability?

2 Conceptual Framework and Theoretical Foundations

2.1 Internet of Things and IoT security

An IoT architecture can be described in several layers: perception, network, edge/fog, and application. The perception layer includes sensors, actuators, and embedded devices. The network layer carries flows through Wi-Fi, Bluetooth, Zigbee, 5G, industrial Ethernet, or IoT protocols. Edge and fog layers perform preprocessing close to the sources, while the cloud provides large-scale computation and storage. Attacks may target each layer: sensor capture, spoofing, DDoS, botnets, malware, routing attacks, exfiltration, or gateway compromise [3, 4].

2.2 Intrusion detection systems

IDSs can be grouped into three families. Signature-based IDSs detect known patterns; they are fast and interpretable but less effective against unknown attacks. Anomaly-based IDSs learn normal behavior or decision boundaries; they are better suited to novel attacks but may increase false positives. Hybrid IDSs combine signatures, rules, ML, DL, and sometimes XAI to improve coverage and operational trust [4].

2.3 Deep learning for IDSs

CNNs are effective for extracting local patterns from flow vectors or temporal windows. LSTM and GRU networks model the temporal order of events. Autoencoders learn a compact representation of normal behavior and detect anomalies through reconstruction error. Transformers use self-attention to model long-range dependencies; they can model multivariate flows but remain costly for edge deployment [6, 7]. GNNs and Graph-Transformers are relevant when network structure or relationships between devices can be modeled.

2.4 Federated learning

In FL, each client k owns a local dataset D_k and trains a local model without transmitting its raw data. The global objective is:

$$\min_w F(w) = \sum_{k=1}^{K} \frac{n_k}{N} F_k(w), \quad N = \sum_{k=1}^{K} n_k, \tag{1}$$

where F_k is the local loss and n_k is the size of the client's dataset. FedAvg aggregates local weights using a weighted average [8]. FedProx adds a proximal term to reduce local drift in heterogeneous settings [24]. FedNova normalizes contributions when clients perform different numbers of local steps [25]. SCAFFOLD uses control variates to reduce client drift [26]. In IoT IDSs, these mechanisms are essential because clients do not have the same attacks, volumes, or distributions [11, 15].

2.5 Meta-learning and few-shot learning

Meta-learning aims to learn how to learn quickly. In MAML, the model learns an initialization θ that can be adapted to a new task $\mathcal{T}_i$ with a few gradient steps:

$$\theta_i' = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(\theta), \tag{2}$$

and the initialization is then optimized over several tasks:

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_i \mathcal{L}_{\mathcal{T}_i}(\theta_i'). \tag{3}$$

The main families are optimization-based methods such as MAML, metric-based methods such as prototypical networks, and memory- or attention-based methods. In an IoT IDS, a task $\mathcal{T}_i$ may be defined by a device, a site, a protocol, an attack family, or a collection period. The initialization θ therefore does not represent only an abstract set of parameters: it corresponds to a global model that must be adapted locally under bandwidth, memory, communication-frequency, and label-availability constraints. In this review, Equations (2)–(3) are not treated as a simple theoretical reminder. They operationalize a verifiable methodological distinction between: (i) studies that explicitly define a task or client $\mathcal{T}_i$ and perform few-shot or meta-federated local adaptation; (ii) studies that learn a global federated model without a local adaptation step comparable to θ_i'; and (iii) studies that mention adaptation without documenting computational cost, number of local steps, communication, or label availability. Table 12 operationalizes this distinction.

2.6 Formalization of non-IID data

Non-IID means that two clients i and j do not necessarily follow the same distribution:

$$P_i(x, y) \neq P_j(x, y). \tag{4}$$

In IoT, this phenomenon takes several forms: *label skew*, when clients observe different classes; *feature skew*, when the same classes have different characteristics; *quantity skew*, when local dataset sizes vary; *concept drift*, when the relationship between features and labels changes over time; and *device skew*, when the distribution depends on the device or gateway. IoT IDSs should therefore be evaluated on realistic partitions, not only on random IID splits.

3 Systematic Review Methodology

3.1 Methodological framework

The review follows a process inspired by PRISMA 2020 by documenting identification, screening, eligibility, and inclusion [20]. The recommendations of Kitchenham and Charters are also used to structure the research questions, inclusion criteria, and extraction grid [21]. Methodological quality is then assessed using AMSTAR 2 as the main recognized instrument.

3.1.1 Scientific justification

This review focuses on studies that document at least one of the following dimensions: deep learning for detection, federated learning, statistical heterogeneity, few-shot adaptation or meta-learning, federated personalization, edge/communication cost, and methodological quality.

The review therefore excludes general bibliometric elements that do not directly contribute to the IoT/IIoT IDS problem. The tables and syntheses that follow are centered on the search protocol, eligibility criteria, risk of bias, non-IID classification, and reproducibility indicators.

3.2 Documentary sources

The main sources selected are IEEE Xplore, ScienceDirect/Elsevier, and Springer Nature. They cover the major journals and conferences used in this article: IEEE Internet of Things Journal, IEEE Transactions on Industrial Informatics, IEEE Access, Computer Networks, Future Generation Computer Systems, Computers & Security, Expert Systems with Applications, Journal of Parallel and Distributed Computing, International Journal of Information Security, Cluster Computing, and Discover Internet of Things. Scopus and Semantic Scholar may be used as complementary indexing sources to verify metadata and identify duplicates, but the reference table distinguishes the primary publication source.

3.3 Search strategy

The search strategy is structured into five conceptual blocks: IoT/IIoT, IDS, federated learning, meta-learning/few-shot/personalization, and deep learning. The string below corresponds to the canonical query; it was then translated syntactically in each database in order to respect its operators, fields, and filters.

(“Internet of Things” OR IoT OR IIoT OR “Industrial Internet of Things”) AND (“intrusion detection” OR IDS OR “anomaly detection” OR “malware detection” OR DDoS) AND (“federated learning” OR “federated deep learning” OR FedAvg OR FedProx OR FedNova OR SCAFFOLD OR “personalized federated learning”) AND (“meta-learning” OR MAML OR “few-shot learning” OR “federated meta-learning” OR “rapid adaptation”) AND (“deep learning” OR CNN OR LSTM OR GRU OR Transformer OR attention OR autoencoder OR GNN).

The three main databases were filtered symmetrically over the period 2017–2026. Document-type filters were adapted to the options available in each database: research articles and review articles for ScienceDirect; journals and conferences for IEEE Xplore; articles, research articles, review articles, and conference papers for Springer Nature. Complementary targeted queries focused on “federated intrusion detection IoT”, “meta-learning IDS IoT”, “non-IID federated malware detection”, and “few-shot IoT intrusion detection”.

Table 3: Optimized search strategies for the systematic review

Database	Optimized search strategy
ScienceDirect	Search 1: (“federated learning” AND “intrusion detection” AND IoT) OR (“federated learning” AND “anomaly detection” AND IoT) OR (“federated learning” AND “malware detection” AND IoT). Search 2: (“meta-learning” AND “intrusion detection” AND IoT) OR (“few-shot learning” AND “intrusion detection” AND IoT) OR (MAML AND “intrusion detection”). Filters: research articles, review articles, 2017–2026, Computer Science, Engineering.
IEEE Xplore	((“Internet of Things” OR IoT OR IIoT OR “Industrial Internet of Things” OR AIoT) AND (“intrusion detection” OR IDS OR “anomaly detection” OR “malware detection” OR DDoS) AND (“federated learning” OR “federated deep learning” OR FedAvg OR FedProx OR FedNova OR SCAFFOLD OR “personalized federated learning” OR “federated meta-learning” OR “few-shot learning” OR MAML OR “meta-learning”)). Filters: journals, conferences, 2017–2026.
Springer Nature	(“federated learning” AND “intrusion detection” AND IoT) OR (“federated learning” AND “anomaly detection” AND “Internet of Things”) OR (“meta-learning” AND “intrusion detection” AND IoT) OR (“few-shot learning” AND “IoT intrusion detection”) OR (“personalized federated learning” AND “intrusion detection”). Filters: article, research article, review article, conference paper, English, 2017–2026; subject areas: Machine Learning, Internet of Things, Artificial Intelligence, Security Science and Technology.

3.4 Inclusion and exclusion criteria

Table 4: Inclusion and exclusion criteria

Type	Criteria
Inclusion	Peer-reviewed studies; IoT, IIoT, edge/fog, or distributed-network context; intrusion, anomaly, botnet, or malware detection task; use of ML, DL, FL, meta-learning, few-shot learning, personalization, or attention; reported experimental results.
Exclusion	Studies outside IoT/IIoT; purely conceptual papers; absence of evaluation; insufficient dataset description; duplicates; preliminary versions replaced by a final version; full text unavailable.

3.5 Selection process

The study selection process followed the stages recommended by PRISMA 2020: identification, screening, eligibility, and inclusion. References from IEEE Xplore, ScienceDirect/Elsevier, and Springer Nature were first exported and imported into Zotero to create a unified bibliographic database and remove duplicates.

The identification phase retrieved 6536 references from the three main sources. Ten additional references, derived from expert knowledge, were added to include foundational work needed for methodological and theoretical formalization, particularly on PRISMA, federated learning, meta-learning, and explainability metrics. After removing 258 duplicates, 6288 references were retained for screening.

Screening was carried out independently by two reviewers. Each reviewer examined titles and abstracts to assess relevance against the inclusion and exclusion criteria defined in Table 4. Studies clearly outside the scope of IoT/IIoT, intrusion detection, deep learning, federated learning, or meta-learning were excluded at this stage. In total, 6186 references were excluded after title and abstract screening.

The remaining 102 full texts were then assessed for eligibility by the two reviewers. This stage verified the presence of experimental validation, dataset description, methodological clarity, type of approach, and relevance to the review’s guiding thread, namely the analysis of federated IoT IDSs facing non-IID data, rare attacks, and edge constraints.

Disagreements between the two reviewers were resolved through discussion and joint rereading of the concerned articles. When a divergence remained, the final decision was made by consensus after re-examining the inclusion and exclusion criteria. This procedure helped limit selection bias and ensure consistent application of the review protocol.

After full-text reading, 58 texts were excluded. The main exclusion reasons were: study outside the IoT/IIoT context, absence of an IDS or anomaly-detection task, incomplete experimental protocol, absence of deep learning, federated learning, or meta-learning, insufficiently detailed results, or inability to clearly connect the study to the research questions. Finally, 44 unique studies were included in the qualitative synthesis.

Table 5: Results of the document search

Source	**Identified**	**Candidates**	**Applied filters**
Springer Nature	1604	27	Article, review, conference paper, 2017–2026, English
IEEE Xplore	2346	35	Journals, conferences, 2017–2026
ScienceDirect/Elsevier	2586	40	Research article, review article, 2017–2026
Total main databases	**6536**	**102**	Before deduplication
Expert knowledge	10	10	Methodological references
Total before deduplication	**6546**	–	–
Unique studies included	–	**44**	After full-text reading

Candidate articles correspond to references judged potentially relevant after an initial reading of titles and abstracts in each source. These numbers are not added to obtain the final corpus because several references may appear in multiple databases. The final corpus includes 44 unique studies after duplicate removal and full-text reading.

The study selection process follows the PRISMA 2020 stages: identification, screening, eligibility, and inclusion. Figure 1 presents the complete selection flow.

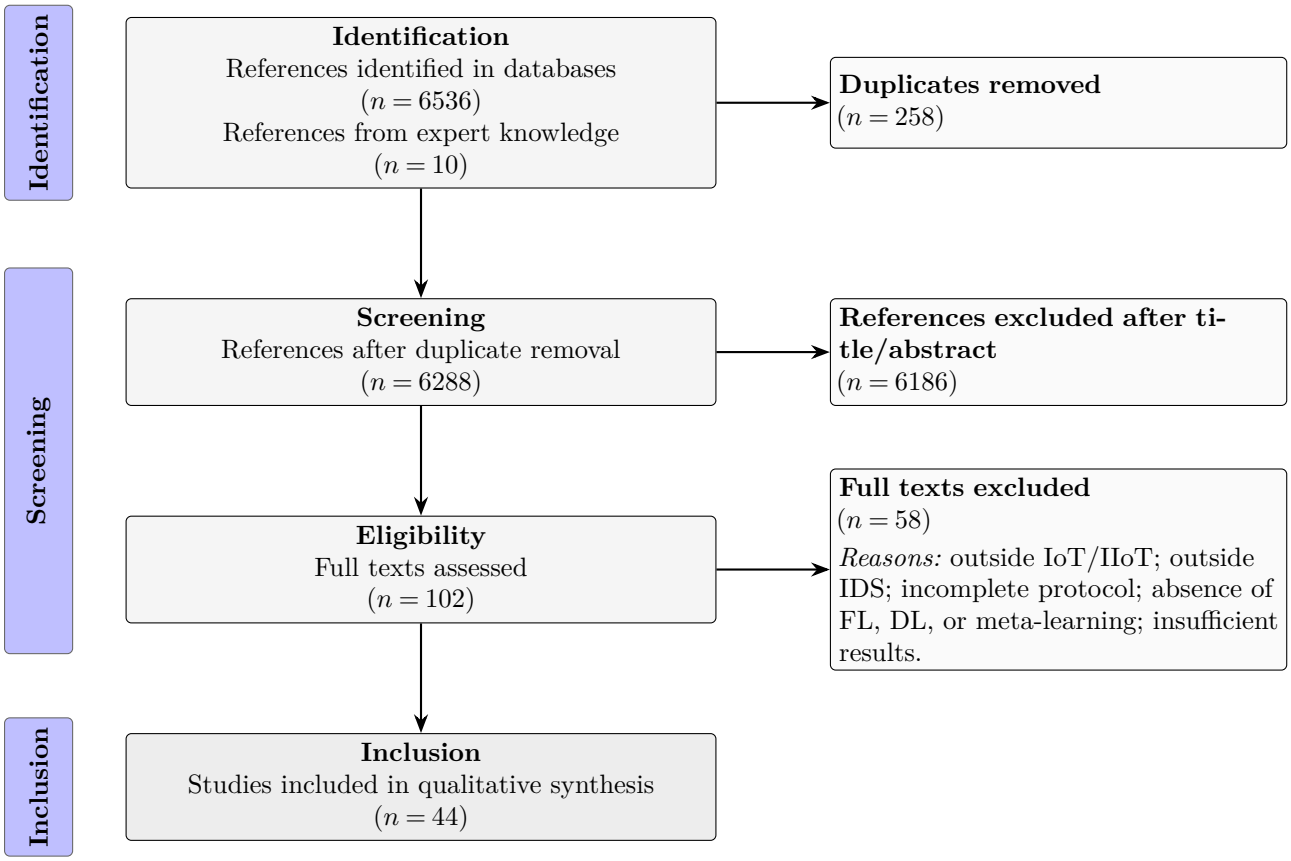


Fig. 1: PRISMA flow diagram of the study selection process.

The 44 cited references were grouped according to their main publication source. Most came from ScienceDirect/Elsevier with 16 articles, followed by IEEE Xplore with 14 articles and Springer Nature with 4 articles. These three sources represent 34 references, the central part of the corpus used to construct the review. The remaining 10 references come from complementary sources, including PMLR, NeurIPS, ACM, MDPI, PLOS ONE, BMJ, SCITEPRESS, MLSys, Now Publishers, and methodological reports.

3.6 Methodological quality assessment according to AMSTAR 2

To avoid the use of a non-validated assessment tool, the methodological quality of this review is appraised using the AMSTAR 2 checklist [27]. AMSTAR 2 is a recognized instrument for the critical appraisal of systematic reviews. It includes 16 items covering the definition of research questions, the existence of an a priori protocol, justification of included study types, search strategy, study selection and data extraction, justification of exclusions, description of included studies, risk-of-bias assessment, consideration of heterogeneity, publication bias, and conflicts of interest.

In this review, AMSTAR 2 is not replaced by a customized scoring grid. The 16 items are retained as the main assessment structure. Responses are coded using four categories: “Yes”, “Partial”, “No”, and “N.A.” when the item is not applicable. Items 2, 4, 7, 9, 11, 13, and 15 are considered critical according to AMSTAR 2. Because this review provides a qualitative and taxonomic synthesis, without statistical meta-analysis, Items 11 and 12 related to meta-analysis methods are considered not applicable. This non-applicability is explicitly stated to avoid methodological ambiguity.

The objective of this assessment is not to produce a global numerical score, but to identify the methodological strengths and weaknesses of the review. Results are interpreted qualitatively by distinguishing critical from non-critical weaknesses. This approach strengthens methodological transparency while avoiding the arbitrary addition of points.

Table 6: Methodological assessment of the review according to AMSTAR 2

Item	AMSTAR 2 criterion	Critical	Response	Justification in this review
1	Clearly defined research question and inclusion criteria	No	Yes	The research questions and inclusion/exclusion criteria are explicitly formulated. The PICO framework is transposed to the computer-science context: population = IoT/IIoT environments; intervention = DL, FL, and MetaL methods; comparators = centralized, federated, or classical approaches; outcomes = IDS metrics, robustness, reproducibility, and edge constraints.
2	Protocol established before the review and justified deviations	Yes	Partial	The selection protocol is described in the article, but no external prior registration was carried out. Methodological choices are therefore documented in the manuscript rather than in a preregistered protocol.
3	Justification of the study designs included	No	Yes	The included study types are justified by their connection with IoT/IIoT IDSs, deep learning, federated learning, meta-learning, personalization, edge constraints, or methodological foundations required by the review.

Item	AMSTAR 2 criterion	Critical	Response	Justification in this review
4	Comprehensive and documented search strategy	Yes	Yes	The main databases, search strings, harmonized 2017–2026 time filters, document types, and complementary queries are documented.
5	Study selection performed in duplicate	No	Yes	Title/abstract screening and full-text assessment were performed by two reviewers, with disagreements resolved through discussion and consensus.
6	Data extraction performed in duplicate	No	Partial	Extracted variables are described and divergences were checked through joint rereading. However, independent duplicate extraction is not documented in as much detail as study selection.
7	List of excluded studies with justification	Yes	Yes	The 58 full texts excluded after full-text reading are listed with an exclusion reason linked to the inclusion and exclusion criteria.
8	Detailed description of included studies	No	Yes	Included studies are described according to source, context, approach, learning paradigm, non-IID treatment, metrics, reproducibility, and relevance to the research questions.
9	Satisfactory technique for assessing risk of bias in included studies	Yes	Partial	Risk of bias is discussed using methodological dimensions relevant to IoT IDS studies: dataset description, experimental protocol, partitions, metrics, reproducibility, code availability, non-IID scenarios, and edge constraints. However, no single specialized tool fully validated for IoT IDS/FL studies was applied to primary studies.
10	Sources of funding for included studies reported	No	Partial	Funding sources are not systematically available or reported in all included studies; this limitation is declared.
11	Appropriate methods for meta-analysis	Yes	N.A.	No statistical meta-analysis is performed because of heterogeneity in datasets, metrics, models, partitions, and experimental protocols. The review relies on a qualitative and taxonomic synthesis.
12	Impact of risk of bias on the meta-analysis	No	N.A.	Not applicable, because no statistical meta-analysis is conducted.

Item	AMSTAR 2 criterion	Critical	Response	Justification in this review
13	Risk of bias considered when interpreting results	Yes	Yes	Conclusions take methodological limitations into account: non-comparable protocols, weak reproducibility, missing non-IID client partitions, incomplete metrics, limited code availability, and limited edge validation.
14	Satisfactory explanation of heterogeneity	No	Yes	Heterogeneity is discussed in terms of dataset, architecture, experimental protocol, non-IID distribution, metrics, IoT/IIoT context, and edge constraints.
15	Investigation of publication bias	Yes	Partial	The risk of publication bias is recognized as high, especially because there was no systematic search of grey literature, theses, preprints, or project deliverables. No funnel plot or Egger test is used because the review does not perform a quantitative meta-analysis.
16	Declaration of conflicts of interest	No	Yes	The review authors' conflicts of interest are declared in the dedicated section.

The application of AMSTAR 2 highlights several methodological strengths: explicit research questions, a documented search strategy, duplicate study selection, justification of exclusions, description of included studies, and consideration of heterogeneity. The main weaknesses concern the absence of a preregistered protocol, data extraction only partially documented in duplicate, incomplete information on funding sources in primary studies, and the absence of a systematic search in the grey literature. These limitations are taken into account in the final interpretation of the results.

Risk-of-bias assessment of included studies. AMSTAR 2 evaluates the methodological quality of the systematic review. For the included studies, risk of bias is appraised qualitatively from the information available in each article: clarity of the dataset, description of the split, experimental protocol, metrics, baselines, consideration of non-IID data, reproducibility, code availability, communication cost, and edge constraints. This appraisal is not presented as an autonomous AMSTAR 2 tool applied to primary studies, but as a complementary critical synthesis required to interpret evidence in IoT/IIoT IDS research.

Foundational, methodological, or contextual references, such as articles on PRISMA, FedAvg, MAML, datasets, or explainability, are not assessed as empirical IoT IDS studies. They are marked as not applicable in the synthesis tables.

Overall, studies directly related to FL for IoT IDSs, non-IID data, and meta-learning show low to moderate risk of bias, with one provisional high-risk exception when the information available before submission does not yet allow sufficient verification of metadata, protocol, and reproducibility. Moderate risks mainly stem from partial treatment of non-IID data, limited reproducibility, missing edge measurements, or incomplete communication-cost analysis. This conclusion concerns the risk of bias of individual studies; it is distinct from publication bias at the review level, which remains high because there was no systematic search in the grey literature.

Table 7: Qualitative appraisal of risk of bias in included studies

Study	**Methodological criteria considered**	**Risk**	**Synthetic justification**
Zarpelão et al. (2017)	Protocol, IoT IDS scope, FL/Meta-FL currency	Moderate	Structuring review on IoT IDSs, but prior to recent work on FL, non-IID data, and meta-learning.
Abduvaliyev et al. (2013)	Protocol, WSN/IDS context, IoT transferability	Moderate	Useful reference on IDSs in sensor networks, but older and indirect for federated IoT IDSs.
Chen et al. (2022)	Dataset, metrics, architecture, reproducibility	Moderate	Relevant study on Transformer-based IoT anomaly detection, but not centered on FL or personalization.
Mothukuri et al. (2021)	FL scope, security, IoT IDS applicability	Moderate	Detailed review on FL security and privacy, but not exclusively oriented toward IoT IDSs.
Campos et al. (2022)	Dataset, FL protocol, non-IID data, metrics	Low	Study directly related to FL for IoT IDSs, with a clear discussion of challenges and non-IID data.
Belenguer et al. (2025)	Taxonomy, protocol, methodological limitations	Low	Recent review focused on FL applications to IDSs, with taxonomy and methodological limitations.
Hamdi (2023)	Dataset, FL protocol, reproducibility, edge	Moderate	Direct study on FL for IoT IDSs, but edge cost and reproducibility are only partially documented.
Mahmoodi et al. (2023)	Dataset, distributed protocol, non-IID data, reproducibility	Moderate	FL-based distributed IDS, but some details on non-IID data, cost, and reproducibility remain limited.
Djaidja et al. (2024)	Experimental protocol, FL scenario, metrics	Low	Recent experimental study on FL for 5G/B5G IDSs, with analysis of FL benefits and limits.
Lu et al. (2023)	Few-shot, MAML, adaptation protocol, metrics	Low	Central study on MAML and few-shot IoT IDSs, relevant to rare attacks.
Han et al. (2024)	Non-IID data, meta-sampling, FL protocol, metrics	Low	Study directly centered on non-IID data and meta-sampling for federated IDSs.
Yan et al. (2025)	Federated personalization, Meta-FL protocol, metrics	Low	Central study on federated personalization and meta-learning for IoT IDSs.
Xu et al. (2025)	Few-shot, attention, convolution, validation	Moderate	Relevant study on few-shot IDS with attention and convolution, but not centered on FL.

Study	Methodological criteria considered	Risk	Synthetic justification
Li et al. (2021)	FL protocol, DL architecture, edge, reproducibility	Moderate	DeepFed applies deep federated learning to industrial systems, but edge measurements remain limited.
Idrissi et al. (2023)	Autoencoder, FL protocol, metrics, non-IID data	Low	Fed-ANIDS directly addresses federated anomaly detection with a clear experimental methodology.
Jin et al. (2024)	Incremental FL, emerging attacks, reproducibility	Moderate	Incremental FL is relevant to emerging attacks, but comparability and reproducibility need strengthening.
Jin et al. (2025)	Aggregation, heterogeneity, FL protocol, metrics	Low	Recent study on efficient aggregation for heterogeneous data, directly related to non-IID data.
Ruzafa-Alcázar et al. (2023)	Privacy, IIoT protocol, metrics, robustness	Low	Relevant study on privacy-preserving FL for IIoT, with a well-aligned security scenario.
Puviarasu and Sudha (2026)	Metadata, protocol, partitions, reproducibility	Provisional high	Very recent study on FL and privacy; metadata, experimental protocol, client partitions, and reproducibility elements must be checked before using it as strong evidence.
Lee and Lee (2025)	Lightweight FL, IIoT validation, reproducibility, edge	Moderate	Relevant study on lightweight FL for IIoT, but reproducibility and edge validation details require strengthening.
Fraihat et al. (2026)	Hybrid DL, feature optimization, validation	Moderate	Recent study on hybrid DL and feature optimization for IIoT, but not centered on FL or Meta-FL.
Zhang et al. (2025)	FL-IDS review, taxonomy, limitations	Low	Recent review on FL in IDSs, useful for synthesizing architectures, challenges, and limitations.
Buyuktanir et al. (2025)	FL-IDS review, synthesis, perspectives	Low	Recent review on advances and perspectives of FL for IDSs, relevant to the research agenda.

References not assessed in this qualitative appraisal: Atzori et al. (2010), Zanella et al. (2014), Sicari et al. (2015), Vaswani et al. (2017), McMahan et al. (2017), Kairouz et al. (2021), Page et al. (2021), Shea et al. (2017), Kitchenham and Charters (2007), Li et al. (2020), Wang et al. (2020), Karimireddy et al. (2020), Moustafa and Slay (2015), Koroniotis et al. (2019), Alsaedi et al. (2020), Sharafaldin et al. (2018), Neto et al. (2023), Finn et al. (2017), Hospedales et al. (2022), Arrieta et al. (2020), Ribeiro et al. (2016), Lundberg and Lee (2017). These references are used as theoretical, methodological, algorithmic, or dataset foundations, and not as empirical IoT IDS studies to be assessed.

Synthesis of risk of bias in included studies. The risk-of-bias assessment shows that the methodological quality of the included studies is variable. Foundational, methodological, or dataset-related references were not assessed as empirical IDS studies because they do not directly propose an IoT/IIoT IDS experimental protocol. They were therefore considered not applicable to this appraisal.

Among the empirical studies analyzed, some show low risk of bias, particularly studies that clearly describe their dataset, experimental protocol, metrics, and federated scenario. These studies are especially useful for supporting conclusions related to federated learning, non-IID data, meta-learning, and personalization. Studies classified as moderate risk remain relevant but have limitations such as partial description

of the train/test split, limited treatment of non-IID data, missing communication-cost measurement, or incomplete reproducibility.

One very recent study, Puviarasu and Sudha (2026), is classified as provisional high risk, not because its contribution is invalid, but because metadata, client partitions, experimental protocol, and reproducibility elements must be verified before final submission. The other central empirical studies are classified as low or moderate risk. Several methodological limitations remain frequent: dependence on a single dataset, dominant use of accuracy, absence of external validation, weak description of client partitions, lack of edge measurements, and limited code availability. These limitations reduce direct comparability of results and justify the choice of a qualitative synthesis rather than a statistical meta-analysis.

Overall, confidence is higher for general conclusions about the value of federated learning and deep architectures in IoT IDSs. It is more moderate for conclusions related to non-IID data, robustness, and communication, and more limited for meta-learning, federated personalization, explainability, and real deployment on edge devices. Details of the study-level appraisal and non-applicable methodological references are presented in Table 20.

3.7 Data extraction

For each study, the following variables were extracted: authors, year, source, IoT/IIoT context, dataset, attack type, architecture, FL aggregator, definition of non-IID data, number of clients, metrics, robustness, communication cost, edge constraints, and code availability. Data extraction was carried out by two reviewers using a structured grid. Divergences or ambiguous information were checked through joint rereading of the concerned articles and then harmonized by consensus.

3.8 Publication-bias assessment

Publication bias was assessed qualitatively because the included studies show strong heterogeneity in datasets, experimental protocols, metrics, architectures, and federated scenarios. Consequently, a formal statistical analysis of publication bias, such as a funnel plot or Egger's test, was not considered appropriate because no quantitative meta-analysis was performed.

The search in IEEE Xplore, ScienceDirect/Elsevier, and Springer Nature reduces the risk of missing part of the literature published in indexed journals and conferences. However, this argument is not sufficient to conclude that the risk is moderate. This review did not conduct a systematic search in the grey literature, including preprints, theses, technical reports, project deliverables, or institutional repositories. These sources may contain negative results, experimental failures, unpublished evaluations, or protocols less favorable to the proposed methods.

Consequently, the risk of publication bias is considered **high**. This appraisal does not invalidate the qualitative synthesis, but it requires cautious conclusions, especially for effects reported as positive on performance, robustness, few-shot learning, or federated personalization.

3.9 Assessment of confidence in the evidence

The overall confidence in the evidence was assessed from the AMSTAR 2 methodological appraisal presented in Table 6, the qualitative appraisal of risk of bias in the included studies presented in Table 7, and the consistency of results by theme. This appraisal is not an arithmetic score; it considers the methodological quality of the review, the transparency of experimental protocols, the diversity of datasets, the treatment of non-IID data, reported metrics, baselines, robustness, edge/communication cost, and reproducibility.

Three confidence levels were used. A high level indicates that studies in a theme sufficiently describe their data, partitions, metrics, baselines, and experimental parameters. A moderate level indicates that results are usable, but that one or more critical domains are incomplete, for example reproducibility, non-IID analysis, or communication-cost evaluation. A low level indicates that evidence remains limited by a small number of studies, incomplete experimental protocols, weak preprocessing transparency, or evaluation far from real IoT/IIoT conditions.

The level of evidence was then synthesized by theme in order to avoid overinterpreting individual results. This synthesis considers study quality, consistency of results across works, diversity of datasets, and proximity of experimental protocols to real IoT/IIoT deployment conditions. Because of the strong

heterogeneity of datasets, non-IID partitions, metrics, and architectures, this assessment should be interpreted as a qualitative appraisal of confidence, not as a statistical meta-analysis.

Table 8: Overall confidence level of evidence by theme

Theme	**Evidence level**	**Justification**
Deep Learning for IoT IDSs	Moderate to high	CNN, LSTM/GRU, autoencoder, Transformer, and hybrid architectures are widely studied. However, performance remains difficult to compare because of differences in datasets, splits, and metrics.
Federated Learning for IoT IDSs	Moderate	Studies show the value of FL for privacy and distributed collaboration, but results strongly depend on the degree of non-IID data, number of clients, aggregator used, and experimental protocol.
Handling non-IID data and client drift	Moderate	Non-IID data are recognized as a central factor of performance degradation. However, partition strategies vary strongly across studies, which limits comparability.
Meta-learning, few-shot learning, and federated personalization	Low to moderate	Recent results are still preliminary for rare attacks and rapid adaptation; the number of studies is limited and large-scale validation in real IoT conditions remains insufficient.
FL robustness and security	Moderate	Several works discuss poisoning, malicious clients, differential privacy, or robust aggregation. However, few IoT IDS studies systematically evaluate these threats under realistic scenarios.
Edge deployment, communication, and TinyML	Low	Communication cost, memory, latency, and energy are still rarely reported. Evidence remains limited regarding real feasibility on constrained IoT devices.
Explainability of federated IoT IDSs	Low	XAI methods are recognized as important, but they are rarely integrated into federated IoT IDSs and even less often evaluated operationally.

4 Proposed Taxonomy

4.1 Global view

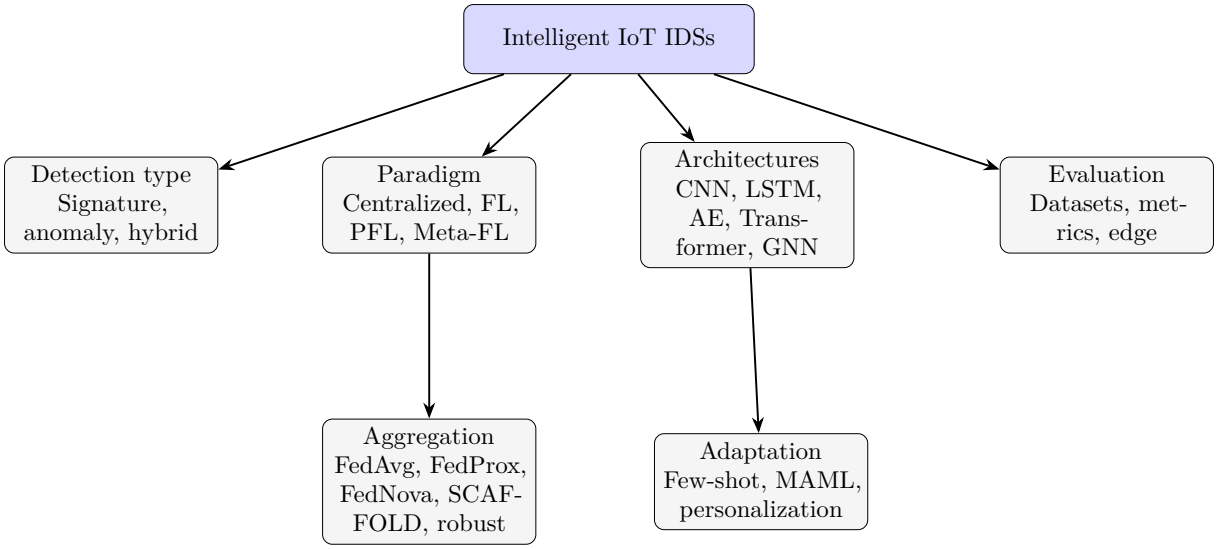


Fig. 2: Proposed taxonomy of IoT IDSs based on DL, FL, and meta-learning.

4.2 Taxonomy axes

Table 9: Axes of the proposed taxonomy

Axis	**Categories**	**Scientific role**
Detection type	Signature, anomaly, hybrid	Distinguish IDSs oriented toward known attacks from IDSs able to detect unknown behavior.
Learning paradigm	Centralized, FL, PFL, Meta-FL	Link privacy, collaboration, personalization, and adaptation.
Architectures	CNN, LSTM/GRU, autoencoder, Transformer, GNN	Compare representation capacity, cost, and edge relevance.
Aggregation	FedAvg, FedProx, FedNova, SCAFFOLD, robust	Measure resistance to non-IID data, client drift, and malicious updates.
Adaptation	Fine-tuning, MAML, few-shot, personalization	Address rare attacks, new devices, and domain changes.
Evaluation	Datasets, metrics, robustness, edge	Ensure comparability, reproducibility, and operational validity.

5 Datasets and Experimental Protocols

5.1 Classical and recent datasets

Datasets strongly influence conclusions. UNSW-NB15 and CICIDS2017/CICIDS2018 are used as general network baselines; Bot-IoT and TON_IoT are closer to IoT/IIoT scenarios; CICIoT2023 provides a recent large-scale basis; Edge-IIoTset complements edge/IIoT scenarios. However, datasets may contain imbalances, artifacts, or redundancies that artificially favor certain models [28–32].

Table 10: Main datasets used in IoT and IIoT IDSs

Dataset	Domain	Strengths	Limitations
UNSW-NB15	General network	Varied attacks, documented features	Not specifically IoT; risk of aging.
Bot-IoT	IoT/botnet	Botnet, DDoS, and scanning scenarios	Very strong imbalance; results sometimes overestimated.
TON_IoT	IoT/IIoT	Network data, telemetry, and logs	Heterogeneous preprocessing across studies.
CICIDS2017/2018	General network	Widely used for IDS comparison	Possible artifacts; not fully IoT.
CICIoT2023	Recent IoT	Large scale and numerous IoT attacks	Voluminous; close and imbalanced subclasses.
Edge-IIoTset	Edge/IIoT	Edge and IIoT scenarios	Less standardized use in reviews.

5.2 Metrics

Accuracy alone is not sufficient. Rare attacks may be poorly detected despite high accuracy. Recommended metrics include macro-F1, weighted-F1, per-class recall, AUC, MCC, false-positive rate, normalized confusion matrix, inter-client variance, communication cost, and inference latency.

5.3 Methodological issues

Frequent errors include global normalization before splitting, random line-by-line separation when correlated flows come from the same capture, undocumented removal of rare classes, mixing clients between train and test sets, and the absence of multiple seeds. These errors may produce non-generalizable results.

6 Deep Learning for IoT IDSs

6.1 CNN

1D CNNs are well suited to flow vectors and temporal windows. They capture local patterns and remain less costly than Transformers. Their limitation is the difficulty of modeling long-range dependencies without a hybrid architecture.

6.2 LSTM/GRU and hybrid CNN–LSTM models

LSTM and GRU networks process sequences of packets or flows. CNN–LSTM architectures combine local extraction and temporal memory, making them relevant for IoT traffic. DeepFed illustrates this orientation by combining CNN/GRU in a federated framework for cyber-physical industrial systems [33].

6.3 Autoencoders

Autoencoders are useful for anomaly detection, especially when attack data are rare. Fed-ANIDS explored a federated approach based on autoencoders for network anomaly detection and highlighted the relevance of FedProx in distributed environments [34].

6.4 Transformers and attention

Transformers learn global dependencies but are costly. Attention mechanisms may nevertheless help identify relevant features or time steps. Transformer models applied to IoT time series and attention-convolution approaches, such as MACML, show the relevance of combining global and local representation for few-shot IDSs [7, 19].

6.5 GNN

GNNs are relevant when relationships among sensors, devices, or network flows are known. Their use remains less mature in federated IoT IDSs because graph construction can be costly and domain-dependent.

7 Federated Learning for IoT IDSs

7.1 FL architectures: cross-device and cross-silo

IoT IDSs can be federated in cross-device mode, when many objects or gateways participate with limited resources, or in cross-silo mode, when a few industrial sites, hospitals, or organizations collaborate. Cross-device FL raises issues of availability, bandwidth, and heterogeneity; cross-silo FL improves stability but requires institutional agreements.

7.2 FL algorithms

Table 11: Federated algorithms and relevance to IoT IDSs

Method	**Principle**	**IoT IDS relevance**
FedAvg	Weighted average of local models	Simple baseline but sensitive to non-IID data.
FedProx	Proximal term limiting deviation from the global model	Reduces client drift in heterogeneous settings.
FedNova	Normalization of local updates	Useful when clients perform different numbers of local epochs.
SCAFFOLD	Control variates to correct drift	Relevant but more costly in communication/memory.
Robust aggregation	Median, trimmed mean, Krum, outlier detection	Defense against malicious clients or noisy updates.
PFL/Meta-FL	Global model + local adaptation	Adaptation mechanism for non-IID data and rare attacks, to be evaluated with edge cost and external validation.

7.3 Non-IID data in IoT

Non-IID data are the central problem of federated IoT IDSs. Campos et al. evaluate FL for IoT IDSs with different distributions on TON_IoT [11]. Djaidja et al. show, in a 5G/B5G context, that FL can improve privacy but may also suffer from heterogeneous distributions and from the choice of experimental strategy [15]. Methods such as FedProx, meta-sampling, dual correction, or incremental FL seek to mitigate these effects [17, 35, 36].

7.4 FL robustness and security

FL does not prevent all attacks. Clients may inject incorrect labels, malicious gradients, or backdoors. Updates may also leak information. Defenses include differential privacy, secure aggregation, homomorphic encryption, malicious-client detection, and robust aggregation [10, 37, 38].

7.5 Communication efficiency

Edge IDSs must reduce the number of rounds, transmitted volume, and model size. Solutions include compression, quantization, sparsification, distillation, pruning, and lightweight learning. Recent work on lightweight federated IDSs for IIoT emphasizes the need to measure real feasibility on constrained devices [39, 40].

8 Meta-Learning and Federated Personalization

8.1 Why MetaL?

Zero-day attacks, new malware families, and traffic changes create scenarios where labeled data are scarce. Meta-learning therefore aims to produce a model able to adapt from few examples. This capability is particularly important when each IoT client observes only part of the attacks.

8.2 MAML, ProtoNet, and Reptile

MAML is the most frequently cited method for optimization-based rapid adaptation [41, 42]. Metric methods, such as prototypical networks, learn a space in which classes are separable from few examples. Reptile simplifies the double loop of MAML. In IoT IDSs, the choice depends on the computational cost that can be accepted on the client.

8.3 Personalized FL and Meta-FL

Unique global FL models are often insufficient when clients are very different. Federated personalization preserves global knowledge while adapting the model locally. Yan et al. show that the question of "global or local adaptation" is central for a personalized IoT IDS [18]. Lu et al. and Xu et al. confirm the value of few-shot learning for IoT IDSs when attack classes are rare [16, 19].

8.4 Meta-federated IDS

A meta-federated IDS can be viewed as a combination of three levels: global inter-client learning, fast local adaptation, and robustness to non-IID distributions. This approach differentiates the present review from syntheses that treat FL or meta-learning separately. It remains insufficiently validated at scale and requires evaluation protocols that measure edge cost, communication, and leakage risks.

8.5 Operational mapping of adaptation and cost

Table 12 links the formal MAML framework introduced in Section 2.5 to the operational criteria used in this review. This avoids treating meta-learning equations as disconnected theoretical material.

Table 12: Operational mapping between MAML formalism, IoT IDS adaptation, and edge cost

Formal element	**IoT IDS interpretation**	**Extraction criterion used in this review**
$\mathcal{T}_i$	Local task defined by a device, site, protocol, attack family, or collection period.	Does the study define the client/task structure and separate support/query or train/test data without leakage?
θ	Global initialization or shared model learned across clients or tasks.	Is the global model, initialization, or aggregation step explicitly described?
θ_i'	Local adaptation after one or several gradient steps.	Does the study report local adaptation, number of steps, local data size, and adaptation protocol?
α, β	Inner and outer learning rates.	Are hyperparameters, seeds, and optimization settings reported?
Communication rounds	Exchange between clients and server during FL or Meta-FL.	Are rounds, transmitted volume, model size, or participation frequency measured?
Edge constraints	Memory, latency, energy, and computation on gateway or device.	Are inference latency, memory footprint, energy, or edge feasibility reported?

9 Qualitative Comparative Synthesis

This section is not a statistical meta-analysis because the studies use very different datasets, splits, metrics, and protocols. It instead proposes a qualitative comparative synthesis and a numerical comparison when results are available.

Before the cross-cutting synthesis, the results of the included studies were examined according to their main contribution. Representative studies were grouped by methodological family: deep models for IoT IDSs, federated learning for distributed IDSs, approaches addressing non-IID data, robustness and privacy mechanisms, few-shot/meta-learning methods, and edge-oriented approaches. For each group, the analysis considered the dataset used, model type, experimental protocol, reported metrics, and methodological limitations. This organization makes it possible to compare contributions without assuming sufficient homogeneity for a statistical meta-analysis.

Table 13: Characteristics of studies contributing to each synthesis and associated risk of bias

Synthesis	Main study characteristics	Overall risk of bias	Synthetic justification
Deep Learning for IoT IDSs	Studies using CNN, LSTM/GRU, autoencoders, Transformers, GNNs, or hybrid models to detect intrusions, anomalies, or malware in IoT/IIoT flows.	Moderate	Architectures are well described, but comparisons remain limited by dataset, split, and metric heterogeneity.
Federated Learning for IoT IDSs	Studies applying FedAvg, FedProx, FedNova, SCAFFOLD, or federated variants to distributed IDS scenarios.	Moderate	FL is directly related to the topic, but protocols vary according to number of clients, data partition, aggregator, and reported metrics.
Non-IID data, client drift, and robustness	Studies analyzing statistical heterogeneity, label skew, feature skew, malicious clients, poisoning, backdoors, or gradient leakage.	Moderate	Non-IID data are recognized as a central factor but are often simulated differently across studies, limiting comparability.
Few-shot, meta-learning, and federated personalization	Studies using MAML, few-shot learning, personalized FL, or Meta-FL for adaptation to rare attacks or new clients.	Low to moderate	Some protocols report encouraging results, but the number of studies remains limited and large-scale validation in real IoT contexts is still insufficient.
Edge deployment, communication, and lightweight models	Studies reporting communication cost, latency, memory, number of rounds, compression, or edge feasibility.	Low	Few studies (26%) systematically evaluate memory consumption, energy, inference latency, or real deployment on constrained devices.
Explainability and interpretability of IDSs	Studies using or discussing XAI, SHAP, LIME, feature importance, or IDS alert interpretation.	Low	Explainability is recognized as necessary, but it is rarely integrated and evaluated in federated IoT IDSs.

9.1 Performance comparison

Deep models often achieve very high accuracies on Bot-IoT, TON_IoT, or CICIoT2023, but these results should be interpreted with caution. Majority classes may dominate the metric, and IID splits may hide operational difficulty.

9.2 Impact of non-IID data

FedAvg may experience a substantial accuracy decrease when label skew is strong, as reported in non-IID FL evaluations such as **(author?)** [11]. FedProx, meta-sampling, dual correction, personalization, and pretraining can improve convergence, but their cost and effectiveness conditions remain variable [11, 17, 36].

Table 14 makes explicit the classification used in this review. When a study mentions heterogeneity without defining the partition or reporting client-level measurements, it is classified as "unspecified non-IID". When no information makes it possible to identify a type of heterogeneity, the study is classified as "not documented" in order to avoid overinterpretation.

Table 14: Qualitative mapping of non-IID types addressed in IoT IDS/FL studies

Type of non-IID data	**Operational criterion used**	**Representative studies in the corpus**
Label skew	Attack-class distribution differs across clients; some classes are rare or absent on a given client.	Campos et al. [11], Han et al. [17], Yan et al. [18], Lu et al. [16].
Feature skew	Same label space but different traffic characteristics according to device, protocol, site, or network environment.	Djaidja et al. [15], Jin et al. [36], Ruzafa-Alcázar et al. [37].
Quantity skew	Strongly different number of samples across clients or imbalanced client participation.	FedAvg/FedProx/FedNova serve as references for analyzing this case [8, 24, 25]; several IDS studies simulate it without always measuring it explicitly.
Concept drift	Temporal change in traffic, attacks, or normal behavior across observation periods.	Jin et al. [35] and studies on emerging attacks and incremental adaptation.
Unspecified non-IID	The study states that heterogeneous or federated data are used without specifying label skew, feature skew, quantity skew, or drift.	FL/IDS studies whose protocols do not describe client partitioning in sufficient detail.
Not documented	The study mainly addresses centralized DL, a dataset, or a methodological reference; no non-IID type can be evaluated.	Context references, datasets, and algorithmic articles used as theoretical foundations.

9.3 Impact of few-shot learning

Few-shot approaches are relevant in scenarios where some attack classes are very rare. They must be evaluated using clearly defined N-way K-shot protocols, separating training and test classes to avoid semantic leakage [16, 19].

9.4 Accuracy/communication trade-off

Federated methods can reduce raw-data sharing, but not necessarily the total network cost. Therefore, studies should report rounds, transmitted volume, compression, participation frequency, and model size.

9.5 Bias related to missing results

The risk of bias related to missing results was assessed qualitatively. No formal statistical analysis, such as a funnel plot or Egger's test, was performed because the included studies show strong heterogeneity in datasets, experimental protocols, metrics, architectures, and federated scenarios. In addition, this review does not perform a quantitative meta-analysis.

The assessment shows a high risk of bias related to missing results. This risk mainly comes from the fact that some studies may fail to report negative results, experimental failures, per-class performance, communication costs, or results obtained under difficult non-IID distributions. Several works also favor global metrics, such as accuracy, at the expense of more informative measures such as macro-F1, per-class recall, MCC, latency, or edge cost.

To partially limit this bias, the review included studies from several indexed scientific sources and did not restrict the analysis to work reporting only high performance. Studies presenting methodological limitations, mixed results, communication constraints, non-IID difficulties, or edge-deployment problems were also considered when they met the inclusion criteria. Despite these precautions, the possibility that some negative, inconclusive, unpublished, or insufficiently reported studies were not identified remains high, because the grey literature was not searched systematically.

10 Cross-Cutting Critical Analysis

10.1 Limited comparability

Results are difficult to compare because studies differ in datasets, cleaning, retained classes, splits, number of clients, metrics, and architectures. Two studies using the same dataset may reach opposite conclusions if one uses an IID split and the other uses a device-based partition.

10.2 Insufficient reproducibility

Too many studies do not publish code, seeds, hyperparameters, preprocessing scripts, or client partitions. This limitation weakens confidence in the results and complicates comparisons.

Table 15: Reproducibility indicators extracted from the empirical studies

Indicator	**Observed proportion**	**Interpretation**
Public source code available	5/23 studies (22%)	Most empirical studies do not provide complete code, which limits independent reproduction.
Hyperparameters sufficiently described	11/23 studies (48%)	Several studies report model settings, but not always optimizer, learning rate, rounds, or client configuration.
Random seed or repeated runs reported	7/23 studies (30%)	Performance variability is rarely quantified across multiple seeds.
Client partition or non-IID split documented	9/23 studies (39%)	FL studies often mention heterogeneity without a precise operational description of client partitions.
Communication, latency, memory, or energy measured	6/23 studies (26%)	Edge feasibility remains weakly documented despite being central for IoT/IIoT deployment.

These indicators show that the reproducibility of the corpus remains limited. The lack of code, seeds, and client partitions makes it difficult to verify whether reported performance comes from the model itself, from preprocessing choices, from data leakage, or from a favorable split. This limitation supports the recommendation to publish code, preprocessing scripts, client partitions, hyperparameters, random seeds, and per-class results.

10.3 Unrealistic evaluations

Many papers evaluate simulated clients on a single centralized dataset that is then artificially partitioned. Yet, in a real IoT system, clients differ in hardware, firmware, protocol, traffic, schedules, and exposure to attacks.

10.4 Weak generalization

IDSs evaluated on a single dataset may learn artifacts. A robust evaluation should include cross-scenario validation, external testing, and per-class analysis.

10.5 Ignored edge cost

Accuracy is often reported without latency, memory, energy, or communication volume. Yet an IoT IDS must be deployable on gateways or constrained objects.

10.6 Missing explainability

Deep and federated models remain opaque. SHAP, LIME, feature importance, and per-class analysis should be used to verify that the model learns relevant behavior rather than artifacts [43–45].

11 Open Challenges and Research Agenda

Table 16: Open challenges and research directions

Challenge	Problem	Recommended direction
Realistic non-IID data	Artificial partitions do not reflect real clients	Partition by device, site, time, or protocol; heterogeneity measurement.
Personalized FL	A unique global model fails on atypical clients	PFL, federated clustering, local adaptation, meta-initialization.
Few-shot zero-day	Rare attacks are underrepresented	MAML, ProtoNet, support/query sets, rapid adaptation.
Robust aggregation	Malicious or noisy clients	Trimmed mean, median, Krum, reputation, cross-client validation.
Privacy	Gradients may leak information	DP, secure aggregation, homomorphic encryption, leakage audit.
TinyML/edge	Models are too heavy for objects or gateways	Pruning, quantization, distillation, CPU/memory/latency measurement.
Explainable IDS	Opaque alerts	SHAP/LIME, hybrid rules, understandable alert reports.
Benchmarks	Lack of standardized protocols	Public splits, seeds, per-class metrics, code, and anti-leakage checklist.

12 Limitations of the Review

This review has several limitations that should be acknowledged to interpret the results correctly. First, although the search strategy covered three major scientific sources, namely IEEE Xplore, ScienceDirect/Elsevier, and Springer Nature, some relevant studies indexed in other databases, published as preprints, or available in less visible proceedings may not have been retained. The review therefore sought to prioritize scientific quality and source traceability, but it does not claim to exhaustively cover all existing work on IoT IDSs based on deep learning, federated learning, and meta-learning.

Second, the substantial heterogeneity of the included studies limits the ability to directly compare reported performance. The analyzed works differ in datasets, preprocessing strategies, retained attack classes, partitioning methods, architectures, metrics, and federated scenarios. Thus, two studies using the same dataset may obtain results that are difficult to compare if one applies a random IID split whereas the other adopts a non-IID partition by client, device, or attack class. This heterogeneity justifies the choice of a qualitative synthesis rather than a statistical meta-analysis.

Third, the methodological quality of the included studies remains variable. Several works do not sufficiently describe preprocessing steps, hyperparameters, seeds, train/validation/test partitions, or reproduction scripts. In addition, source code and experimental configurations are not always available. This partial reproducibility reduces confidence in some conclusions and complicates the precise identification of factors responsible for observed performance.

Fourth, some limitations concern the reported metrics. Many studies still favor accuracy or weighted-F1, whereas these measures can hide poor detection of rare or minority attacks. In an IoT IDS context, per-class metrics, macro-F1, minority-class recall, MCC, normalized confusion matrices, and false-positive rate should be reported systematically. The absence of these measures in some studies limits the real evaluation of model robustness against rare attacks, zero-day attacks, and imbalanced distributions.

Fifth, operational constraints related to edge deployment are still insufficiently documented. Few studies report inference latency, memory consumption, energy cost, communication volume, number of federated rounds, or the size of transmitted models in detail. Yet these elements are essential for judging the real feasibility of an IDS in IoT or IIoT environments composed of resource-constrained devices.

Finally, conclusions related to meta-learning, few-shot learning, and federated personalization should be interpreted with caution. These approaches may support adaptation to rare attacks and non-IID environments, but they remain less studied than classical deep architectures and basic federated methods such as FedAvg or FedProx. Large-scale validations across multiple datasets and realistic IoT scenarios remain limited. Therefore, this review identifies them as insufficiently validated approaches whose general superiority in IoT IDS contexts cannot yet be established.

13 Conclusion

This review has shown that IoT IDSs can no longer be analyzed solely as centralized classification models. IoT and IIoT environments are distributed, heterogeneous, exposed to rare attacks, and constrained by edge deployment. FL improves privacy and collaboration, but remains fragile when facing non-IID data, client drift, adversarial attacks, and communication cost. The synthesis also shows that meta-learning, few-shot learning, and federated personalization remain insufficiently validated at scale. Their relevance should therefore be confirmed through reproducible, robust, explainable, and edge-realistic protocols that report client partitions, per-class metrics, communication cost, latency, memory, and code availability.

14 Conflict-of-Interest and Funding Statement

We declare that we have no conflict of interest related to this review. We received no specific funding for this study.

References

1. A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, IEEE Internet of Things Journal **1**(1), 22 (2014). DOI 10.1109/JIOT.2014.2306328
2. L. Atzori, A. Iera, G. Morabito, Computer Networks **54**(15), 2787 (2010). DOI 10.1016/j.comnet.2010.05.010
3. S. Sicari, A. Rizzardi, L.A. Grieco, A. Coen-Porisini, Computer Networks **76**, 146 (2015). DOI 10.1016/j.comnet.2014.11.008
4. B.B. Zarpelão, R.S. Miani, C.T. Kawakani, S.C. de Alvarenga, Journal of Network and Computer Applications **84**, 25 (2017). DOI 10.1016/j.jnca.2017.02.009
5. A. Abduvaliyev, A.S.K. Pathan, J. Zhou, R. Román, W.C. Wong, IEEE Communications Surveys & Tutorials **15**(3), 1223 (2013). DOI 10.1109/SURV.2012.121912.00006
6. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Advances in Neural Information Processing Systems **30** (2017). DOI 10.48550/arXiv.1706.03762
7. Z. Chen, D. Chen, X. Zhang, Z. Yuan, X. Cheng, IEEE Internet of Things Journal **9**(12), 9179 (2022). DOI 10.1109/JIOT.2021.3100509
8. H.B. McMahan, E. Moore, D. Ramage, S. Hampson, B. Agüera y Arcas, Proceedings of the 20th International Conference on Artificial Intelligence and Statistics pp. 1273–1282 (2017)
9. P. Kairouz, H.B. McMahan, B. Avent, A. Bellet, M. Bennis, A.N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, et al., Foundations and Trends in Machine Learning **14**(1–2), 1 (2021). DOI 10.1561/2200000083
10. V. Mothukuri, R.M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, G. Srivastava, Future Generation Computer Systems **115**, 619 (2021). DOI 10.1016/j.future.2020.10.007
11. E.M. Campos, P.F. Saura, A. González-Vidal, J.L. Hernández-Ramos, J. Bernal Bernabe, G. Baldini, A.F. Skarmeta, Computer Networks **203**, 108661 (2022). DOI 10.1016/j.comnet.2021.108661

12. A. Belenguer, J.A. Pascual, J. Navaridas, Computer Networks **258**, 111023 (2025). DOI 10.1016/j.comnet.2024.111023
13. N. Hamdi, International Journal of Information Security **22**, 1937 (2023). DOI 10.1007/s10207-023-00727-6
14. A.B.Z. Mahmoodi, S. Sheikhi, E. Peltonen, P. Kostakos, IEEE Access **11**, 121325 (2023). DOI 10.1109/ACCESS.2023.3327922
15. T.E.T. Djaidja, B. Brik, A. Boualouache, S.M. Senouci, Y. Ghamri-Doudane, Computers & Security **139**, 103707 (2024). DOI 10.1016/j.cose.2024.103707
16. C. Lu, X. Wang, A. Yang, Y. Liu, Z. Dong, IEEE Internet of Things Journal **10**(24), 21309 (2023). DOI 10.1109/JIOT.2023.3283408
17. W. Han, J. Peng, J. Yu, J. Kang, J. Lu, D. Niyato, IEEE Internet of Things Journal **11**(8), 13340 (2024). DOI 10.1109/JIOT.2023.3337755
18. H. Yan, X. Lin, S. Li, H. Peng, B. Zhang, IEEE Transactions on Information Forensics and Security **20**, 279 (2025). DOI 10.1109/TIFS.2024.3516548
19. C. Xu, J. Yang, P. Li, PLOS ONE **20**(8), e0331065 (2025). DOI 10.1371/journal.pone.0331065
20. M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, et al., BMJ **372**, n71 (2021). DOI 10.1136/bmj.n71
21. B. Kitchenham, S. Charters, Guidelines for performing systematic literature reviews in software engineering. Tech. Rep. EBSE 2007-001, Keele University and Durham University Joint Report (2007)
22. H. Zhang, J. Ye, W. Huang, X. Liu, J. Gu, Journal of Parallel and Distributed Computing **195**, 104976 (2025). DOI 10.1016/j.jpdc.2024.104976
23. B. Buyuktanir, S. Altinkaya, G. Karatas Baydogmus, K. Yildiz, Cluster Computing **28**, 473 (2025). DOI 10.1007/s10586-025-05325-w
24. T. Li, A.K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, V. Smith, Proceedings of Machine Learning and Systems **2**, 429 (2020)
25. J. Wang, Q. Liu, H. Liang, G. Joshi, H.V. Poor, Advances in Neural Information Processing Systems **33**, 7611 (2020)
26. S.P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, A.T. Suresh, Proceedings of the 37th International Conference on Machine Learning pp. 5132–5143 (2020)
27. B.J. Shea, B.C. Reeves, G. Wells, M. Thuku, C. Hamel, J. Moran, D. Moher, P. Tugwell, V. Welch, E. Kristjansson, D.A. Henry, BMJ **358**, j4008 (2017). DOI 10.1136/bmj.j4008
28. N. Moustafa, J. Slay, 2015 Military Communications and Information Systems Conference pp. 1–6 (2015). DOI 10.1109/MilCIS.2015.7348942
29. N. Koroniotis, N. Moustafa, E. Sitnikova, B. Turnbull, Future Generation Computer Systems **100**, 779 (2019). DOI 10.1016/j.future.2019.05.041
30. A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, A. Anwar, IEEE Access **8**, 165130 (2020). DOI 10.1109/ACCESS.2020.3022862
31. I. Sharafaldin, A.H. Lashkari, A.A. Ghorbani, Proceedings of the 4th International Conference on Information Systems Security and Privacy pp. 108–116 (2018). DOI 10.5220/0006639801080116
32. E.C.P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, A.A. Ghorbani, Sensors **23**(13), 5941 (2023). DOI 10.3390/s23135941
33. B. Li, Y. Wu, J. Song, R. Lu, T. Li, L. Zhao, IEEE Transactions on Industrial Informatics **17**(8), 5615 (2021). DOI 10.1109/TII.2020.3023430
34. M.J. Idrissi, H. Alami, A. El Mahdaouy, A. El Mekki, S. Oualil, Z. Yartaoui, I. Berrada, Expert Systems with Applications **234**, 121000 (2023). DOI 10.1016/j.eswa.2023.121000
35. Z. Jin, J. Zhou, B. Li, X. Wu, C. Duan, Future Generation Computer Systems **151**, 57 (2024). DOI 10.1016/j.future.2023.09.019
36. Z. Jin, Y. Ding, X. Wu, X. Chen, G. Li, Computer Networks **259**, 111116 (2025). DOI 10.1016/j.comnet.2025.111116
37. P. Ruzafa-Alcázar, P. Fernández-Saura, E. Mármol-Campos, A. González-Vidal, J.L. Hernández-Ramos, J. Bernal Bernabe, A.F. Skarmeta, IEEE Transactions on Industrial Informatics **19**(2), 1145 (2023). DOI 10.1109/TII.2021.3126728
38. A. Puviarasu, V.K. Sudha, Ain Shams Engineering Journal (2026). DOI 10.1016/j.asej.2025.103866
39. S.J. Lee, I.G. Lee, ICT Express (2025). DOI 10.1016/j.icte.2025.05.002
40. S. Fraihat, Q. Yaseen, Y. Sanjalawe, A. Abu-Errub, S.N. Makhadmeh, M.A. Al-Betar, Discover Internet of Things **6**, 34 (2026). DOI 10.1007/s43926-026-00284-z
41. C. Finn, P. Abbeel, S. Levine, Proceedings of the 34th International Conference on Machine Learning pp. 1126–1135 (2017)

42. T. Hospedales, A. Antoniou, P. Micaelli, A. Storkey, IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(9), 5149 (2022). DOI 10.1109/TPAMI.2021.3079209
43. A.B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, Information Fusion **58**, 82 (2020). DOI 10.1016/j.inffus.2019.12.012
44. M.T. Ribeiro, S. Singh, C. Guestrin, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining pp. 1135–1144 (2016). DOI 10.1145/2939672.2939778
45. S.M. Lundberg, S.I. Lee, Advances in Neural Information Processing Systems **30** (2017)

A PRISMA 2020 Checklist

Section and topic	Item	Checklist element	Location in the manuscript
TITLE			
Title	1	Identify the report as a systematic review.	Page 1, line 1: yes. Title includes "systematic review, taxonomy, and research agenda".
ABSTRACT			
Abstract	2	See the PRISMA 2020 checklist for abstracts.	Page 1: yes. Abstract includes objective, main sources, PRISMA, taxonomy, results, and limitations.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Yes. Sections 1–1.2: IoT/IIoT growth, IDSs, FL, non-IID data, rare attacks, and edge constraints.
Objectives	4	Provide an explicit statement of the objectives or questions addressed by the review.	Yes. Sections 1.2–1.5: guiding thread, contributions, and research questions RQ1–RQ6.
METHODS			
Eligibility criteria	5	Specify inclusion and exclusion criteria and how studies were grouped for synthesis.	Yes. Section 3.4, Table 4. Grouping by DL, FL, non-IID data, meta-learning, datasets, and edge.
Information sources	6	Specify all databases, registers, and other sources searched and the date of the last search.	Section 3.2: IEEE Xplore, ScienceDirect/Elsevier, Springer Nature. Consultation/export date: 27-05-2026.
Search strategy	7	Present the complete search strategies for all databases, registers, and other sources.	Yes. Section 3.3, Table 3: queries by source and applied filters.
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria, including the number of reviewers and independence.	Section 3.5: selection follows PRISMA stages; title/abstract screening and full-text assessment were performed independently by two reviewers.
Data collection process	9	Specify methods used to collect data from reports.	Section 3.7: structured extraction grid and harmonization by consensus.
Data items	10	List and define all outcomes and variables for which data were sought.	Section 3.7: authors, year, source, context, dataset, attack type, architecture, FL aggregator, non-IID definition, clients, metrics, robustness, cost, edge constraints, and code availability.
Study risk of bias	11	Specify methods used to assess risk of bias in the included studies.	Section 3.6 and Table 6: AMSTAR 2 methodological assessment.
Effect measures	12	Specify for each outcome the effect measures used.	Sections 5.2 and 9: accuracy, macro-F1, weighted-F1, recall, AUC, MCC, false positives, communication cost, and latency.
Synthesis methods	13	Describe the methods used to decide which studies were eligible for each synthesis and how results were presented.	Sections 3.4–3.7; tables and figures summarize search strategy, criteria, PRISMA flow, quality assessment, taxonomy, datasets, FL algorithms, and research agenda.
Reporting bias assessment	14	Describe methods used to assess risk of bias due to missing results.	Section 3.8: publication bias was assessed qualitatively.
Certainty assessment	15	Describe methods used to assess confidence in the evidence.	Section 3.9: confidence was assessed qualitatively from methodological quality, consistency, dataset diversity, and realism.
Study selection	16	Describe the results of the search and selection process, ideally using a flow diagram, and cite excluded studies with reasons.	Section 3.5, Table 5, Figure 1, and Appendix D.
Study characteristics	17	Present characteristics of included studies.	Sections 5–8 and the representative-study appendix.

Table 17: PRISMA 2020 checklist

Section and topic	Item	Checklist element	Location in the manuscript
Risk of bias in studies	18	Present risk-of-bias assessment for each included study.	Tables 7 and 20.
Results of individual studies	19	Present results for each study in a summarized form.	Sections 6–9 and Appendix C.
Results of syntheses	20	Summarize characteristics of studies contributing to each synthesis and present synthesis results.	Section 9, Table 13; no statistical meta-analysis was performed.
Reporting biases	21	Present the assessment of bias due to missing results.	Section 9: moderate qualitative risk.
Certainty of evidence	22	Present the confidence level of evidence.	Table 8: evidence levels by theme.
Discussion	23	Provide general interpretation, limitations, and implications.	Sections 10–13.
Registration and protocol	24	Provide registration information, protocol access, and amendments.	Not registered; no specific funding or known conflict of interest declared.
Funding	25	Describe funding sources and the role of funders.	No specific funding declared.
Competing interests	26	Declare potential conflicts of interest.	No conflict of interest.
Availability of data, code, and materials	27	Indicate availability of data, code, and other materials.	Not applicable.

Table 18: PRISMA 2020 checklist

Section and topic	Item	Checklist item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objectives or questions addressed by the review.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g., databases, registers) used to identify studies and the date when they were last searched.	Yes
Risk of bias	5	Specify the methods used to assess the risk of bias in the included studies.	Yes
Results synthesis	6	Specify the methods used to present and synthesize the results.	Yes
Included studies	7	Indicate the total number of included studies and participants, and summarize the relevant characteristics of the studies.	Yes (44)

Table 19: PRISMA 2020 Checklist for Abstract

B Risk-of-Bias Assessment of Included Studies and Associated Methodological References

Table 20: Risk-of-bias assessment of included studies

Study	Type	Risk	Synthetic justification
Atzori et al. (2010)	IoT context	N/A	Context reference on IoT; no experimental IDS evaluation.
Zanella et al. (2014)	IoT context	N/A	Reference on smart cities; not assessed with the IDS/FL grid.
Sicari et al. (2015)	IoT security	N/A	Context reference on security, privacy, and trust in IoT.
Zarpelão et al. (2017)	IoT IDS review	Moderate	Useful review for IoT IDS taxonomy, but prior to recent FL and Meta-FL work.
Roman et al. (2013)	WSN/IDS review	Moderate	Context study on sensor networks; indirect relevance to modern IoT.
Vaswani et al. (2017)	DL reference	N/A	Foundational reference on Transformers; not an IoT IDS study.
Chen et al. (2022)	DL/IoT Transformer	Moderate	Relevant study on IoT anomaly detection with Transformers, but not centered on FL or Meta-FL.
McMahan et al. (2017)	Foundational FL	N/A	Foundational FedAvg paper; not evaluated as an IDS study.
Kairouz et al. (2021)	FL review	N/A	Review reference on open FL problems; not an empirical IDS study.
Mothukuri et al. (2021)	FL security review	Moderate	Important review on FL security and privacy, but not specific only to IoT IDSs.
Campos et al. (2022)	FL IoT IDS	Low	Directly related to FL for IoT IDSs, with discussion of challenges and non-IID data.
Belenguer et al. (2025)	FLIDS review	Low	Recent review on FL applied to IDSs, useful for taxonomy and methodological limits.
Hamdi (2023)	FL IoT IDS	Moderate	Directly related to federated IoT IDSs, but edge cost and reproducibility may remain partial.
Mahmoodi et al. (2023)	Distributed FL IDS	Moderate	Relevant study on distributed IDSs, but IoT/IIoT generalization and non-IID details should be checked according to protocol.

Study	Type	Risk	Synthetic justification
Djaidja et al. (2024)	5G/B5G FL IDS	Low	Recent experimental study useful for discussing FL benefits and limitations in IDSs.
Lu et al. (2023)	Few-shot/MAML IoT IDS	Low	Central study for MAML and few-shot IoT IDSs; highly relevant for rare attacks.
Han et al. (2024)	Non-IID/meta-sampling	Low	Directly related to non-IID heterogeneity and federated learning for IDSs.
Yan et al. (2025)	Personalized Meta-FL	Low	Central study for federated personalization and meta-learning in IoT IDSs.
Xu et al. (2025)	Few-shot IoT IDS	Moderate	Relevant study on few-shot learning and attention-convolution, but not centered on FL.
Page et al. (2021)	PRISMA	N/A	Methodological reference; not evaluated as an IDS study.
Kitchenham and Charters (2007)	SLR methodology	N/A	Methodological reference for systematic reviews.
Li et al. (2020)	FedProx	N/A	Foundational algorithmic reference; not specific to IoT IDSs.
Wang et al. (2020)	FedNova	N/A	Foundational algorithmic reference; not specific to IoT IDSs.
Karimireddy et al. (2020)	SCAFFOLD	N/A	Foundational algorithmic reference; not specific to IoT IDSs.
Moustafa and Slay (2015)	IDS dataset	N/A	UNSW-NB15 dataset reference; not a federated IoT IDS model study.
Koroniotis et al. (2019)	IoT dataset	N/A	Bot-IoT dataset reference; used to discuss data and limitations.
Alsaedi et al. (2020)	IoT/IIoT dataset	N/A	TON_IoT dataset reference; not evaluated as an IDS method.
Sharafaldin et al. (2018)	IDS dataset	N/A	CICIDS dataset reference; useful for experimental protocols.
Neto et al. (2023)	IoT dataset	N/A	CICIoT2023 dataset reference; used to discuss recent datasets.
Li et al. (2021)	Industrial DeepFed IDS	Moderate	Relevant study on federated IDSs in industrial systems, but edge cost and reproducibility require verification.
Idrissi et al. (2023)	Fed-ANIDS	Low	Recent study directly related to FL for network anomaly detection.
Jin et al. (2024)	Incremental FL IDS	Moderate	Relevant for incremental learning, but comparison and reproducibility require verification.
Jin et al. (2025)	Heterogeneous aggregation	Low	Recent study centered on efficient aggregation in heterogeneous-data contexts.
Ruzafa-Alcázar et al. (2023)	Privacy-preserving FL IIoT	Low	Relevant study on privacy, FL, and industrial IDSs.
Puviarasu and Sudha (2026)	Privacy-preserving FL	Moderate	Very recent study; metadata and protocol should be rechecked before final submission.
Lee and Lee (2025)	Lightweight IIoT FL	Moderate	Relevant for lightweight deployment, but experimental details and reproducibility require verification.
Fraihat et al. (2026)	Hybrid IIoT DL	Moderate	Recent study on hybrid DL and feature optimization; not centered on FL/Meta-FL.
Finn et al. (2017)	MAML	N/A	Foundational meta-learning reference; not specific to IoT IDSs.
Hospedales et al. (2022)	Meta-learning review	N/A	Review reference on meta-learning; not specific to IoT IDSs.
Arrieta et al. (2020)	XAI	N/A	Review reference on explainability; not specific to federated IoT IDSs.
Ribeiro et al. (2016)	LIME	N/A	Foundational XAI reference; not evaluated as an IDS study.
Lundberg and Lee (2017)	SHAP	N/A	Foundational XAI reference; not evaluated as an IDS study.
Zhang et al. (2025)	FLIDS review	Low	Recent review on FL in IDSs, useful for synthesizing challenges and architectures.
Buyuktanir et al. (2025)	FLIDS review	Low	Recent review on advances, applications, and future directions in FLIDS.

C Representative Studies and Thematic Synthesis

Table 21: Representative studies included in the qualitative synthesis

Study	Source	Theme	Result or role in the review
Hamdi (2023)	Springer	FL for IoT IDSs	Shows the relevance of FL for IoT intrusion detection.
Mahmoodi et al. (2023)	IEEE	Autonomous FL	Studies a distributed IDS based on autonomous federated learning.
Lu et al. (2023)	IEEE	Few-shot/MAML	Shows the relevance of MAML for rapid adaptation to rare attacks.
Djaidja et al. (2024)	ScienceDirect	FL for 5G/B5G IDSs	Experimentally analyzes the advantages and limits of FL in IDS contexts.
Han et al. (2024)	IEEE	Non-IID/meta-sampling	Addresses data heterogeneity in federated IDSs.
Yan et al. (2025)	IEEE	Personalized Meta-FL	Studies local and global adaptation for personalized IoT IDSs.
Belenguer et al. (2025)	ScienceDirect	FLIDS review	Provides a recent synthesis of FL applications to IDSs.
Zhang et al. (2025)	ScienceDirect	FLIDS review	Analyzes architectures, challenges, and perspectives of FL in IDSs.
Fraihat et al. (2026)	Springer	Hybrid IIoT DL	Shows the relevance of hybrid DL and feature optimization.
Puviarasu and Sudha (2026)	ScienceDirect	Privacy-preserving FL	Addresses privacy and real-time IoT intrusion detection.

D Studies Assessed in Full Text and Then Excluded

Table 22 presents the 58 full texts assessed and then excluded from the final qualitative synthesis. Unlike a generic justification, each row specifies the dominant operational reason observed for the concerned study: absence of quantified non-IID partitioning, incomplete client protocol, short publication, redundancy with a more complete included study, application context distant from the research questions, or inability to obtain a complete extraction sheet for the critical variables. For studies whose titles already mention IoT, IIoT, FL, or IDS, the exclusion reason should not be read as an absence of thematic relevance, but as a methodological or documentary limitation preventing their integration into the main synthesis.

Table 22: Studies excluded after full-text assessment and operational exclusion reasons

No	Study	Identified title	Source	Main reason for exclusion
1	Dawit Dejene Bikila et al. (2026)	*FedPGO : A personalized federated learning framework with gradient orthogonalization for IoT intrusion detection*	ScienceDirect/Elsevier	The study is thematically relevant to PFL applied to IoT IDSs, but it was not integrated into the main corpus because a complete and verifiable extraction sheet was unavailable for the critical variables: client partitions, operational definition of non-IID data, per-class metrics, communication cost, and reproducibility.
2	Zhu et al. (2026)	*PFLSE: A Personalized Federated Learning Framework Based on Shannon Entropy Metric for Intrusion Detection in IIoT*	IEEE Xplore	The study is thematically relevant to personalized IIoT IDSs; it is excluded from the main synthesis because of redundancy with included studies on PFL/non-IID data and incomplete documentation of secondary variables required by the extraction grid: code availability, edge cost, client partition details, and reproducibility.
3	Shimin Sun et al. (2025)	*Robust intrusion detection based on personalized federated learning for IoT environment*	ScienceDirect/Elsevier	The study addresses an axis close to intelligent security, but it does not provide a complete IoT/IIoT IDS evaluation with comparable attack classes, baselines, and validation protocol.
4	Lu et al. (2025)	*Adaptive Neuron Honeypot: Trapping Malicious Backdoors in Federated Learning*	IEEE Xplore	Methodological redundancy with included studies more directly centered on few-shot/MAML IoT IDSs; the additional contribution to the DL–FL–MetaL taxonomy is limited.
5	Zhang et al. (2023)	*Federated Learning for Distributed IIoT Intrusion Detection Using Transfer Approaches*	IEEE Xplore	Short conference paper; insufficient information on the dataset, train/test split, per-class metrics, and protocol reproducibility.
6	Nandini Arimanda et al. (2026)	*FL-IDS++: A dynamic federated learning framework for intrusion detection with personalized non-IID data, adversarial resilience and energy-efficient lightweight models*	ScienceDirect/Elsevier	The study is very close to the inclusion criteria; it is not retained in the current synthesis because the complete extraction sheet could not be stabilized to compare personalization, adversarial robustness, energy cost, non-IID partitions, and per-class metrics simultaneously with the included studies.
7	Jia et al. (2025)	*Personalized Federated Learning for Cross-Area Vehicle Trajectory Anomaly Detection*	IEEE Xplore	The short format does not sufficiently describe the clients, aggregation strategy, training parameters, and metrics required by the QA grid.
8	Kushwaha et al. (2026)	*Energy-Aware Federated Learning for IoT Intrusion Detection Using Latent Feature Encoding*	IEEE Xplore	The IoT IDS validation is limited; the study lacks a reproducible federated scenario and an operational description of non-IID data or personalization.
9	Ma et al. (2026)	*Federated Learning and Resource-Aware Graph Neural Network for Intrusion Detection in 6G-IoT Driven Healthcare System*	IEEE Xplore	The study is close to the criteria, but it was not retained in the main synthesis because the 6G-healthcare scenario and resource-oriented GNN architecture do not provide sufficiently comparable extraction on attack classes, client partitions, per-class metrics, and edge/communication costs.
10	Baich et al. (2025)	*A Federated Learning-Based Intrusion Detection System Using Dynamic Ensemble Aggregation for IoT Networks*	IEEE Xplore	The contribution is redundant with included studies and does not provide a sufficiently detailed client protocol to classify the type of non-IID data.
11	Raja Waseem Anwer et al. (2025)	*Advanced intrusion detection in the industrial Internet of Things using federated learning and LSTM models*	ScienceDirect/Elsevier	The article does not allow extraction of a quantified non-IID partition or per-class metrics for a clearly defined IoT/IIoT IDS task.
12	Chen et al. (2025)	*An Improved Proximal Optimization Method for Cloud–Edge–End Collaboration in Railway Surveillance Networks*	IEEE Xplore	Short publication; insufficient details on the dataset, split, hyperparameters, seeds, and federated scenario.
13	Rajesh Kalakoti et al. (2025)	*Federated Learning of Explainable AI(FedXAI) for deep learning-based intrusion detection in IoT networks*	ScienceDirect/Elsevier	Coverage of the DL–FL–MetaL thread is partial; clients, split, and the type of non-IID data are not operationally described.
14	Prabu Selvam et al. (2025)	*Federated learning-based hybrid convolutional recurrent neural network for multi-class intrusion detection in IoT networks*	Springer Nature	Interesting adaptation approach, but the main variables cannot be extracted: dataset, classes, clients, comparable metrics, and adaptation cost.
15	Babatunde Olanrewaju-George et al. (2025)	*Federated learning-based intrusion detection system for the internet of things using unsupervised and supervised deep learning models*	ScienceDirect/Elsevier	The FL/DL study does not report the IoT IDS protocol, baselines, training parameters, and reproducibility elements with sufficient detail.
16	Kaur et al. (2025)	*A Review of Computational Intelligence Approaches for Fault-Tolerant IoT Federated Learning Systems*	IEEE Xplore	The validation context does not sufficiently match the scope of federated IoT/IIoT IDSs; client scenario and exploitable non-IID measurements are absent.
17	Kong et al. (2025)	*PFedGRID: Personalized Federated Graph-Based Robust Intrusion Detection for IoT Networks*	IEEE Xplore	Short publication; the description of the dataset, classes, metrics, and client protocol is insufficient for comparative analysis.
18	Hamad et al. (2025)	*Multi-Partner Project: CyberSecDome - Framework for Secure, Collaborative, and Privacy-Aware Incident Handling for Digital Infrastructure*	IEEE Xplore	A project-synthesis article; it does not provide a directly exploitable experimental IDS protocol for the methodological extraction grid.
19	Yu Lu et al. (2024)	*Adaptive Intrusion Detection in Internet of Things Using Quantum Classifier*	ScienceDirect/Elsevier	The study is centered on a specific quantum model and not on the DL–FL–MetaL axis retained for the main synthesis.
20	Dhakal et al. (2024)	*Adaptive Secure Monitoring to Detect Anomaly Using Federated Learning in Internet of Things*	IEEE Xplore	The study does not sufficiently specify client partitioning, non-IID distribution, per-class metrics, or reproducibility parameters.
21	Bhavsar et al. (2024)	*Anomaly based Intrusion Detection for IoT using Federated Learning*	IEEE Xplore	The approach is relevant to federated IoT anomaly detection, but the article does not provide enough information on baselines, client protocol, and metrics.
22	Mehrnoosh Nobakht et al. (2024)	*Approach for Detection of Anomalies in IoT Using the Convolutional Neural Networks and Federated Learning*	ScienceDirect/Elsevier	The method is relevant to DL/FL, but the evidence is not sufficiently documented for extracting the non-IID protocol, comparison conditions, and edge cost.
23	Abdel-Basset et al. (2023)	*Deep learning-empowered anomaly detection in cyber-physical systems using federated learning and blockchain*	IEEE Xplore	The study is redundant with more complete included studies and does not provide sufficiently precise extraction on client partitions, per-class metrics, and communication cost.
24	Anandraj et al. (2026)	*Federated and Swarm Learning for Intrusion Detection System in the Internet of Vehicles*	IEEE Xplore	The application domain is focused on Internet of Vehicles rather than the IoT/IIoT scope retained in this review; cross-client comparability remains limited.
25	Hasan et al. (2026)	*Federated Discriminative Label Embedding for Medical Vision Language Model: A Reinforcement Learning Approach*	IEEE Xplore	The application context is distant from IDS IoT/IIoT; the study does not provide a directly usable IDS protocol.
26	Dubey et al. (2026)	*Federated Learning Against Byzantine Attacks in IoT Networks*	IEEE Xplore	The evaluation does not clearly distinguish centralized, federated, and personalized learning; communication cost is not reported.
27	Saleem et al. (2026)	*Lightweight Federated Few-Shot Learning-Based Network Intrusion Detection for Resource-Constrained IoT Devices*	IEEE Xplore	The study is relevant to FL few-shot learning and constrained IoT devices; it is excluded from the main corpus because a complete and comparable extraction of per-class metrics, few-shot task definitions, client partitions, and real inference/communication cost is missing.
28	Singh et al. (2026)	*A Lightweight Hybrid Approach for DoS Defense in Software Defined Networks*	IEEE Xplore	The application scope does not sufficiently match the IoT/IIoT IDS criteria; the DL–FL–MetaL variables are not all extractable.
29	Shrivastava et al. (2025)	*A Federated Learning and Deep Learning-Based Intrusion Detection Mechanism for IoT Networks*	IEEE Xplore	Personalization or adaptation is mentioned without an N-way K-shot protocol, without defined client tasks, and without a clear train/test separation.
30	Rezaei et al. (2025)	*Adapt-LFA: Adaptive Gradient-Guided Label Flipping Attack Against Federated Learning-based Intrusion Detection in IoT*	IEEE Xplore	Mainly centered on an FL attack or defense; it does not provide a complete IoT IDS evaluation with dataset, classes, and IDS baselines.
31	Gupta et al. (2025)	*Adaptive Federated Learning for Real-Time Anomaly Detection in IoT-Enabled Communication Systems*	IEEE Xplore	The results do not sufficiently document per-class recall, macro-F1, or errors on minority classes; there is a risk of interpretation based on global accuracy.
32	Hasan et al. (2025)	*Adaptive Defense Against Packet-Mutation Adversarial Attacks in Network Intrusion*	IEEE Xplore	Operational description of clients and the type of non-IID data is absent; the protocol does not allow classification as label skew, feature skew, or quantity skew.

No	Study	Identified title	Source	Main reason for exclusion
33	Kalaneya et al. (2025)	*Bio-Inspired Spiking Neural Networks for Energy-Efficient Anomaly Detection in Wearable Embedded IoT Devices: A Federated Learning Approach*	IEEE Xplore	Secondary application context; the connection with IoT/IIoT IDSs, network attacks, and federated learning is not direct enough.
34	Hajla et al. (2025)	*HFEL: A hybrid federated ensemble learning framework for intrusion detection in IoT networks*	Springer Nature	The elements required for complete extraction are absent or incomplete: split, classes, per-class metrics, client partitions, and reproducibility.
35	Darwish et al. (2025)	*Comparative Analysis of Federated Learning, Deep Learning, and Traditional Machine Learning Techniques for IoT Malware Detection*	IEEE Xplore	The study does not report a sufficient comparison with strong IDS baselines or a clear analysis of edge or communication cost.
36	Li et al. (2025)	*A Cloud-Edge Collaborative Framework for Industrial IoT Device Fault Detection and Prediction*	IEEE Xplore	Short publication; details on dataset, split, clients, aggregation, and metrics are insufficient for the PRISMA/QA grid.
37	Herlambang et al. (2025)	*Federated Learning Approaches for IoT Intrusion Detection Based on FedAvg and FedProx on IID and Non-IID Data*	IEEE Xplore	The evaluation does not sufficiently specify attack classes, imbalance, per-class metrics, and reproducibility conditions.
38	Tiwari et al. (2025)	*A Federated Learning Approach for Time Series Forecasting and Robust Querying in Distributed Data Systems*	IEEE Xplore	The application context is distant from the main scope; no directly usable federated IoT/IIoT IDS evaluation is provided.
39	Bhavsar et al. (2025)	*Federated Learning using Deep Learning-Intrusion Detection system (FL DL-IDS) for IoT Applications*	IEEE Xplore	The protocol does not document client partitions, drift measurements, or per-client results required for non-IID analysis.
40	Khandaker et al. (2025)	*Federated Learning-Based IoT Intrusion Detection Using Multi-Objective Client Selection with Attention-Enhanced CNN*	IEEE Xplore	The study does not provide enough information on preprocessing, split, hyperparameters, and execution conditions to be reproducible.
41	Desai et al. (2025)	*Federated Learning-based Intrusion Detection Framework for Cyber Physical System-based IIoT*	IEEE Xplore	The study is relevant to IIoT, but the contribution does not allow complete extraction of client partition, non-IID conditions, per-class metrics, and reproducible parameters needed for comparison.
42	Kapil Kumar et al. (2025)	*Federated active meta-learning with blockchain for zero-day attack detection in industrial IoT*	Springer Nature	Secondary application axis; the protocol does not sufficiently connect model, IoT dataset, IDS tasks, and federated/edge constraints.
43	Alnajjar et al. (2025)	*Generative AI-Enabled Cybersecurity Using Federated Learning for Malware Detection in Industrial IoT Environment*	IEEE Xplore	Insufficient justification of baselines and split; the contribution does not allow robust evaluation of personalization or non-IID data.
44	Huynh et al. (2025)	*Hybrid Federated Learning with TabTransformer and FedMADE-GSA for IoT Intrusion Detection*	IEEE Xplore	Short publication; information on clients, aggregation, metrics, and reproducibility is insufficient.
45	Sen et al. (2025)	*Intrusion Detection System on IoT Devices via Federated Learning*	IEEE Xplore	The study does not report a task/client adaptation protocol; the link with MAML, few-shot learning, or personalization remains undocumented.
46	Alkabkabi et al. (2025)	*Intrusion Detection Systems for IoT Using Federated Learning*	IEEE Xplore	The results do not isolate the effect of non-IID data or client drift; per-client measurements are absent.
47	Elouardi et al. (2025)	*OptiFLIDS: Optimized Federated Learning for Energy-Efficient Intrusion Detection in IoT*	IEEE Xplore	The IDS protocol is incomplete for the synthesis: classes, splits, baselines, and critical metrics are not all reported.
48	Cakin et al. (2025)	*Privacy-Preserving Intrusion Detection in Industrial IoT Using Federated Learning*	IEEE Xplore	The study does not sufficiently document adaptation cost, number of rounds, latency, or communication volume.
49	Gao et al. (2025)	*Enhancing Privacy in 6G Networks with Federated Learning for Data Analytics*	IEEE Xplore	Clear description of IoT/IIoT data and experimental constraints is absent; results are difficult to compare with the included studies.
50	Wang et al. (2025)	*A Threshold Multi-CA Authentication Model for Privacy-Preserving Federated Semi-Supervised Learning in Healthcare IoT*	IEEE Xplore	The method is relevant, but the protocol does not allow simultaneous extraction of client scenario, non-IID data, per-class metrics, and reproducibility.
51	Rguibi et al. (2025)	*Hybrid Edge-Based Intrusion Detection for IoT Using Adaptive Lightweight SVM with Dynamic Feature Pruning*	IEEE Xplore	The link with federated IoT IDSs remains partial; the protocol does not sufficiently document personalization, adaptation, or communication cost.
52	Mutlu et al. (2025)	*Towards Explainable and Distributed Intrusion Detection in IoTs: A Federated Learning-Based Approach*	IEEE Xplore	The evaluation does not provide enough information on rare attacks, task separation, or minority-class results.
53	K et al. (2025)	*Title not found in the provided BibTeX exports*	IEEE Xplore	Secondary application context; a complete relationship among IoT intrusion detection, federated learning, and edge constraints is absent.
54	Samson et al. (2024)	*A Federated Learning-Based Hierarchical Intrusion Detection System for Underwater IoT Wireless Sensor Networks*	IEEE Xplore	The study is distant from the retained DL–FL–MetaL thread; the main variables of the extraction grid are unavailable.
55	Do et al. (2024)	*A Horizontal Federated Learning Approach to IoT Malware Traffic Detection: An Empirical Evaluation with N-BaIoT Dataset*	IEEE Xplore	Short conference format; insufficient details on client partitions, hyperparameters, seeds, and preprocessing scripts.
56	Benameur et al. (2024)	*A Novel Federated Learning Based Intrusion Detection System for IoT Networks*	IEEE Xplore	The study does not sufficiently describe data distribution, per-class metrics, and comparison conditions with FL/IDS baselines.
57	Sundar et al. (2024)	*A Smart and Secured Anomaly Detection Scheme for Internet of Things (IoT) Powered Devices Using Federated Learning Technology*	IEEE Xplore	The study is relevant to federated IoT anomaly detection, but the protocol does not allow comparable extraction of the adaptation mechanism, type of non-IID data, per-class metrics, and reproducibility parameters.
58	Hamdouchi et al. (2024)	*Comprehensive Evaluation of Federated Learning Configurations for Intrusion Detection in IoT Contexts*	IEEE Xplore	The study is relevant to FL configurations in IoT IDS contexts, but it was not retained because the extracted information remains insufficient to precisely compare dataset, split, clients, per-class metrics, non-IID data, and reproducibility.