A systematic literature review on feature selection for machine learning-based attack classification for IoT security

Jing Li

Address (ADR)

Abstract: Feature selection plays a crucial role in creating effective machine learning-based attack classification models for IoT security. However, to fully understand the significance and characteristics of feature selection using IoT data sets, it is important to systematically review the literature in this area. This paper presents a systematic review of feature selection approaches used in IoT data sets for IoT security. The aim of this review is to provide researchers and practitioners with a comprehensive overview of the state-of-the-art feature selection techniques and to assist in selecting appropriate methods to pre-process data sets for creating machine learning-based security models. The review covers studies published between January 2018 and December 2022, focusing on machine learning-based attack classification models. Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines are employed to collect and analyse data from 1272 studies across six scientific databases, resulting in 63 primary studies that meet the inclusion criteria. The primary studies are analysed and categorised to answer research questions relating to current practices, feature selection methods, benchmark IoT datasets, feature selection validation methods, limitations, challenges, and future directions. The review provides a valuable reference for researchers and practitioners seeking to incorporate effective feature selection approaches for attack classification models in IoT security.

Keywords: Internet of Things; feature selection; IoT data set; attack detection; classification; IoT security; systematic literature review, machine learning, deep learning.

Reference to this paper should be made as follows: Jing Li (2023) 'A systematic literature review on feature selection for machine learning-based attack classification for IoT security', *Int. J. Sensor Networks*, Vol. X, No. Y4, pp.000–000.

Biographical notes: (ABS)

1 Introduction

Internet of Things (IoT) provides a promising opportunity to seamlessly bring the real world and digital world to form a larger, more intelligent network. The term IoT was initially proposed to refer to uniquely identifiable interoperable connected objects with radio-frequency identification (RFID) technology (Ashton 2009). After development and evolution for many years, IoT has become a compound technology involving embedded devices with sensors, wireless sensor network, operating system, software, data communication, middleware, big data and AI technologies for various applications over the internet (Xu et al., 2014). By 2030, there will be estimated 26 billion connections on the Internet of Things (IoT), according to a recent report from Statista (Statista 2022). These IoT devices are integrated into appliances based on IoT infrastructure that support a variety of protocols for communication over a public network. The characteristics of IoT including multiple-layer infrastructure, ubiquitous inter-connected IoT devices and less powerful capability in IoT system caused that IoT infrastructure is more vulnerable to various attacks compared with traditional IT infrastructure.

Many factors contribute to the vulnerability of IoT networks. IoT infrastructure that comprises four layers: perception, network, middleware, and application, each of which has its security vulnerabilities, including the gateways that connect them. Cyber criminals can exploit these vulnerabilities, making IoT security a major concern (Hassija et al. 2019). In addition, IoT has extensive applications in various sectors, including smart home, healthcare, manufacturing, agriculture, logistics, autonomous vehicles, and smart cities, with the ability to exchange data between the real and digital worlds. However, these devices are susceptible to security threats due to the large amount of confidential user data they carry. Ensuring the integrity, confidentiality, and availability of the IoT system is critical to protecting the user data (Kouicem et al., 2018). Moreover, the increasing demand for IoT devices has led to their deployment worldwide, but their low-cost design, coupled with limited resources such as bandwidth, battery, computational resources, and memory, restricts the integration of security applications (Hassija et al. 2019). Therefore, above concerns on IoT security, like many attack surfaces with multiple layers in IoT infrastructure, extensive applications among all sectors and limited resources on IoT devices, cause IoT vulnerable to various attacks by cyber criminals (Ahmad and Alsmadi 2021).

Attack detection ahead of time is quite crucial in IoT security domain and IDS is the one of the most effective methods to protect IoT network from attacks. Recently, attack detection from IoT datasets, generated from real or simulated IoT networks, is an increasingly popular and attractive area of study for many researchers. It depends on developing more effective classification models that can be used for classifying any unseen IoT data after training the model over a specific training dataset. Detecting and classifying attacks using IoT data, have posed a huge challenge for researchers in the field of computer science, as this kind of datasets contains a large number of features and huge number of examples, because there are so many IoT devices deployed all around the world with huge number of data generated every day (Al-Garadi et al. 2020). However, many of these features are considered irrelevant or redundant, and they must be removed by using an effective and efficient feature selection method to improve the performance and reduce the computational resources of the classification model. Therefore, researchers have employed much effort in coming up with more effective feature selection techniques that can increase classification's accuracy and decrease the computation time using a smaller number of features in detection of attacks in IoT networks.

Recently, there are many review studies focusing machine learning and deep learning on IoT datasets to find the patterns and classify various attacks in IoT. For example, (Nagaraja and Kumar 2018) reviewed early research on intrusion detection and highlights the lack of specificity regarding the measures used to detect intrusion, and presents various methods for feature selection and computation of high-dimensional data for identifying network intrusion. Then, (Hussain et al. 2020) presented the potential impact of IoT on our lives and the security challenges posed by resource-constrained IoT networks, and proposed the use of machine learning and deep learning techniques to address security problems. After that, (Bojarajulu et al., 2021) discussed how machine learning models with optimal feature selection can mitigate these attacks by detecting malicious network traffic, however, there is just a cursory analysis of FS in IoT data. Recently, (Ahmetoglu and Das 2022) discussed the growing amount of cyberattacks and how machine learning may automate the detection and prediction of attacks in network traffic using techniques including detection, classification, clustering, and anomaly analysis, but they did not concentrate on IoT-specific data. As a result, there are few reviews that specifically focused on feature selection techniques for IoT datasets in IoT security.

In this research, we aim to help other researchers by making a systematic literature review of related studies that used feature selection in machine learning-based attack classification models in IoT security in the last five years. Also, this research will help to specify the current situation, challenges and future directions in feature selection for IoT security. The main contributions of this systematic review can be summarized as follows:

1. Comprehensive literature searches and investigations are conducted, with a focus on feature selection using IoT data for attack detection in IoT security. There are 1272 papers collected with the specified keywords search strategy from six databases (Web of Science, IEEE Xplore, Scopus, ScienceDirect, ACM, SpringerLink, and Wiley Online Library) in the literature published from January 2018 to December 2022 on feature selection in IoT data for IoT security.

2. An extensive literature review using the methodology based on the standard PRISMA process is employed. This study proposes five research questions (twelve sub-questions in total). We initially formulate research questions and motivations, collect the related studies, define the including and excluding criteria, and finally identify the primary studies focusing on the topics with an acceptable quality score. Data items for each research question are extracted for intensive data analysis to provide answers and discussion to the proposed research questions.

3. Intensive explanation and discussion have been implemented in the primary studies. Each paper is reviewed based on the extracted data items to answer the proposed research questions, involving the current situation and the objectives of FS, FS methods, IoT datasets, FS validation methods, limitations, challenges, and future directions. This aims to provide researchers and practitioners who want to pursue research on feature selection using IoT data to build machine learning or deep learning-based attack classification models in IoT networks.

The remainder of this paper is structured as follows: The research methodology involving research questions and research protocols is presented in Section 2. Section 3 presents the results of this literature review to provide answers and discussion to the research questions. The limitations of this study are stated in Section 4. Finally, a conclusion is provided in Section 5.

2 Methodology

To achieve the objectives of the current SLR study, we followed the standard and original guidelines proposed by (Keele and others 2007). This section presents the method used to undertake the current SLR of feature selection using IoT data for IoT security. The Figure 1 guides the stages of our methodology include:

A. Planning
B. Formulation of research questions and motivations
C. Search process
D. Inclusion and exclusion criteria
E. Quality assessment
F. Data collection
G. Data analysis

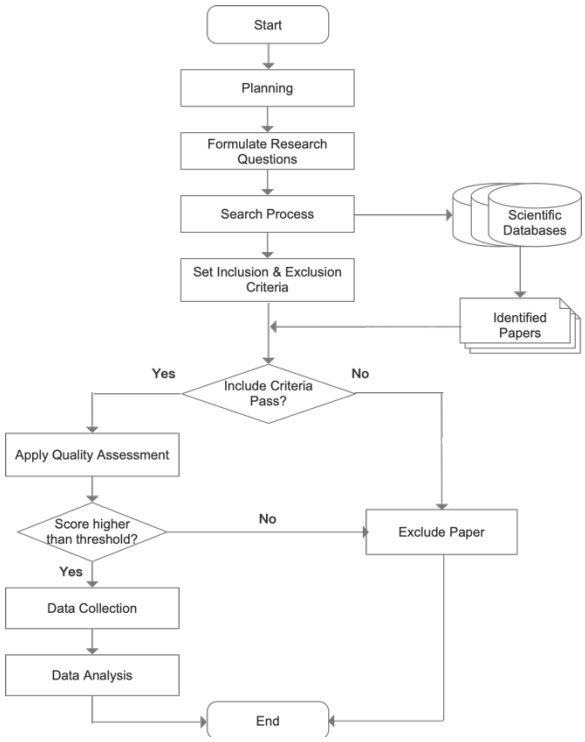

FIGURE 1. Search process flowchart.

A. PLANNING

During the planning phase, we determined the actions required to achieve the objectives of our study, which are related to the topic of how feature selection can contribute to the effect of machine learning or deep learning-based security models in IoT security. We ensured that both the strategic and technical strategies were appropriately developed at this early stage. This would guarantee that the remaining parts of the suggested technique were carried out in an organized and consistent way. This planning step laid the groundwork for the projected SLR methodology's effective implementation.

B. FORMULATION OF RESEARCH QUESTIONS AND MOTIVATIONS

In this section, we present the research questions investigated in the current SLR study, as well as the motivations behind them. The motivations come from the noteworthy achievements of the work using feature selection approaches for attack classification in the IoT security domain. For instance, most studies have indicated that feature selection is essential to processing heterogenous IoT data sets to improve the performance of attack detection and classification. Thus, the research questions RQs investigated in this study can be found in Table 1.

TABLE 1. Research questions and their motivations.

No.	Research Questions	Motivations
RQ1	What is current situation of FS approaches applied for machine learning or deep learning-based attack classification model for IoT security?	
RQ1.1	RQ1.1 What is the trend of the studies that applied FS for IoT security in recent five years?	The motivation of RQ1 is to identify highly related empirical studies during past five years, in order to acquire the situations of the studies on the topic in recent years
RQ1.2	What are the distributions of the studies according to databases and publishers?	
RQ1.3	What are the main objectives of applying feature selection for IoT security?	
RQ2	What are the FS methods and techniques applied on attack classification models for IoT security?	
RQ2.1	What are the main types of feature selection approaches applied for attack classification model?	RQ2 is motivated by the need to identify the main objectives and the FS methods applied in recent studies.
RQ2.2	What are the specific techniques utilized for each type of feature selection used in IoT security models?	
RQ3	What are the characteristics of related factors for FS methods for IoT security?	
RQ3.1	What are the IoT datasets as the benchmark by the studies when applying FS method?	RQ3 is motivated by the need to the idea on how the feature selection method can be applied concerning to the characteristics of datasets and attacks.
RQ3.2	What are mapping of datasets and attacks to various FS methods among the studies?	
RQ4	What are the verification methods to evaluate the effectiveness of proposed FS approaches?	
RQ4.1	What are the traditional machine learning and deep learning algorithms mapping to FS methods?	RQ4 is motivated by the need to the idea of how to verify the effectiveness of proposed FS methods.
RQ4.2	What are the performance metrics used for validation of FS approaches?	
RQ4.3	What are the methods of the validation of FS in studies?	
RQ5	What are the challenges and future directions using FS to the classification models in IoT security?	
RQ5.1	What are the limitations of the proposed FS in studies?	RQ5 is motivated by the need to find the limitations, challenges, and the research in future direction.
RQ5.2	What are the major challenges of applying FS methods?	
RQ5.3	What are the future research directions of FS in terms of the classification model for IoT Security?	

C. SEARCH PROCESS

In this subsection, we present how each article was identified for this study. For the purpose of extracting relevant experimental studies on FS method implementation using IoT data for attack classification in IoT security, a number of digital scientific databases were considered and accessed. The list of databases, their corresponding fields, search strings, and initial number of studies are presented in Table 2. Each article was extracted from a digital database using the conventional manual search process for journal articles and conference proceedings, respectively.

TABLE 2. Repositories and their corresponding search strings.

Digital Database	Field	Search Strings	No.
Web of Science	All	"feature selection" AND ("iot" OR "internet of things") AND "security" AND ("machine learning" OR "deep learning")	178

IEEE Xplore	All	"feature selection" AND ("iot" OR "internet of things") AND "security" AND ("machine learning" OR "deep learning")	113
Scopus	TITLE -ABS-KEY	"feature selection" AND ("iot" OR "internet of things") AND "security" AND ("machine learning" OR "deep learning")	258
ScienceDirect	All	"feature selection" AND "iot security" AND ("machine learning" OR "deep learning")	168
ACM Digital Library	All	("feature selection") AND ("iot" OR "internet of things") AND ("security") AND ("machine learning") 2018-2022	338
SpringerLink	All	"feature selection" AND "iot security" AND ("machine learning" OR "deep learning") 2018-2022	183
Wiley online Library	All	"feature selection" AND "iot security" AND ("machine learning" OR "deep learning")	34
Overall			1272

D. INCLUSION AND EXCLUSION CRITERIA

The studies included in this SLR were based on certain criteria that determined whether a study could meet the condition for inclusion; otherwise, such a study would be excluded. One of the important conditions each article was expected to meet was that it needed to be written in accurate and understandable English. Articles written in a different language were not included in the current study because such articles would be difficult to read and understand. The list of criteria for inclusion and exclusion is presented in Table 3.

TABLE 3. Inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
(a). Each article must be focused and related to the topic of feature selection applied in machine learning or deep learning-based classification model for IoT security.	(a). The studies that are not related to the topic, or only related to any one sub-topic like feature selection, machine learning or deep learning-based attack classification, or IoT security but not all above.
(b). Each article must be a proof of an empirical study addressing the research questions of the related topic.	(b). Technical reports, government reports, letters and editorials, short notes, book chapters, survey or review papers, and experimental papers that deviate from answering the research questions.
(c). The data set utilized by each article must be based on at least one public IoT data set or the IoT data extracted from IoT scenarios.	(c). The studies focusing on datasets not generated from IoT scenarios.
(d). Each article must be written in a simple and understandable English reported in a publication article, which can be accessed.	(d). Article must not be written in a different language than English, or cannot be accessed.
(e). Each article must be published within January 2018 – December 2022.	(e). Be published outside the period of time specified.

E. QUALITY ASSESSMENT

Besides the filtering by inclusion and exclusion criteria, we need to further screen the studies using quality assessment in order to narrow the scope for the following data collection and analysis. In order to assess the quality of the studies, our focus was primarily on finding pieces of evidence about how deeply and thoroughly authors have answered the research questions presented in this SLR. The quality assessment also helped to accurately extract the data from the rest of the research studies after excluding irrelevant research. The quality assessment criteria presented in Table 4 is formulated based on five quality assessment questions related to our research questions.

TABLE 4. Quality assessment questionnaire.

Quality Assessment Question	Relevant to the research question
QA1: Whether the complete data items can be extracted and how does the author(s) explain their research problems?	RQ1
QA2: How does the author(s) present the implementation of feature selection methods the studies?	RQ2
QA3: How does the author(s) present the concerning factors to feature selection methods in research methodology?	RQ3
QA4: How does the author(s) conduct comprehensive validation for the proposed research method?	RQ4
QA5: How does the author(s) present the study findings, limitations and future direction?	RQ5

Research papers were evaluated against the quality assessment questions above and were assigned a score based on the quality assessment scoring matrix presented in Table 5. A score of more than 3 qualifies a paper to remain included or accepted in this research review; otherwise, it is excluded.

TABLE 5. Quality assessment scoring criteria.

Quality Assessment Scoring Criteria	Score
The author(s) have provided a detailed, explicit, and clear understanding on the answers of the specific RQ.	High = H = 1
The author(s) have provided some and general explanation, but not detailed, explicit, and clear understanding on the answers to the specific RQ.	Medium = M = 0.5
The author(s) have provided no and very few technical details to the specific RQ.	Low = L = 0

F. DATA COLLECTION

After the quality assessment has been performed on the research papers, we will exclude all those that do not answer our research questions. Data extraction is implemented, which requires thorough analysis, identification, and collection of important data from each selected research paper after quality assessment. This returns a narrow list of papers with focused meta data that can be used for paper classification and further analysis. Table 6 presents the metadata of the information extracted from the qualified or primary research studies.

TABLE 6. Data Item collection form.

Data Fields	Description	Respected Research Question
Reference ID	It Provides the unique ID of each primary study	For documentation purpose
No. of Primary Studies	It provides the number of studies qualified after quality scoring scheme	RQ1
Year	Year of publication	RQ1
Publication Source	It Provides the name of publication venue and type of primary study and its publisher	RQ1
Databases	It provides the digital database name for each primary study.	RQ1
Objectives	It Provides the major objectives of FS implemented in each primary study.	RQ2
FS Approaches	It Provides the type of feature selection utilized in each primary study.	RQ2
FS Techniques	It Provides FS technique in its FS direction from the primary study.	RQ2
Datasets	It provides the IoT dataset used for analysis.	RQ3
Attacks	It presents the types of attacks detection or classification in each study.	RQ3
No. of Feature sets	It introduces the original features from dataset and list the number of features.	RQ3
No. of Instances	It introduces the network flow instances of datasets.	RQ3
Machine learning algorithm(s)	It provides the type of machine learning algorithms.	RQ4
Deep learning algorithms(s)	It provides the type of deep learning algorithms.	RQ4
Performance metrics	It provides the metrics used for FS validation.	RQ4
Compared with that of full features	It provides the result compared with that of full feature sets.	RQ4
compared with existing FS techniques	It provides the result compared with that of the-state-of-the-art FS techniques.	RQ4
Limitations	It describes the limitations of the proposed FS in the primary studies.	RQ5
Major Challenges	It introduces the major challenges for FS applied.	RQ5
Future Direction	It introduces the future direction for FS applied for attack classification for IoT security.	RQ5

With the objective in mind, we ensured that the current SLR study applied the Preferred Reporting Items for Systematic Review and Meta Analysis (PRISMA) (Page et al. 2021). The PRISMA is implemented in the current study as a measure to show detailed information concerning the total number of articles considered in the current study, as presented in Figure 2.

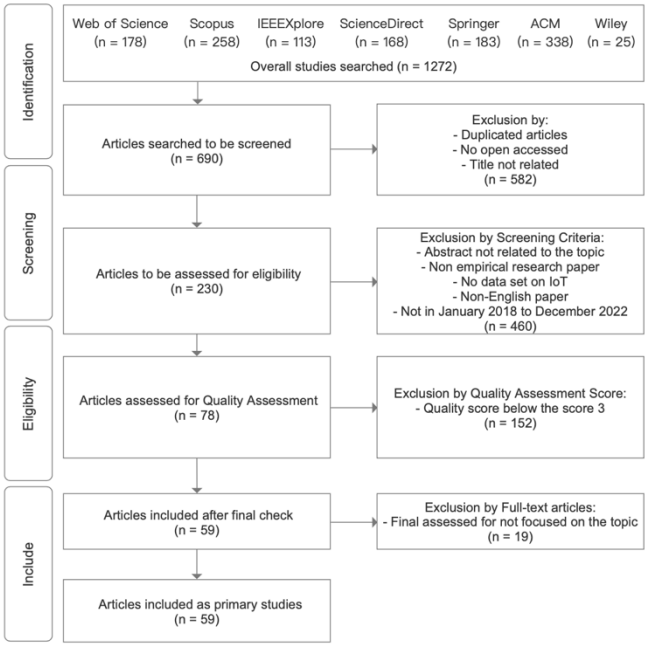

FIGURE 2. PRISMA Diagram.

G. DATA ANALYSIS

This subsection is the final step of the SLR methodology: data analysis to utilize the primary studies and the data extracted from each study to conduct data analysis in order to answer the six main research questions raised in this research study. The quality assessment report can be found in Table 7, which presents the primary studies, the main idea, and the quality score for each study.

3 Result and Analysis

In this systematic review, after the PRISMA process from initial collected 1272 studies in scientific databases (Web of Science, IEEE Xplore, Scopus, ScienceDirect, ACM, SpringerLink, and Wiley Online Library), we finally identified 63 papers as the primary studies for this systematic literature review, in order to find out the answers to the research questions proposed at the beginning of this literature review. We can refer to the Table 7, 8, and 9 for the brief information of the primary studies with corresponding quality scores. The classification of the primary studies based on the meta data extraction are provided and the finding results are presented in the part that follows.

TABLE 7. Primary studies from 2018 to 2020.

INDEX	TITLE	REF	PUBLISHER	TYPE	QUALITY ASSESSMENT SCORING					
					QA1	QA2	QA3	QA4	QA5	SCORE
PS01	Dimensionality Reduction for Machine Learning Based IoT Botnet Detection	(Bahsi et al., 2018)	Scopus	Conference	1	1	0.5	0.5	0.5	3.5

PS02	Machine Learning DDoS Detection for Consumer Internet of Things Devices	(Doshi et al., 2018)	IEEE	Conference	0.5	1	0.5	0.5	1	3
PS03	Deep Abstraction and Weighted Feature Selection for Wi-Fi Impersonation Detection	(Aminanto et al. 2018)	IEEE	Journal	1	1	1	1	0.5	4.5
PS04	Unsupervised Anomaly Based Botnet Detection in IoT Networks	(Nomm and Bahsi 2018)	IEEE	Conference	1	1	0.5	0.5	0.5	3.5
PS05	DEMISe: Interpretable Deep Extraction and Mutual Information Selection Techniques for IoT Intrusion Detection	(Parker et al. 2019)	Scopus	Conference	1	1	1	0.5	0.5	4
PS06	Hybrid Feature Selection Models for Machine Learning Based Botnet Detection in IoT Networks	(Guerra-Manzanares et al., 2019)	IEEE	Conference	1	1	1	0.5	0.5	4
PS07	Towards the Integration of a Post-Hoc Interpretation Step into the Machine Learning Workflow for IoT Botnet Detection	(Guerra-Manzanares et al., 2019)	IEEE	Conference	1	1	1	0.5	0.5	4
PS08	IoT botnet attack detection based on optimized extreme gradient boosting and feature selection	(Alqahtani et al., 2020)	Scopus	Journal	1	1	1	0.5	0.5	4
PS09	Averaged dependence estimators for DoS attack detection in IoT networks	(Baig et al. 2020)	ScienceDirect	Journal	0.5	1	0.5	0.5	0.5	3
PS10	Feature Selection Improves Tree-based Classification for Wireless Intrusion Detection	(Bhandari et al. 2020)	Scopus	Conference	1	1	1	0.5	0.5	4
PS11	Machine Learning-based Intrusion Detection for IoT Devices in Smart Home	(Li et al., 2020)	IEEE	Conference	1	0.5	1	0.5	0.5	3.5
PS12	Augmented whale feature selection for IoT attacks: Structure, analysis and applications	(Mafarja et al. 2020)	ScienceDirect	Journal	1	1	0.5	0.5	0.5	3.5
PS13	IoT malicious traffic identification using wrapper-based feature selection mechanisms	(M. Shafiq et al. 2020)	ScienceDirect	Journal	1	1	0.5	0.5	0.5	3.5
PS14	Implementing Lightweight IoT-IDS on Raspberry Pi Using Correlation-Based Feature Selection and Its Performance Evaluation	(Y.N. Soe et al. 2020)	Springer	Journal	1	1	0.5	1	0.5	4
PS15	Detecting botnet by using particle swarm optimization algorithm based on voting system	(Asadi et al. 2020)	ScienceDirect	Journal	1	1	1	1	0.5	4.5
PS16	Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection	(Al Shorman et al., 2020)	Springer	Journal	1	1	0.5	0.5	0.5	3.5
PS17	Machine Learning-Based IoT-Botnet Attack Detection with Sequential Architecture dagger	(Soe et al. 2020)	Scopus	Journal	1	1	0.5	1	0.5	4
PS18	Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks	(Davahli et al., 2020)	Springer	Journal	1	1	0.5	1	0.5	4

TABLE 8. Primary studies in 2021.

INDEX	TITLE	REF	PUBLISHER	TYPE	QUALITY ASSESSMENT SCORING					
					QA1	QA2	QA3	QA4	QA5	SCORE
PS19	IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization	(Fatani et al. 2021)	IEEE	Journal	0.5	1	1	0.5	0.5	3.5
PS20	Intrusion Detection System Using Machine Learning for Vehicular Ad Hoc Networks Based on ToN-IoT Dataset	(Gad et al., 2021)	IEEE	Journal	1	1	0.5	0.5	0.5	3.5
PS21	Comparing Machine Learning and Deep Learning for IoT Botnet Detection	(Gandhi and Li 2021)	IEEE	Conference	0.5	1	0.5	0.5	0.5	3
PS22	A Machine Learning Framework for Intrusion Detection System in IoT Networks Using an Ensemble Feature Selection Method	(Guo 2021)	Scopus	Conference	0.5	1	0.5	0.5	0.5	3
PS23	A Comparative Study of Machine	(Ismail et	IEEE	Conference	0.5	1	0.5	0.5	0.5	3

	Learning Models for Cyber-attacks Detection in Wireless Sensor Networks	al. 2021)								
PS24	Three-layer hybrid intrusion detection model for smart home malicious attacks	(Shi et al., 2021)	ScienceDirect	Journal	0.5	1	1	0.5	0.5	3.5
PS25	An Agile Approach to Identify Single and Hybrid Normalization for Enhancing Machine Learning-Based Network Intrusion Detection	(Siddiqi and Pak 2021)	IEEE	Journal	0.5	0.5	1	0.5	0.5	3
PS26	Efficiency Enhancement of Intrusion Detection in Iot Based on Machine Learning Through Bioinspire	(Samdekar et al., 2021)	IEEE	Conference	0.5	0.5	1	0.5	0.5	3
PS27	CorrAUC: A Malicious Bot-IoT Traffic Detection Method in IoT Network Using Machine-Learning Techniques	(Shafiq et al. 2021)	IEEE	Journal	1	1	1	0.5	0	3.5
PS28	Toward Design of an Intelligent Cyber Attack Detection System using Hybrid Feature Reduced Approach for IoT Networks	(Kumar et al., 2021)	Springer	Journal	1	1	0.5	1	0.5	4
PS29	Classifier Performance Evaluation for Lightweight IDS Using Fog Computing in IoT Security	(Khater et al. 2021)	Scopus	Journal	1	0.5	0.5	0.5	1	3.5
PS30	Effective combining of feature selection techniques for machine learning-enabled IoT intrusion detection	(Rahman et al. 2021)	Springer	Journal	0.5	1	1	0.5	0.5	4
PS31	Fault-tolerant AI-driven Intrusion Detection System for the Internet of Things	(Medjek et al. 2021)	Web of Science	Journal	0.5	1	0.5	0.5	0.5	3
PS32	Toward lightweight intrusion detection systems using the optimal and efficient feature pairs of the Bot-IoT 2018 dataset	(Ozer et al., 2021)	Springer	Journal	1	1	0.5	0.5	0.5	3.5
PS33	An effective genetic algorithm-based feature selection method for intrusion detection systems	(Halim et al. 2021)	Web of Science	Journal	1	1	0.5	0.5	0.5	3.5
PS34	IoT Dataset Validation Using Machine Learning Techniques for Traffic Anomaly Detection	(Vigoya et al. 2021)	Scopus	Journal	0.5	1	0.5	0.5	0.5	3

TABLE 9. Primary studies in 2022.

INDEX	TITLE	REF	PUBLISHER	TYPE	QUALITY ASSESSMENT SCORING					
					QA1	QA2	QA3	QA4	QA5	SCORE
PS35	Anomaly Detection for Internet of Things Cyberattacks	(Alanazi and Aljuhani 2022)	Scopus	Journal	1	0.5	0.5	0.5	0.5	3
PS36	A discrete time-varying greywolf IoT botnet detection system	(Alazab 2022)	ScienceDirect	Journal	1	1	1	0.5	0.5	4
PS37	Intelligent IoT-BOTNET Attack Detection Model with optimized Hybrid Classification Model	(Bojarajulu et al., 2022)	ScienceDirect	Journal	0.5	1	0.5	1	0.5	3.5
PS38	Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique	(Disha and Waheed 2022)	Springer	Journal	1	1	0.5	1	0.5	4
PS39	Advanced feature extraction and selection approach using deep learning and aquila optimizer for iot intrusion detection system	(A. Fatani et al. 2022)	Scopus	Journal	1	1	1	0.5	0.5	4
PS40	Machine and deep learning amalgamation for feature extraction in Industrial Internet-of-Things	(Jayalaxmi et al. 2022)	Web of Science	Journal	1	0.5	1	0.5	0.5	3.5
PS41	A new Intelligent Satellite Deep Learning Network Forensic framework for smart satellite networks	(Koroniotis et al., 2022)	ScienceDirect	Journal	1	1	0.5	0	0.5	3
PS42	Machine learning-based early detection of IoT botnets using network-edge traffic	(A. Kumar et al. 2022)	ScienceDirect	Journal	1	0.5	0.5	0.5	0.5	3
PS43	An intellectual intrusion detection system using Hybrid Hunger Games Search and Remora Optimization Algorithm for IoT wireless networks	(R. Kumar et al., 2022)	ScienceDirect	Journal	1	1	0.5	1	0.5	4
PS44	IoT information theft prediction using ensemble feature selection	(Leevy et al. 2022)	Springer	Journal	1	1	0.5	0.5	0.5	4

PS45	Decision Tree with Pearson Correlation-based Recursive Feature Elimination Model for Attack Detection in IoT Environment	(Padmashree and Krishnamoorthi 2022)	Scopus	Journal	0.5	1	0.5	0.5	0.5	3
PS46	Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system	(Ravi et al., 2022)	ScienceDirect	Journal	1	1	0	0.5	0.5	3
PS47	A feature selection-based method for DDoS attack flow classification	(Zhou et al. 2022)	ScienceDirect	Journal	1	1	0	0.5	0.5	3
PS48	A bio-inspired hybrid deep learning model for network intrusion detection	(Moizuddin and Jose 2022)	Web of Science	Journal	1	1	0.5	0.5	0.5	3.5
PS49	IoT intrusion detection technology based on Deep learning	(Cao et al. 2022)	IEEE	Conference	1	1	0.5	0.5	0	3
PS50	Lightweight Internet of Things Botnet Detection Using One-Class Classification	(Malik et al. 2022)	Scopus	Journal	1	1	0	0.5	0.5	3
PS51	Analyzing IoT Attack Feature Association with Threat Actors	(Shafiq et al. 2022)	Wiley	Journal	1	1	0.5	0	1	3.5
PS52	CHAMELEON: Optimized feature selection using particle swarm optimization and ensemble methods for network anomaly detection	(Chohra et al. 2022)	Web of Science	Journal	1	1	0.5	1	1	4.5
PS53	Injection attack detection using machine learning for smart IoT applications	(Gaber et al., 2022)	Web of Science	Journal	1	1	0.5	1	0.5	4
PS54	A Distributed Intrusion Detection System using Machine Learning for IoT based on ToN-IoT Dataset	(Gad et al. 2022)	Scopus	Journal	0.5	1	0.5	1	0.5	3.5
PS55	ML-Based IDPS Enhancement with Complementary Features for Home IoT Networks	(Illy et al. 2022)	IEEE	Journal	1	1	0	0.5	0.5	3
PS56	NFDLM: A Lightweight Network Flow based Deep Learning Model for DDoS Attack Detection in IoT Domains	(Saurabh et al. 2022)	IEEE	Conference	1	1	0	0.5	0.5	3
PS57	FI-PCA for IoT Network Intrusion Detection	(Abdulkareem et al. 2022)	IEEE	Conference	1	1	0.5	0.5	0.5	3.5
PS58	Machine Learning based IoT-BotNet Attack Detection Using Real-time Heterogeneous Data	(Ahmed and Tjortjis 2022)	IEEE	Conference	0.5	1	0.5	0.5	0.5	3
PS59	Examining the Suitability of NetFlow Features in Detecting IoT Network Intrusions	(Awad et al. 2022)	Scopus	Journal	1	1	0.5	1	0.5	4
PS60	VMFCVD: An Optimized Framework to Combat Volumetric DDoS Attacks using Machine Learning	(Prasad and Chandra 2022)	Springer	Journal	1	1	0.5	1	0.5	4
PS61	A GPU-based machine learning approach for detection of botnet attacks	(Motylinski et al. 2022)	Web of Science	Journal	0.5	1	0.5	0.5	0.5	3
PS62	Fast, Lightweight IoT Anomaly Detection Using Feature Pruning and PCA	(Carter et al., 2022)	ACM	Conference	0.5	1	0.5	0.5	0.5	3
PS63	An Effective Feature Selection Model Using Hybrid Metaheuristic Algorithms for IoT Intrusion Detection	(Kareem et al. 2022)	Scopus	Journal	0.5	1	0.5	1	1	4

We used VOSviewer to conduct a brief data analysis on the key terms found in the title and abstract fields of 63 primary studies in order to determine the key aspects and relationships between them. As shown in Figure 3, there are three clusters with red, blue, and green color highlighted. For the red flag items, "internet", "thing", "device", "iot device" means the term "internet of things" or IoT in academics and industry; additionally, "attack," "detection," and "botnet" mean security terms; thus, the output of the above means IoT security. Among the green flag terms are "dataset," "feature," "technique," "algorithm," and "deep learning," with the additional words "intrusion detection system" and "intrusion detection," or IDS in academics and industry. As for the blue cluster, one term is "experimental result," which means the experimentation and validation of the proposed system. Therefore, we can conclude that the brief background and methodology used in primary studies are valid, and we can argue that the primary studies are topic-focused, and the review based on these studies is valid.

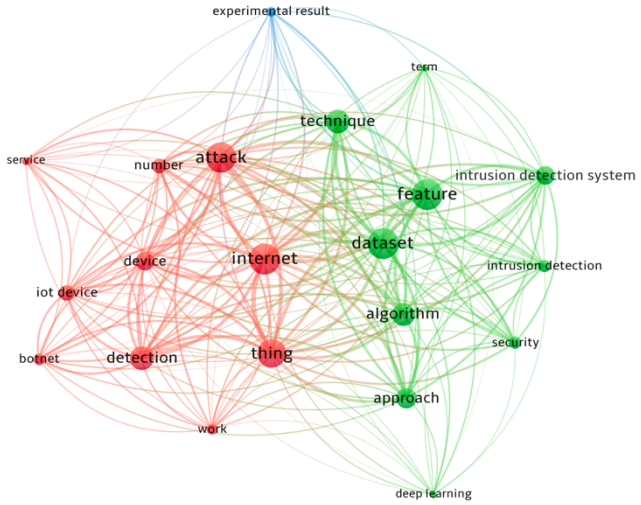

FIGURE 3. The key items clustering and network visualization of primary studies by VOSviewer

RQ1: What is current situation of FS approaches applied for machine learning or deep learning-based attack classification model for IoT security?

RQ1.1 What is the trend of the studies that applied FS for IoT security in recent five years?

After learning from the studies, two main factors contribute to this increasing trend. The first is the growing development of AI or machine learning-based cyber security models in IoT networks, which is driven by the limitations of traditional security models, which can only detect and classify known attacks and are ineffective at recognizing unknown or zero attacks. Since the massive IoT devices deployed all over the world can create any possible variant of a known attack, any type of zero-day attack could threaten the critical infrastructure of IoT networks. Furthermore, followed by the second factor, the high performance of a security model using machine learning algorithms is highly dependent on the representative features needed to map to any type of attack. While the optimal features can be selected using various feature selection techniques in various manners, In the last five years, appropriate feature selection methods and machine learning algorithms have contributed to more effective and efficient security models in the IoT.

According to Figure 4 and the discussion above, FS research in IoT security has been expanding since it started in 2020, and it is still a hot issue among researchers in the field of IoT security.

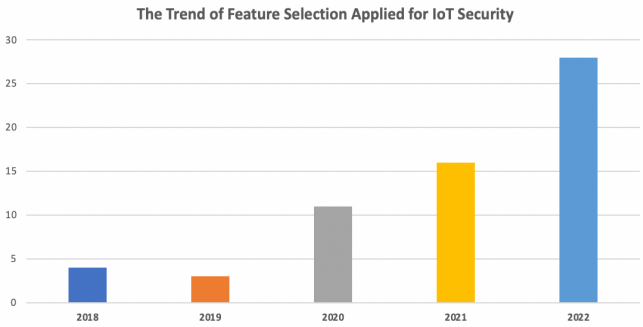

FIGURE 4. The trend of FS applied for IoT security over the years

RQ1.2 What are the distributions of the studies according to databases and publishers?

As we can see from Figure 5, the distribution of the number of primary studies in each journal is presented. Three groups of journals can be classified; the journals that published equal or more than five primary studies are "Computer and Security" and "Sensors," which are in the first group. The second group of journals has equal to or more than three published primary studies involving "Future Generation Computer Systems," "Computers and Electrical Engineering," and "IEEE Access." Besides, we can see the third group of journals below the three primary studies, involving "Knowledge-based Systems," "Electronics," "International Journal of Distributed Sensor Networks," and so on. Based on Figure 5 and the explanation above, we can identify which journals focus more on the topic of feature selection in IoT security and which journals can be future candidates for paper publication in the same domain.

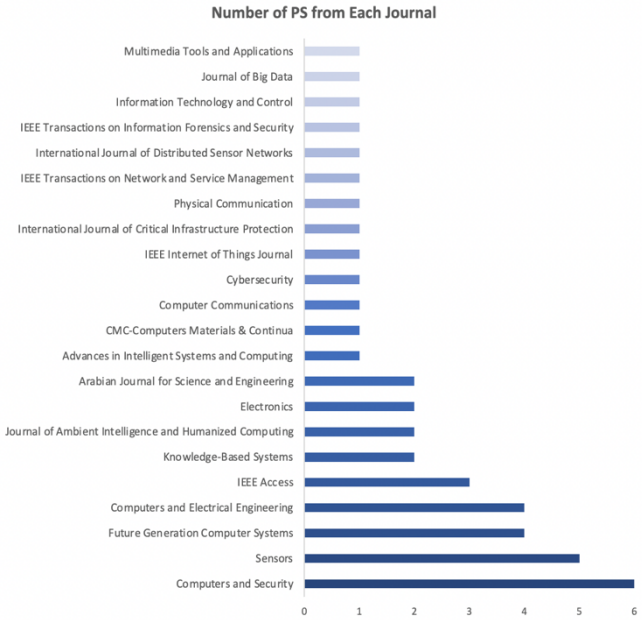

FIGURE 5. The number of PS from each journal

RQ1.3: What are the main objectives of applying feature selection for IoT security?

Generally, the purpose of feature selection in machine learning is to identify a subset of the most relevant features from the original feature set, which can then be used to train a model. This can improve the performance of the model and reduce overfitting, as well as make the model more interpretable by identifying the most important factors that contribute to the outcome. Additionally, it can also help in reducing the computational cost and time of training a model.

Figure 6 presents the distribution of the primary studies according to purpose over the past five years. The primary studies are classified by five categories: improving performance, reducing complexity, preventing overfitting, model interpretability, and comparison.

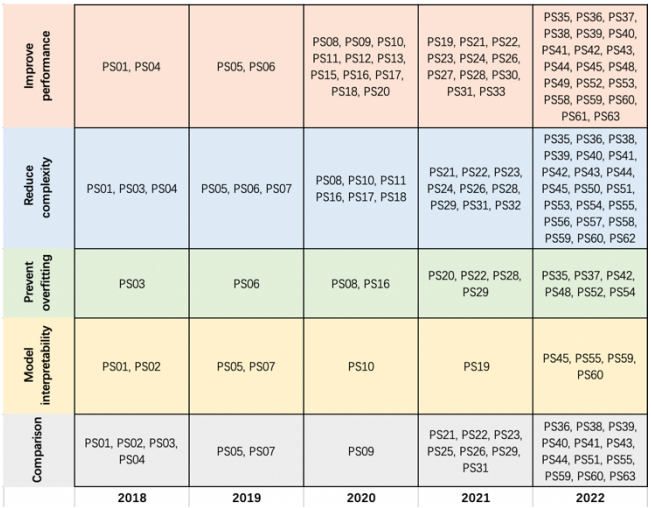

FIGURE 6. The purpose of feature selection for IoT security in recent years

Many purposes are described with various terms; thus, the rules for categorising studies are as follows: Studies focusing on an effective attack detection model, improving the detection rate, achieving high accuracy, lowering the misclassification rate, and so on are considered first class. Studies that use feature selection techniques to reduce model cost and time, model prediction time, and model efficiency in resource-constrained IoT systems are considered to fall under the category of reducing complexity. While the works emphasising overfitting problems are classified as being of the third class, some studies focusing on the model's interpretability after using feature selection are classified as being of the fourth class. Finally, some studies that employed various feature selection techniques with various machine learning algorithms for effective models are classified as the last class. Since some studies employed feature selection to improve the performance and reduce the complexity of the proposed model as well, they fall into multiple categories with corresponding purposes.

Based on the data extracted in Figure 6, we can see that the primary studies mostly fell into improving performance (45 out of 62), which means the performance metrics on attack detection and classification are the most concerning points for the academics. Followed by the second category is reducing complexity (42 out of 62), which shows that producing lightweight models is also vital, particularly for resource-limited IoT networks in recent years. The interesting point is that more studies are focusing on developing lightweight models with significant performance metrics, rather than focusing solely on performance. Since there are so many feature selection techniques that can be employed before training learning models, many studies have tried to employ various feature selection techniques combined with different machine learning and deep learning algorithms to build state-of-the-art models, particularly in the past two years.

Feature selection can prevent overfitting and contribute to a model that can be generalised to the same scenarios; thus, the implementation of the most primary studies can address overfitting to some extent; however, not many studies emphasise overfitting in their studies (14 out of 62). Few studies (10 out of 62) argued for model interpretability in their proposed feature selection approaches; however, there is an increasing trend toward the implementation of model interpretability in 2022.

RQ2: What are the FS methods and techniques applied on attack classification models for IoT security?

RQ2.1: What are the main types of feature selection approaches applied for attack classification model?

In theory, based on the mechanism of selecting the optimal features, feature selection involves filter, wrapper, embedded and hybrid methods. The filter method uses a statistical test to evaluate the importance of each feature and select a subset of the most relevant features. Examples include chi-squared test, mutual information, and correlation-based feature selection. Wrapper method employs a specific machine learning algorithm to evaluate the performance of different feature subsets and select the best subset based on a performance metric. Examples include forward selection, backward elimination, and recursive feature elimination. Embedded methods use a specific machine learning algorithm to select features as part of the model training process. Examples include Lasso and Ridge regression, which have a built-in feature selection process through the regularization term.

Hybrid methods combine the strengths of multiple feature selection methods to achieve better performance. Examples include combining filter and wrapper FS to construct a more powerful feature selection method. Ensemble approaches combine multiple feature selection techniques that use the same, different parts, or even the entire dataset, and then make decisions using algorithms such as interaction or majority voting based on the candidate feature list from multiple FS techniques. Some studies tried to explore various FS techniques to obtain the model with the best performance; we classify these types of studies into the comparison category.

Table 10 shows the distribution of primary studies over various feature selection approaches applied for IoT security. We found there are six types of feature selection methods used to identify the most informative features before building a security model. The result shows that the filter-based feature selection method is the top FS approach used in attack classification for IoT security. Most of the studies used filter FS by using various statistical techniques to select the feature sets based on two rules: one was to eliminate the features that have high correlation with each other, while the other was to obtain the features that have high correlation with the target classes for model training and prediction (Soe et al. 2020). Furthermore, the feature sets of the two rules may overlap, so appropriate threshold settings for the correlation are vital for determining the final feature sets for building high performance models (Saurabh et al. 2022). Moreover, considering most of the datasets for IoT security are large volumes of network datasets with big data characteristics (Koroniotis et al., 2022), while most IoT systems are equipped with limited computational and storage resources, filter FS was chosen in the majority of research because it is computationally light, which fulfils the need for making the overall model lightweight for IoT networks (Awad et al. 2022).

Wrapper FS was the primary study's second method, followed by the Filter method. Wrapper FS selects the optimal feature subsets rather than identifying individual features implemented by the filter method. Furthermore, wrapper works with machine learning algorithms to evaluate feature subsets based on the performance metrics required by the model. Since the feature subsets that are evaluated and identified are performance-oriented, the wrapper method can achieve better performance and contribute to higher generalization capability than filter FS; however, its demerit is that it requires much more computational resources than filter FS, which can be a significant concern for resource-constrained IoT devices. (Vigoya et al. 2021) employed a recursive feature elimination method to evaluate the feature subsets among all the features of the dataset exhaustively. It uses a specific machine learning algorithm to iteratively remove the least important features based on a performance metric until the desired number of features is reached. Because the search space of feature subsets is extremely large for IoT datasets, which contain too many features with a large number of instances in network flow records, heuristic algorithms were mostly applied with learning algorithms using the wrapper method. For example, (Al Shorman et al., 2020) proposed the Augmented Whale Optimization (AWO) algorithm search algorithm to find the best subset of features with OCSVM as a learning algorithm to reach the best False Positive Rate (FPR). (Halim et al. 2021) proposed an enhanced genetic algorithm (GA) that outperformed recursive feature elimination (RFE), sequential feature selection (SFS), and correlation-based feature selection (CFS). (Chohra et al. 2022) propose the Particle Swarm Optimization (PSO) algorithm by combining both swarm intelligence and ensemble learning techniques to select the optimal settings for feature subsets and hyperparameters of the model. However, most of the work needs to focus more on the balance between the searching efficiency in feature space and performance metrics such as accuracy or F1-score.

Hybrid FS method is a combination of multiple feature selection methods to achieve better performance. The idea behind hybrid methods is to leverage the strengths of multiple feature selection techniques to achieve a more accurate and robust solution. For example, a hybrid method might first use a filter method to reduce the number of features based on a statistical test, then use a wrapper method to further refine the feature set based on the performance of a machine learning model. For example, (Guerra-Manzanares et al., 2019) proposed a hybrid FS by first using filter FS and then employing wrapper FS to identify the optimal feature subset based on the features obtained by filter FS. Similarly, (Li et al., 2020) employed a two-step hybrid FS using the Kolmogorov-Smirnov (KS) test to select important features from the candidate feature set, then used Pearson correlation to remove redundant features to obtain the informative features. Moreover, (Padmashree and Krishnamoorthi 2022) (Zhou et al. 2022) (Malik et al. 2022) implemented the same idea of hybrid FS to identify optimal feature subsets by combining the strengths of multiple feature selection methods, resulting in improved performance and robustness.

Alternatively, ensemble FS is another type of hybrid feature selection method that combines the results of multiple feature selection techniques to produce a final feature set. In ensemble feature selection, multiple feature selection methods are applied to the same data or subsets of the same data, and the results are combined to form a final feature set. This combination can be performed in a number of ways, such as by taking the union or intersection of the feature sets produced by each method or using a voting scheme where each method votes for the features it considers important. For example, (Kumar et al., 2021) combined the correlation coefficient, random forest mean decrease accuracy, and gain ratio techniques to obtain three different feature sets, which were then combined using a suitably designed mechanism (AND operation) to obtain single optimized feature sets. Similarly, (Guo 2021) conducted a combination of five FS techniques: information gain (IG), gain ratio (GR), chi-squared (CS), Pearson correlation coefficient (PCC), and symmetric uncertainty (SU), then, based on majority voting, the features that appear more than three times in the five subsets are extracted as the final feature sets for the next step of model training. However, it is important to carefully consider the choice of FS techniques to be combined, as well as the specific details of the combination, in order to achieve the best results.

In embedded FS, the feature selection process is incorporated into the optimization objective of the machine learning model. Researchers (Doshi et al., 2018), (Shi et al., 2021), and (Disha and Waheed 2022) employed random forest as the algorithm for selecting the feature sets based on Gini Impurity Scores, then use multiple machine learning algorithms respectively to implement attack classification. Embedded methods can produce models with improved interpretability, as the most important features are automatically identified and selected as part of the training process. Alternatively, some studies that used different techniques to obtain feature sets are classified as others. For example, (Ozer et al., 2021) iteratively selected and evaluated only 2 features, named feature pairs, from the original 12 features using multiple machine learning algorithms to build and evaluate a lightweight model, respectively. (Ravi et al., 2022) bypassed the feature selection process and directly used deep learning algorithms named RNN, LSTM, and GRU to extract and select features for model building. However, (Carter et al., 2022) employed principal component analysis (PCA) to reduce the features, and the results showed a significant advantage to using PCA for both traditional machine learning algorithms (SVM) and neural network-based deep learning algorithms for anomaly detection.

Considering there are various FS approaches and so many specific techniques that can be selected for various IoT datasets to build models, some primary studies implement comparison using various FS methods to evaluate and obtain the best feature sets. For example,

(Parker et al. 2019), (Guerra-Manzanares et al., 2019) implemented an FS comparison of filter, wrapper, and hybrid methods, and the results showed hybrid produced the best accuracy results. Moreover, (Samdekar et al., 2021) evaluated multiple FS methods involving filter and wrapper using one machine learning algorithm named SVM to determine the best model. (Gaber et al., 2022) compared filter and wrapper FS, which involved constant removal and constant removal combined with recursive feature elimination. Furthermore, (Illy et al. 2022) evaluated the impact of various selected features and various machine learning algorithms on the accuracy of the models, and the result showed that the accuracy of the detection model depends more on the feature sets than the ML methods.

TABLE 10. Primary studies based on FS methods.

FS Types	# of PS	Primary Studies
Filter	18	PS01, PS08, PS09, PS10, PS14, PS17, PS20, PS21, PS29, PS37, PS41, PS42, PS54, PS56, PS57, PS58, PS59, PS61
Wrapper	13	PS03, PS12, PS13, PS16, PS19, PS33, PS34, PS36, PS39, PS43, PS48, PS52, PS63,
Embedded	3	PS02, PS24, PS38,
Hybrid	10	PS06, PS11, PS15, PS18, PS27, PS45, PS47, PS49, P50, PS60
Ensemble	6	PS04, PS22, PS23, PS28, PS35, PS44,
Comparison	10	PS05, PS07, PS25, PS26, PS30, PS31, PS40, PS51, PS53, PS55
Others	3	PS32, PS46, PS62

We investigated the trends further by looking at the distribution over the years. We can see in Figure 7 that there were limited studies for each type of FS in the first two years because most of the public IoT datasets were generated in 2018. Then, over the last three years, there has been a consistent use of filter, wrapper, and hybrid FS because researchers proposed various machine learning-based frameworks for various objectives, such as performance metrics-oriented models, lightweight models, or maintaining the balance between two objectives. Embedded FS was introduced from 2021 to 2022, since AI interpretability or model transparency had been a concern for academics in recent years. Furthermore, other types of FS were also introduced in the past two years. It's worth noticing that FS comparison was consistently conducted over the past five years, because many various FS techniques with different methods can be used to identify final features, moreover, many factors such as the objectives of the proposed models, characteristics of datasets, and classifiers trained by various learning algorithms can affect the results of the proposed FS methods.

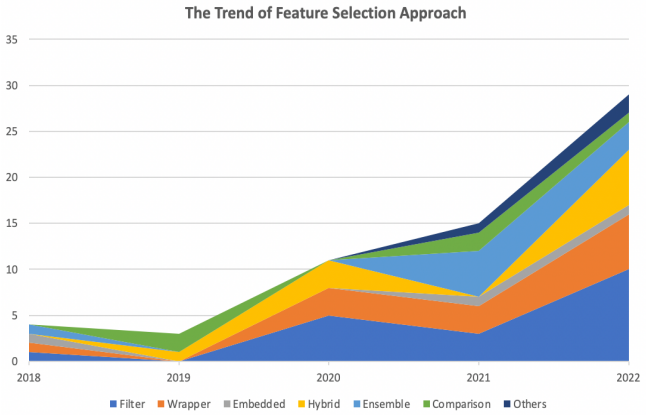

Figure 7. The trend of feature selection approach

RQ2.2: What are the techniques for each type of feature selection applied for IoT security models?

Based on the FS approaches, we further investigated the specific techniques utilized based on each FS category. We can see in Table 11 the techniques employed in the primary studies.

In filter FS, IG, PCC, and CS techniques were widely used as feature ranking techniques to identify final feature sets. The second tier used by primary studies to select features was MI and Fisher's score. The terms GR, SHAP, and PFI were only used once. One point that needs to be mentioned is that the techniques used in filter FS were also utilized in other types of FS, such as PCC, IG, CS, GR, and MI, and were widely used as one step of feature selection in hybrid, ensemble, and comparison modes, particularly for PCC, which was the most commonly employed in the primary studies. PCC is a measure of the linear relationship between two continuous variables. In wrapper FS, the techniques categorized in the table are searching algorithms searching for the optimal feature subsets. We can see that most of the searching techniques are based on heuristic algorithms, which are more efficient for handling large IoT datasets with high-dimensional feature spaces. In hybrid FS, the combination of PCC as the feature ranking technique and RFE as the searching technique was widely used.

TABLE 11. Primary studies based on FS techniques.

FS Types	Techniques/ Algorithms	# of PS	Primary Studies
Filter	Fisher's score	2	PS01, PS08
	CS	5	PS09, PS20, PS21, PS42, PS54
	IG	6	PS09, PS37, PS41, PS57, PS58, PS59,
	GR	1	PS09,
	SHAP	1	PS10,
	PCC	5	PS14, PS17, PS54, PS56, PS59
	MI	2	PS29, PS56,
	PFI	1	PS61
Wrapper	D-FES	1	PS03
	AWO	1	PS12
	Bijective soft set	1	PS13
	GWO	2	PS16, PS48
	TSO	1	PS19
	GA	1	PS33

	RFE	1	PS34
	GWO	1	PS36
	AQU	1	PS39
	HGS	1	PS43
	GTO	1	PS63
	PSO	1	PS52
Embedded	Gini Impurity	2	PS02, PS38
	VIM	1	PS24
Hybrid	KST+PCC	1	PS11
	cooperative game theory and Shipley value	1	PS15
	GA+GWO	1	PS18
	CorrAUC+TOPSIS and Shannon entropy	1	PS27
	PCC+RFE	1	PS47, PS50, PS60
	RFE+PCC	1	PS45, PS49
Ensemble	Hopkins + Variance, then based on Entropy	1	PS04
	IG+GR+CS+PCC+SU, then majority voting	1	PS22
	PCC+MI, then the threshold respectively	1	PS23
	PCC, RF and GR, then intersection	1	PS28
	SVM, DT and NB, then top ranked	1	PS30
	RF+PCC, then intersection	1	PS31
	DT+ET+RF+XGB, then then top ranked	1	PS35
	(WRAPPER+CLS+PCC) + AE, then combined	1	PS40
	IG+GR+CS, then top ranked	1	PS44
Comparison	Filter: MI and Hybrid: MI+J48	1	PS05
	Filter: Fisher's score, PCC; Wrapper: SFFS, SBFE; and Hybrid: Filter + Wrapper	1	PS06
	Filter: PCC and Wrapper: SFFS+DT	1	PS25
	CS, ET, FA, and PCA	1	PS26
	IG, CS, and EFS+DT	1	PS51
	Constant Removal and that with RFE	1	PS53
	Manually selecting and evaluating each type of features	1	PS55
Others	Feature Pairs	1	PS32
	Deep learning algorithms, RNN, LSTM, and GRU	1	PS46
	PCA	1	PS42

RQ3. What are the characteristics of related factors for FS methods for IoT security?

RQ3.1: What are the IoT datasets as the benchmark by the studies when applying FS method?

Since the IoT datasets with representative information can be the benchmark for validation of the attack detection and classification models, Table 12 describes a set of publicly used IoT datasets with basic information such as the year created, the number of features, the total number of instances, and the mapping of each dataset to the primary studies. From the visualized figure 8, we can argue that BoT-IoT (19 out of 63) was the mostly used datasets among the studies, following by N-BaIoT (12 out of 63), TON-IoT (4 out of 63), AWID (4 out of 63), and IoT-23 (3 out of 63). Most of other studies were investigated by just one study except by MedBIoT which was created for medical industry.

TABLE 12. IoT datasets mapping to primary studies along with some basic information

Datasets	Year	# of Features	# of Instances	# of PS	Primary Studies
BoT-IoT	2019	45	72,000,000	19	PS09, PS25, PS26, PS27, PS28, PS32, PS33, PS39, PS41, PS44, PS45, PS48, PS49, PS51, PS56, PS57, PS60, PS61, PS63
TON-IoT	2021	44	22,800,064	6	PS20, PS22, PS38, PS54, PS58, PS59
N-BaIoT	2018	115	7,009,269	12	PS01, PS04, PS06, PS07, PS08, PS12, PS14, PS16, PS21, PS07, PS36, PS60
AWID	2016	156	2,578,524	4	PS10, PS18, PS30, PS03
IoT-23	2020	23	266,910	3	PS21, PS35, PS42
IoT-Zeek	2022	13	2,043,602	1	PS52
Anthi 2019 IoT	2019	14	101,583	1	PS55
Kitsune	2018	115	764,137	1	PS21
MedBIoT	2020	100	842,674	2	PS21, PS50
WSN-DS	2016	18	374,661	1	PS23
DAD	2021	14	67,848	1	PS34
Water Tank / Gas Pipeline	2014 / 2013	23 / 26		1	PS40
SDNIoT	2020		210,000	1	PS46

Among the datasets, Bot-IoT dataset was the mostly used and evaluated dataset by the primary studies, and it has the largest volume of instances (72,000,000) for studies to create and evaluate attack detection models. Followed by N-BaIoT which has more than one hundred features (115) to describe the information of each instance, was the secondly studied by researchers. Six studies worked on TON-IoT created in 2021 with its characteristics of heterogenous sources and multiple layers, to create attack detection models.

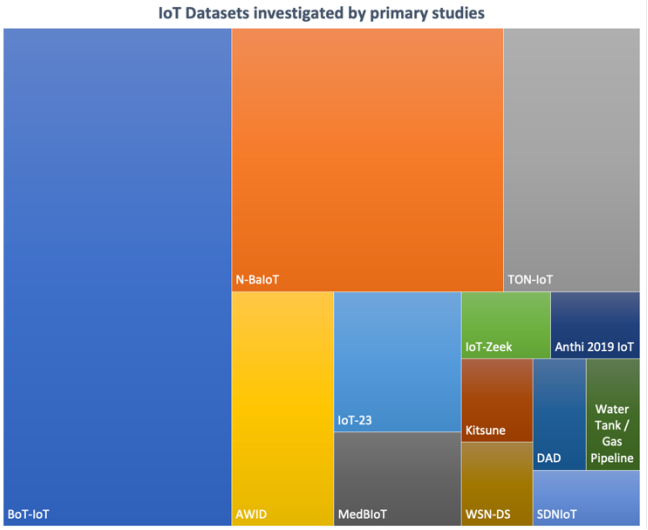

Figure 8. The IoT Datasets investigated by primary studies

RQ3.2: What are mapping of datasets and attacks to various FS methods among the studies?

Since feature selection is the key process of data processing in the machine learning pipeline, the characteristics of the data, particularly the attacks to be detected and classified, are highly related to the FS method to be utilized. Therefore, we investigated the IoT datasets used in the primary studies and what types of FS methods were applied to each of them. Table 12 shows the primary studies based on various IoT datasets, the attack classes in the datasets, and the corresponding FS methods. One point that needs to be mentioned is that the primary studies classified in FS comparison are reclassified into specific FS methods in this section. Since many FS methods are involved in FS comparison, one primary study may be classified into two FS categories. For example, PS25 has implemented both the filter and wrapper methods for comparison; thus, PS25 is added to both the filter and wrapper FS in this table, and the same rule applies to other primary studies that have implemented FS comparison, so the total number of PS in this table is more than 63.

TABLE 12. Primary studies based on various IoT datasets.

Dataset	Attack Classes	FS	# of PS	Primary Studies
BoT-IoT	DoS, DDoS, Reconnaissance, Information Theft	Filter	24	PS09, PS25, PS26, PS41, PS51, PS56, PS57, PS61
		Hybrid		PS27, PS45, PS49, PS60,
		Ensemble		PS28, PS44
		Wrapper		PS25, PS26, PS33, PS39, PS63, PS48, PS51
		Others		PS32
TON-IoT	Scanning, XSS, DoS, DDoS, Backdoor, Injection, Password Cracking, MITM, Ransomware	Filter	6	PS20 , PS54, PS58, PS59
		Ensemble		PS22
		Embedded		PS38
N-BaIoT	Botnet, Gafgyt, Mirai	Filter	13	PS01, PS07, PS08, PS14, PS21,
		Wrapper		PS07, PS12, PS16, PS36
		Hybrid		PS06, PS07, PS60
		Ensemble		PS04
AWID	Injection, Flooding, Impersonation	Filter	4	PS10
		Hybrid		PS18
		Ensemble		PS30
		Wrapper		PS03
IoT-23	Mirai, Torii, Trojan, Gagfyt, Kenjirro, Hakai, Hajime, Okiru	Filter	3	PS21, PS42
		Ensemble		PS35
IoT-Zeek	Malware	Wrapper	1	PS52
Anthi 2019 IoT	Scanning, DoS, IoT-toolkit, and MITM	Others	1	PS55
Kitsune	Mirai, SYN DoS, SSDP Flood etc.	Filter	1	PS21
MedBIoT	Mirai, Bashlite, and Torii	Filter	2	PS21
		Hybrid		PS50
WSN-DS	Grayhole, Blackhole, Flooding, and TDMA scheduling	Ensemble	1	PS23
DAD	Duplication, Interception and Modification on the MQTT message	Wrapper	1	PS34
Water Tank and Gas Pipeline	Naive Malicious Response Injection, Complex Malicious Response Injection, Malicious State Command Injection, Malicious Parameter Command Injection, Malicious Function Code Injection, Denial of Service, Reconnaissance	Ensemble	1	PS40
SDNIoT	DoS, DDoS, Port	Others	1	PS46

	Scanning, OS Fingerprinting, and Fuzzing			
Private Data sets	Ransomware, Cryptominer,	Others	4	PS62
	SYN flood, LowRate, Mirai	Hybrid		PS47
	DOS, DDoS, reconnaissance, exploits, fuzzes, backdoors, Generic, DoS, Shellcode	Filter		PS37
	Routing Protocol for Low-power and Lossy Networks (RPL) Decreased Rank (DR) Sinkhole (SH) Blackhole (BH) Selective forwarding (SF) Hello Flooding (HF) Version number (VN)	Ensemble		PS31

As we can see, the dataset named BoT-IoT was produced by (Koroniotis et al. 2019) and was mostly studied by researchers (24 out of 63). As the FS method to implement feature selection for attack classification, filter and wrapper FS methods dominate. because the data set contains a large amount of data concerning normal and typical abnormal IoT activities. Many researchers have worked on this data set as an IoT scenario benchmark to validate their proposed intrusion detection systems. followed by the N-BaIoT dataset, which was produced by (Meidan et al. 2018) as public IoT dataset for researchers working on the models for IoT security. The data was collected with real commercial IoT devices involved, and it includes traffic information for Gafgyt and Mirai, two of the most well-known IoT-based botnet attacks. Many researchers (13 out of 63) employed filter and wrapper FS to figure out the most appropriate feature sets for the attacks in the dataset in recent years.

TON-IoT was produced recently by (Moustafa 2021) and was also an IoT-specific dataset. The datasets were collected from heterogeneous sources, including telemetry from IoT devices, networking flows, and system logs of the operating system, across multiple layers such as edge, fog, and cloud layers. Some studies have implemented various FS methods to identify the features for various attacks in this data, while the filter method was the most commonly used by studies compared with ensemble and embedded FS. AWID was produced by (Kolias et al. 2016), focusing on wireless data with the Wi-Fi protocol as one of the communication protocols for IoT networks. Except for embedded FS, other methods were used to identify the most suitable feature for classification models.

Similarly, there are increasingly more IoT datasets created and studied by researchers using various FS methods for various attack classifications. For example, (Gandhi and Li 2021) and (A. Kumar et al. 2022) implemented filter FS for IoT datasets named IoT-23, Kitsune, and MedBIoT, respectively. Filter mode can make the feature selection process more lightweight, and to be specific on the technique, Chi-square (CS) was used to identify independent features that can be informative for the models to learn the pattern of attack classes. Similarly, ensemble FS was used by (Ismail et al. 2021), (Alanazi and Aljuhani 2022), and (Jayalaxmi et al. 2022) on datasets named, WSN-DS, IoT-23, and Water Tank and Gas Pipeline, to combine the individual capabilities of FS techniques. Moreover, some researchers created their own datasets for specific purposes. For example, (Medjek et al. 2021) created the dataset focusing on routing type attacks, and proposed ensemble FS mode by combining random forest and Pearson correlation, followed by interaction to select the features.

RQ4: What are the verification methods to evaluate the effectiveness of proposed FS approaches?

RQ4.1: What are the machine learning and deep learning methods used in each type of feature selection?

Because feature selection is an essential component of the data processing pipeline in machine learning, the effectiveness of the feature selection approach can only be assessed when combined with machine learning algorithms to contribute to classification models. Among the primary studies, we investigated machine learning (ML), deep learning (DL), and both algorithms used in each type of FS method. Table 15 shows that the ML and DL were applied for each type of FS method. We can see that most studies combined FS with machine learning algorithms to build the models, while a few studies only employed FS with deep learning algorithms. It means classic machine learning algorithms need processed data after feature selection so that lightweight models with high performance can be built.

TABLE 15. Machine learning and deep learning applied in each type of FS method.

FS Method	ML(s)	DL(s)	ML(s)+DL(s)
Filter	PS01, PS04, PS05, PS07, PS08, PS09, PS10, PS17, PS20, PS21, PS29, PS42,	PS56, PS37	PS41, PS46, PS62

	PS54, PS57, PS58, PS59, PS61		
Wrapper	PS03, PS12, PS13, PS16, PS33, PS34, PS36, PS39, PS43, PS63		PS52, PS19, PS48
Embedded	PS02, PS24,		PS38
Hybrid	PS06, PS07, PS11, PS18, PS27, PS47, PS50, PS60,	PS45, PS49	PS15
Ensemble	PS04, PS22, PS23, PS28, PS30, PS31, PS35, PS44,		PS40
Comparison	PS05, PS26, PS32, PS51, PS53, PS55		PS25

Some studies employed both ML and DL when implementing the FS method. Classical machine learning algorithms dominate all types of FE methods, since feature selection can affect the final classifiers learned by various ML models; thus, multiple ML algorithms were often used with the proposed FS methods. As for deep learning applied with FS, there are two categories of applying deep learning algorithms in primary studies: one type uses DL as the model training algorithms to build classifiers, while the FS method was independently implemented to generate the feature subsets (Moizuddin and Jose 2022), (Saurabh et al. 2022). However, the other type is using DL (Jayalaxmi et al. 2022), (Moizuddin and Jose 2022), (Cao et al. 2022) as feature extraction based on original feature sets, taking the place of feature selection, and combining it with the feature sets selected by FS methods, to comprehensively identify the features for model training.

RQ4.2: What are the performance metrics used for validation of FS approaches?

Evaluating the effect of a feature selection technique on a classification model involves comparing the performance of the model with and without the feature selection. Comparing the performance metrics of the model with and without the feature selection technique can provide insight into the effect of the feature selection on the model's performance. If the performance metrics improve after applying the feature selection technique, it can be concluded that the feature selection has a positive effect on the model. On the other hand, if the performance metrics are degraded after applying the feature selection, it can be concluded that the feature selection has a negative effect on the model.

Table 16 investigated the common performance metrics used by the primary studies to evaluate the performance of a classification model in IoT security:

TABLE 16. Performance metrics evaluated for feature selection.

Metrics	# of PS	Primary Studies
Accuracy	61	PS01 ~ PS15, PS17 ~ PS43, PS45 ~ PS63
Precision	40	PS01, PS08 ~ PS10, PS13, PS15, PS18 ~ PS22, PS24 ~ PS28, PS31, PS34 ~ PS35, PS37 ~ PS39, PS41 ~ PS43, PS45 ~ PS54, PS51~ PS61, PS63
Recall (Sensitivity /DR/TPR)	49	PS01, PS03, PS08 ~ PS11, PS13, PS15, PS16, PS18 ~ PS31, PS33 ~ PS43, PS45 ~ PS54, PS57 ~ PS61, PS63
F1-Score	40	PS01, PS04, PS07 ~ PS11, PS15, PS18 ~ PS22, PS24 ~ PS26, PS28 ~ PS 32, PS34, PS35, PS37, PS39, PS41~PS43, PS45~PS50, PS52~PS54, PS51~PS61
Specificity (TNR)	5	PS13, PS27, PS37, PS51, PS57
AUC-ROC	8	PS21, PS23, PS25, PS29, PS44, PS45, PS48, PS58
FPR	16	PS02, PS16, PS18, PS20, PS23, PS29, PS30, PS36~PS40, PS45, PS48, PS54, PS57
FNR	2	PS37, PS57
MCC	3	PS30, PS37, PS43
G-mean	1	PS16
Model Size	1	PS23
TTB	21	PS04, PS08~PS12, PS15, PS18, PS21~PS23, PS30, PS32, PS35, PS43, PS45, PS50, PS56, PS57, PS61~PS63
TTP	9	PS08, PS16, PS21, PS23, PS29, PS35, PS50, PS57, PS62
CM	6	PS08, PS10, PS19, PS25, PS46, PS54

Accuracy is calculated as the number of correct predictions made by the model divided by the total number of instances in the dataset. The metric provides a general evaluation of the performance of a model, but it can be misleading in cases where the dataset is imbalanced or has a skewed distribution of positive and negative instances (Disha and Waheed 2022). Precision is the proportion of true positive predictions among all positive predictions made by the model. It measures the ability of the model to correctly identify positive or attack instances and avoid false positive or attack predictions in IoT security model. Recall: The fraction of true positive predictions among all positive instances in the dataset, and it is the same as sensitivity, true positive rate (FPR), or detection rate. F1 score means the harmonic mean of precision and recall. The metrics of accuracy, precision, recall, and F1-score were often used together to evaluate classification models, because each metric has its limitation for model evaluation, thus, most studies evaluated all the four metrics to obtain a comprehensive picture of the model's performance.

False positive rate (FPR) measures the proportion of negative instances that are incorrectly classified as positive by the model, in contrast, False negative rate (FNR) is equals (1-FPR), which means the positive case that are incorrectly identified as negative one. Since after FPR is calculated, FNR can be obtained, while only 2 primary studies evaluated FNR, however the two studies also identified FNR, in fact, there is no need to evaluate it since

FPR is obtained. Moreover, AUC and ROC are commonly used performance metrics in binary classification tasks to evaluate the ability of a model to distinguish between positive and negative classes. The ROC curve provides a visual representation of the trade-off between the TPR and FPR for different classification thresholds, while the AUC provides a single scalar value that summarizes the overall performance of the model across all possible classification thresholds.

Matthews Correlation Coefficient (MCC) is also a performance metric used to evaluate the performance of binary classification models. The MCC takes into account the true positive, false positive, false negative, and true negative rates, and it provides a more comprehensive evaluation of the model's performance than accuracy, precision, recall, or F1 score alone. A value of 1 indicates perfect prediction, a value of -1 indicates perfect anti-correlation, and a value of 0 indicates random prediction. However, few primary studies (3 out of 63) used this metric for evaluation. Because the choice of performance metrics will depend on the specific problem being solved and the goals of the evaluation. For example, in some cases (Shafiq et al. 2021), (Saurabh et al. 2022), accuracy may be more important than sensitivity or specificity, while in other cases (Abdulkareem et al. 2022), (Malik et al. 2022) precision or F1-score may be more relevant. In addition, different metrics may be more or less appropriate for different types of datasets. For example, the studies that used highly imbalanced datasets (Vigoya et al. 2021), metrics such as precision and recall may be more appropriate than accuracy, while in datasets with balanced class distribution, accuracy may be a more appropriate metric. Similarly, G-Mean is a combination of sensitivity (TPR) and specificity (TNR), and provides a single scalar value that summarizes the overall performance of the model. It can be especially useful in situations where both TPR and TNR are important. Only 1 out of 63 primary studies investigated G-mean as the performance evaluator.

What's more, model size in memory (bytes) was identified and compared among various classification models only in (Ismail et al. 2021), in order to identify the most efficient model for wireless sensor network. Besides, the evaluators using time to build, or train models (TTB) and time to predict, or test models (TTP) are other aspects to evaluate the efficiency of the models. There are clearly more primary studies focusing more on TTB (21 out of 63) rather than TTP (9 out of 63), because TTB evaluated based on seconds, dominates more time on model training, while TTP evaluated based on µ seconds, takes less time just on test data prediction after the model is built.

Specificity or true negative rate (TNR) is a performance metric used in classification tasks to measure the proportion of true negatives among all negative instances in the dataset. In classification for IoT security, it means the ability of a model to correctly identify normal case in network flow. Only 5 primary studies evaluated this metric, because most studies focused their effort on attacks rather than normal cases detection and classification, however, the benchmark dataset named BoT-IoT researched by the studies, has high imbalanced class distribution between attack and normal cases, which means, very few normal instances while attack flows dominate the cases. Thus, TNR is the metric to evaluate the capability of classifying the normal cases in highly imbalanced class. Other studies (6 out of 63) evaluated the performance of models using confusion matrix (CM), which is a table that summarizes the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions made by the model.

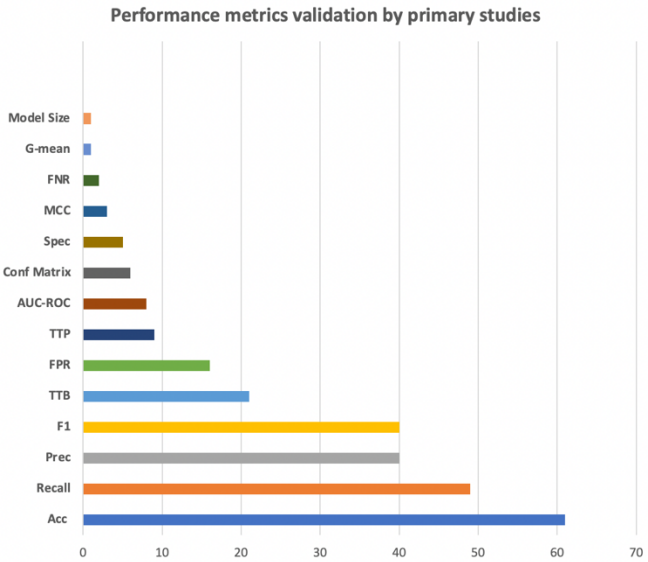

Figure 8. Performance metrics validation by primary studies

We can further refer to Figure 8 for the visualized distribution of the performance metrics discussed above among the primary studies. Accuracy dominates the most, followed by precision, recall, and f1-socre to achieve comprehensive evaluators for classification models. FPR is an important indicator since the cases that are falsely identified as attack classes can lead to more resources to follow up on, such as attack mitigation and preventing intrusions into intrusion prevention systems. Moreover, TTB is a significant indicator to evaluate for the lightweight classification model in resource-limited IoT systems since the number of selected features can directly lead to the cost of building the model. Furthermore, to be specific, the cost for searching, identifying, and selecting the optimal features employed by the methods of feature selection and the FS techniques is another key to the overall cost of machine learning-based classification systems.

RQ4.3: What are the methods of the validation of FS in studies?

We further investigated how primary studies validate the proposed FS to see if comprehensive validation is conducted for each study, besides performance metrics, processing time searching for optimal features, and time for training models and implementing predictions. Table 17 presents the distribution of validation methods for the proposed FS methods of each primary study.

Here are the shortcut names for each character in the table: "Multi-DS" means validated using multiple datasets; in this table, one IoT dataset with any traditional networking, or a non-IoT dataset, is also considered as multiple datasets. "Multi-FS" means various FS techniques were used and compared. Similarly, "multi-MLs" and "multi-DLs" mean multiple machine learning algorithms and multiple deep learning algorithms were used and compared with the proposed FS in studies. "Full features" means the performance of the models based on selected features is compared with the models with full features. Finally, "recent works" means the proposed FS and the models were compared with recent studies on their performance using the same dataset.

The validation of the proposed FS technique using multiple machine learning algorithms is the most validated item. Different algorithms may have different requirements for the input features and may perform better or worse on different types of data. By using and comparing multiple algorithms, the researchers can compare different models using various ML algorithms and identify the most effective model.

TABLE 17. Validation methods for the proposed FS method of each primary study.

FS Validation by	# of PS	Primary Studies
Multi-DS	20	PS09, PS12, PS21, PS25, PS28, PS33, PS38~PS42, PS46~PS48, PS50, PS52, PS54, PS55, PS60, PS63
Multi-FS	12	PS07, PS09, PS11, PS12, PS19, PS26, PS29, PS30, PS45, PS51, PS53, PS56, PS59
Multi-MLs	36	PS01~PS08, PS11~PS14, PS17, PS20, PS22~PS24, PS27~PS29, PS31~PS34, PS38, PS41, PS42, PS44, PS46, PS51, PS53~PS55, PS57~PS61
Multi-DLs	8	PS15, PS37, PS38, PS41, PS46, PS48, PS49, PS56,
Full Features	22	PS05, PS06, PS08, PS10, PS17, PS22, PS26, PS28, PS32, PS33, PS36~PS38, PS43, PS49, PS50, PS53~PS55, PS57~PS59, PS61~PS63
Recent Works	22	PS08, PS12, PS16, PS19, PS28, PS33, PS35~PS39, PS43, PS45, PS46, PS48~PS50, PS53, PS56, PS57, PS59~PS61, PS63

We can refer to Table 17 for detailed information on the validation for each study. Some studies used extensive FS validation methods to consolidate performance results and claim contributions. For example, (Kumar et al., 2021) verified the results except the comparison of multiple deep learning classifiers. Similarly, (Disha and Waheed 2022) validate all the check points in the table except for making comparisons with other FS in the study. Some studies focused more on performance metrics but did not validate the proposed FS method compared with the results of full features and recent studies on the same dataset. For example (Rahman et al. 2021), proposed an ensemble FS to combine the ranked features by three machine learning algorithms, and the proposed FS outperformed any other individual wrapper feature selection with results of 99.95 %, 99.95 %, and 99.90 % accuracy respectively, however, the study does not validate the same model with full features and the results of recent studies.

RQ5: What are the limitations, challenges and future directions of FS to the models in IoT security?

RQ5.1: What are the limitations in current researches?

The IoT data samples included in studies' feature selection techniques may influence the results of the studies. As we concluded from the answers to RQ3, most studies used public IoT dataset to validate their proposed FS methods and machine learning algorithms, while few studies used private data extracted by their experimental environment for specific attack detection. However, the samples varied among some studies even using the same dataset. For example, (Motylinski et al. 2022), (Moizuddin and Jose 2022), (Ozer et al., 2021) used 5% of the whole BoT-IoT dataset, in addition, only 10 features of the original features (45) was used as the candidate features before implementing the proposed FS method. Similarly, (Abdulkareem et al. 2022), (Leevy et al. 2022) used 5% instances of BoT-IoT, with 36 and 29 features, respectively as the sample data. In addition, the proposed model that uses low-frequency IoT datasets gathered in responses to RQ3 requires additional validation using popular datasets.

Moreover, data pre-processing with the domain knowledge of IoT was omitted by some studies, which may cause the performance results of the modes invalid. For example, (A. Fatani et al. 2022) proposed meta-heuristic algorithms AQU searching for the optimal feature subsets, but there is no pre-processing explained in the study. Furthermore. (Ozer et al., 2021) involved the irrelevant features to build anomaly detection systems. (Abdulkareem et al. 2022) conducted dimensionality reduction but did not consider the imbalance class distribution of the BoT-IoT dataset. (Illy et al. 2022) focused on attack and feature analysis and evaluation of manually selected features, however, solid domain knowledge is required to implement manually feature selection. (Kumar et al., 2021) extensively used manual data pre-processing work which may be labor-consuming, and the number of features in the reduced set of features can be further optimized to increase overall accuracy and detection rate.

Various objectives of the classification model drive the implementation of the feature selection methods. Different objectives, such as the chosen performance metrics-oriented model, attacks to be detected and classified, an efficient or lightweight model focused with high accuracy, and so on, may directly influence the FS and machine learning algorithms used for the proposed classification model. For example, the binary classification result can be quite different from that of multi-classification in the same dataset (A. Fatani et al. 2022). Some studies conducted limited performance evaluation, (Ahmed and Tjortjis 2022) evaluated the models with limited performance metrics,

while (Awad et al. 2022), (Saurabh et al. 2022) did not evaluate the effect of the proposed FS with that of original features. Some studies focus only on the specific attacks detection, (Prasad and Chandra 2022), (A. Kumar et al. 2022) focused on binary classification of DDoS and botnet attack among the datasets. Similarly, there is no one FS method or technique that fits all scenarios. Since there are various FS methods applied in primary studies, involving filter, wrapper, embedded, hybrid, and ensemble FS methods, each FS method has its strengths and weaknesses based on the objectives and scenarios.

RQ5.2: What are the major challenges of FS in the primary studies on the IoT security model?

The characteristics of IoT datasets, which are constantly changing, make implementing FS methods and techniques difficult. Large size and high dimensionality are the intrinsic aspects of IoT datasets. For example, BoT-IoT has the maximum of 72,000,000 instances, while the extracted features of N-BaIoT and AWID exceed 100, which can be a challenge for the FS method. Besides, the heterogeneity of IoT datasets can contain a mix of different types of data, such as numerical, categorical, and text data, which can make them challenging to process and analyse. Moreover, IoT datasets can be imbalanced, meaning that the distribution of the target variable may not be evenly distributed. This can pose a challenge for classification models, as they may not accurately reflect the minority class. Furthermore, IoT datasets can be noisy, with errors or missing data points, which can affect the accuracy of the classification models. Finally, since the dynamic characteristics of IoT datasets can change over time, with new data points involving new attack types or zero-day attacks being added or existing data points being updated, this can affect the performance of the classification models that had outstanding performance in previous datasets.

Another significant problem is that the optimised feature scheme lacks sufficient discriminative ability to identify all classes of assaults even with a single dataset. For example, (Shafiq et al. 2021) proposed identified 5 from 45 original features by using hybrid feature selection, which contribute to an effective attack detection performance, however, the sensitivity of specific attack such as data theft attack only achieved 66.67%, compared with more than 99% of other type attacks such as DDoS attack. Moreover, the same FS approach may contribute to opposite output of classification for different datasets. For example, in (Disha and Waheed 2022), the performance of output of UNSW-NB15 and TON-IoT are quite opposite after applying the same FS method, since the class distribution is quite different for this two datasets. Thus, the dataset and the objective of the classification model should be carefully considered before applying the proposed FS method.

Finally, the other challenge is that any FS may have its side effect. Any effective feature selection approach achieving high performance by using specific machine learning algorithm or deep learning algorithm, does not mean the FS scheme provide beneficial effect with other learning algorithms. For example, in (Shafiq et al. 2021), the performance results of C4.5 DT and RF outperform the results of SVM and NB with the same feature selection technique. The result showed the proposed feature selection technique can cause quite different performance result with different machine learning algorithms. In addition, in (Shafiq et al. 2021), the performance results of C4.5 DT and RF outperform the results of SVM and NB with the same feature selection technique. The result showed the proposed feature selection technique can cause quite different performance result with different machine learning algorithms. Similarly, with the same FS approach in (Medjek et al. 2021), RF achieved the best performance compared with other classic machine learning algorithms involving DT, kNN, NB, also outperform the performance of deep learning algorithm MLP, which may be caused by the characteristics of the dataset. Similarly, in (Disha and Waheed 2022), the proposed embedded weighted Gini-based FS showed positive effect on tree-based classifiers, while reduced the performance of neural-based algorithms in terms of accuracy and FAR.

RQ5.3: What are the future research directions of FS in terms of the classification performance of IoT security model?

Since IoT devices are increasingly vulnerable to security threats and attacks, making it vital to develop methods for detecting and classifying these attacks. Various studies applied various feature selection methods and machine learning or deep learning-based frameworks for attack classification in IoT security, which is still a hot area of research, thus, we investigated the future directions after this review study as following.

Diversity and representative of datasets is vital for building generative models for IoT security. Because various zero-day attacks are becoming more common in IoT networks, a dataset that can represent real-world scenarios can be used as a benchmark for attack detection models. (Motylinski et al. 2022), (Abdulkareem et al. 2022), (R. Kumar et al., 2022) suggested multiple public IoT datasets validation is necessary while retaining the distribution of the IoT data in real scenarios when implementing data pre-processing. Moreover, for the specific attacks such as botnet, MITM, and routing attacks, self-created dataset can be used (Prasad and Chandra 2022) (Malik et al. 2022), and (Medjek et al. 2021).

Integration with deep learning models: The integration of feature selection algorithms with deep learning models is a promising direction, as it can provide a more effective way of selecting features for large and complex datasets (Ahmed and Tjortjis 2022). Deep learning algorithms to extract additional features with more characteristics to target attack classes. For example (Rahman et al. 2021) used auto encoder to extract more characteristics of the original features prior to feature selection to improve classification performance. In addition, efficient FS to reduced feature sets can contribute deep learning model with larger architectures to improve the performance result with lower computational cost (Ozer et al., 2021). Lastly, (Gad et al. 2022), (Gaber et al., 2022) argued using more deep learning methods with hyperparameters optimized to create more efficient models.

Semi-supervised and unsupervised feature selection: There is a growing interest in developing semi-supervised

and unsupervised feature selection methods, which can be used when labeled data is scarce or unavailable. (Halim et al. 2021) suggested unsupervised learning algorithms such as clustering to make the machine self-learn new kinds of attacks. (Guo 2021) suggested PCA as the dimensionality reduction technique to examine its performance in IoT IDSs.

Multi-objective optimization: Multi-objective optimization is a growing area of research in feature selection, as it provides a way of balancing different objectives, such as accuracy, interpretability, and computational efficiency (Kareem et al. 2022). Optimized meta-heuristic algorithms such as adaptive PSO can be employed (Chohra et al. 2022). (R. Kumar et al., 2022) suggested highly efficient meta-heuristic methods considering limited energy resource. In addition, improving the effect of FS to predict malicious activities by adjusting the importance score for statistic-based techniques, such as PCC (Awad et al. 2022). Trade-off between the speed and detection rate can be a direction for feature selection in the future. For example, (Shafiq et al. 2021) discussed and proposed an algorithm to evaluate and select many feature subsets considering the trade-off between complexity of the model and attack detection performance. (Kareem et al. 2022) proposed hybrid meta-heuristic algorithms to search for the most optimal feature subsets with reduced searching time, while multi-objective optimization can further be used for hyperparameters of learning algorithms.

Explain ability and interpretability: Explain ability and interpretability are becoming increasingly important in feature selection, as it is important to understand why a model is making certain predictions, and to ensure that the results are not biased or influenced by irrelevant factors. For example, (Guerra-Manzanares et al., 2019) proposed a filter method using Fisher's Score and Local Interpretable Model-Agnostic Explanation (LIME) at feature selection and post-hoc interpretation phases, respectively. Moreover, (Bhandari et al. 2020) conducted a tree-based model with a focus on feature analysis, knowledge of these important features can be used to remove irrelevant features and also to better understand how the models work and what data should be collected in the future. Furthermore, in (Shafiq et al. 2021) and (Kumar et al., 2021), it is promising to observe that the direction of class-wise feature selection, where different results are obtained for different classes using the same feature selection and machine learning techniques, might increase the performance of a particular attack type.

Integration with domain knowledge: The integration of domain knowledge with feature selection is a promising direction, as it can provide a more effective way of selecting features that are relevant to a specific application domain. In order to handle the majority of IoT data, which involve missing values, categorical features, irrelevant features, and distribution imbalance, suitable pre-processing with domain expertise is required. (Siddiqi and Pak 2021) proposed a statistic way to identify the most suitable normalization method for IoT datasets and suggest a direction of hybrid method for normalization to improve machine learning-based IDS. In addition, data pre-processing of handling class imbalance should keep reflecting the class distribution of real scenarios (Motylinski et al. 2022). Moreover, mitigation of the correctly identified attacks implemented as a module in security framework is a promising direction for historic security model in IoT networks (Prasad and Chandra 2022) (Khater et al. 2021).

IoT characteristics and infrastructure: lightweight (energy-efficient) with considerable performance metrics in intrusion detection systems are paramount instead of inefficient and heavyweight intrusion detection systems. (Ozer et al., 2021) intensively reduced the original 12 features optimized by original author, to only 2 features by using feature pair technique. The model with minimal selected features toward specific attack class can contribute to highly efficient or real-time attack detection system. Moreover, reducing FPR is vital to IoT security (Carter et al., 2022), security teams are frequently distracted by false identified attacks because they still require a lot of labour to mitigate in additional intrusion prevention systems (IPS). Moreover, industrial scale with large attack span should be evaluated for the models created from the public datasets (Saurabh et al. 2022), (Illy et al. 2022). Finally, online and incremental feature selection methods are becoming increasingly important, as they can handle large-scale and streaming data in real-time, making them well suited for applications in areas such as sensor networks, social media, and the Internet of Things (IoT). (Kumar et al., 2021), (Ravi et al., 2022) advocate using the optimized models to deploy online for real-time anomaly detection in realistic IoT circumstances.

4. Limitation of The Study

We conducted a systematic literature review about the application of FS methods in machine learning or deep learning-based attack classification models for IoT security based on 63 primary studies between 2018 until 2022. The result of this SLR may have been affected by the coverage of search strategy, the researchers' bias, and inaccuracy of data extraction. These have been discussed and addressed below.

The coverage of search strategy in this study was determined by limited set of keywords that targeted an overview of feature selection methods applied using machine learning or deep learning algorithms for IoT security from the selected academic databases. Thus, there may be the possibility where related studies that using various keywords or the studies from other databases, were not included in this search result.

Another threat to validity is related to the possibility of bias of the researchers. All the primary studies used the at least one data set which was generated in specific IoT environment, as a result, the studies that also employed feature selection and ML or DL-based attack classification were filtered and not entered into following evaluation of quality score criteria. Besides, the score criteria may cause the bias of the quality of the primary studies, because the quality score is designed based on the expected findings on feature selection from the title, abstract, keyword, and full text, but not only general quality of academic study.

Finally, there may be the possibility of inaccuracy of data item extraction on taxonomy of feature selection in this study, such as the category of the purpose of FS, FS

methods, FS validation methods, and performance metrics. However, the data items were extracted based on comprehensive understanding of the studies, and we can ensure the accuracy of this study by providing the detailed searching coverage, data items, and criteria for findings of research questions.

5. Conclusion

Feature selection is an essential step in building accurate and efficient machine learning and deep learning-based classification models to detect and mitigate threats and attacks for IoT security. With the increasing heterogeneous sources in IoT networks and the availability of high-dimensional IoT data, feature selection has become even more important to improving the performance of classification models by reducing the complexity of the models. This study provides a systematic literature review of studies focusing on feature selection methods for IoT security from 2018 to 2023. A total of 63 conferences and articles were reviewed from six research databases: Web of Science, IEEE Xplore, Scopus, ScienceDirect, ACM, SpringerLink, and Wiley Online Library. First, a brief introduction to IoT security and feature selection was provided. The methodology of this research, including research planning, research questions and motivations, the search process, inclusion and exclusion criteria, quality assessment, data collection, and data analysis, was then described. After that, 63 studies were then qualified as primary studies using a quality scoring scheme based on the quality assessment criteria. Thus, in order to answer the research questions, the primary studies published in the past five years on feature selection were then organized and presented based on aspects including, the current situation of feature selection in IoT security, the trend of feature selection, the benchmark datasets, FS validation methods, and metrics. Finally, the limitations, challenges and the future directions for FS were investigated. It is expected that this study will help other researchers, as it provides a systematic review of FS applied to IoT security in recent years.

6. Reference

Abdulkareem, Sulyman Age, Chuan Heng Foh, François Carrez, and Klaus Moessner. 2022. "FI-PCA for IoT Network Intrusion Detection." In *2022 International Symposium on Networks, Computers and Communications (ISNCC)*, 1–6. https://doi.org/10.1109/ISNCC55209.2022.9851723.

Ahmad, Rasheed, and Izzat Alsmadi. 2021. "Machine Learning Approaches to IoT Security: A Systematic Literature Review." *Internet of Things* 14 (June): 100365. https://doi.org/10.1016/j.iot.2021.100365.

Ahmed, Aqeel, and Christos Tjortjis. 2022. "Machine Learning Based IoT-BotNet Attack Detection Using Real-Time Heterogeneous Data." In *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 1–6. https://doi.org/10.1109/ICECET55527.2022.9872817.

Ahmetoglu, Huseyin, and Resul Das. 2022. "A Comprehensive Review on Detection of Cyber-Attacks: Data Sets, Methods, Challenges, and Future Research Directions." *INTERNET OF THINGS*. https://doi.org/10.1016/j.iot.2022.100615.

Al Shorman, Amaal, Hossam Faris, and Ibrahim Aljarah. 2020. "Unsupervised Intelligent System Based on One Class Support Vector Machine and Grey Wolf Optimization for IoT Botnet Detection." *Journal of Ambient Intelligence and Humanized Computing* 11 (7): 2809–25. https://doi.org/10.1007/s12652-019-01387-y.

Alanazi, Manal, and Ahamed Aljuhani. 2022. "Anomaly Detection for Internet of Things Cyberattacks." *CMC-COMPUTERS MATERIALS & CONTINUA*. https://doi.org/10.32604/cmc.2022.024496.

Alazab, Moutaz. 2022. "A Discrete Time-Varying Greywolf IoT Botnet Detection System." *Computer Communications* 192: 405–16. https://doi.org/10.1016/j.comcom.2022.06.016.

Al-Garadi, Mohammed Ali, Amr Mohamed, Abdulla Khalid Al-Ali, Xiaojiang Du, Ihsan Ali, and Mohsen Guizani. 2020. "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security." *IEEE Communications Surveys & Tutorials* 22 (3): 1646–85. https://doi.org/10.1109/COMST.2020.2988293.

Alqahtani, M., H. Mathkour, and M.M. Ben Ismail. 2020. "IoT Botnet Attack Detection Based on Optimized Extreme Gradient Boosting and Feature Selection." *Sensors (Switzerland)* 20 (21): 1–21. https://doi.org/10.3390/s20216336.

Aminanto, Muhamad Erza, Rakyong Choi, Harry Chandra Tanuwidjaja, Paul D. Yoo, and Kwangjo Kim. 2018. "Deep Abstraction and Weighted Feature Selection for Wi-Fi Impersonation Detection." *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*. https://doi.org/10.1109/TIFS.2017.2762828.

Asadi, Mehdi, Mohammad Ali Jabraeil Jamali, Saeed Parsa, and Vahid Majidnezhad. 2020. "Detecting Botnet by Using Particle Swarm Optimization Algorithm Based on Voting System." *Future Generation Computer Systems* 107 (June): 95–111. https://doi.org/10.1016/j.future.2020.01.055.

Ashton, Kevin. 2009. "That '"Internet of Things"' Thing." *RFiD Journal*, 1.

Awad, Mohammed, Salam Fraihat, Khouloud Salameh, and Aneesa Al Redhaei. 2022. "Examining the Suitability of NetFlow Features in Detecting IoT Network Intrusions." *SENSORS*. https://doi.org/10.3390/s22166164.

Bahsi, Hayretdin, Sven Nomm, and Fabio Benedetto La Torre. 2018. "Dimensionality Reduction for Machine Learning Based IoT Botnet Detection." *2018 15TH INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION, ROBOTICS AND VISION (ICARCV)*. International Conference on Control Automation Robotics and Vision.

Baig, Zubair A., Surasak Sanguanpong, Syed Naeem Firdous, Van Nhan Vo, Tri Gia Nguyen, and Chakchai So-In. 2020. "Averaged Dependence Estimators for DoS Attack Detection in IoT Networks." *Future Generation*

Computer Systems 102: 198–209. https://doi.org/10.1016/j.future.2019.08.007.

Bhandari, S., A.K. Kukreja, A. Lazar, A. Sim, and K. Wu. 2020. "Feature Selection Improves Tree-Based Classification for Wireless Intrusion Detection." In *SNTA 2020 - Proceedings of the 3rd International Workshop on Systems and Network Telemetry and Analytics*, 19–26. Association for Computing Machinery, Inc. https://doi.org/10.1145/3391812.3396274.

Bojarajulu, Balaganesh, Sarvesh Tanwar, and Ajay Rana. 2021. "A Synoptic Review on Feature Selection and Machine Learning Models Used for Detecting Cyber Attacks in IoT." In *2021 6th International Conference on Computing, Communication and Security (ICCCS)*, 1–7. https://doi.org/10.1109/ICCCS51487.2021.9776344.

Bojarajulu, Balaganesh, Sarvesh Tanwar, and Thipendra Pal Singh. 2022. "Intelligent IoT-BOTNET Attack Detection Model with Optimized Hybrid Classification Model." *Computers & Security*, 103064. https://doi.org/10.1016/j.cose.2022.103064.

Cao, Bo, Chenghai Li, Junfeng Sun, and Yafei Song. 2022. "IoT Intrusion Detection Technology Based on Deep Learning." In *2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA)*, 284–89. https://doi.org/10.1109/CVIDLICCEA56201.2022.9825291.

Carter, John, Spiros Mancoridis, and Erick Galinkin. 2022. "Fast, Lightweight IoT Anomaly Detection Using Feature Pruning and PCA." In *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, 133–38. SAC '22. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3477314.3508377.

Chohra, Aniss, Paria Shirani, ElMouatez Billah Karbab, and Mourad Debbabi. 2022. "CHAMELEON: Optimized Feature Selection Using Particle Swarm Optimization and Ensemble Methods for Network Anomaly Detection." *COMPUTERS & SECURITY*. https://doi.org/10.1016/j.cose.2022.102684.

Davahli, Azam, Mahboubeh Shamsi, and Golnoush Abaei. 2020. "Hybridizing Genetic Algorithm and Grey Wolf Optimizer to Advance an Intelligent and Lightweight Intrusion Detection System for IoT Wireless Networks." *Journal of Ambient Intelligence and Humanized Computing* 11 (11): 5581–5609. https://doi.org/10.1007/s12652-020-01919-x.

Disha, R.A., and S. Waheed. 2022. "Performance Analysis of Machine Learning Models for Intrusion Detection System Using Gini Impurity-Based Weighted Random Forest (GIWRF) Feature Selection Technique." *Cybersecurity* 5 (1). https://doi.org/10.1186/s42400-021-00103-8.

Doshi, Rohan, Noah Apthorpe, and Nick Feamster. 2018. "Machine Learning DDoS Detection for Consumer Internet of Things Devices." *2018 IEEE SYMPOSIUM ON SECURITY AND PRIVACY WORKSHOPS (SPW 2018)*. https://doi.org/10.1109/SPW.2018.00013.

Fatani, A., A. Dahou, M.A.A. Al-Qaness, S. Lu, and M.A. Elaziz. 2022. "Advanced Feature Extraction and Selection Approach Using Deep Learning and Aquila Optimizer for Iot Intrusion Detection System." *Sensors* 22 (1). https://doi.org/10.3390/s22010140.

Fatani, Abdulaziz, Mohamed Abd Elaziz, Abdelghani Dahou, Mohammed A. A. Al-Qaness, and Songfeng Lu. 2021. "IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization." *IEEE ACCESS*. https://doi.org/10.1109/ACCESS.2021.3109081.

Gaber, Tarek, Amir El-Ghamry, and Aboul Ella Hassanien. 2022. "Injection Attack Detection Using Machine Learning for Smart IoT Applications." *PHYSICAL COMMUNICATION*. https://doi.org/10.1016/j.phycom.2022.101685.

Gad, Abdallah R., Mohamed Haggag, Ahmed A. Nashat, and Tamer M. Barakat. 2022. "A Distributed Intrusion Detection System Using Machine Learning for IoT Based on ToN-IoT Dataset." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*.

Gad, Abdallah R., Ahmed A. Nashat, and Tamer M. Barkat. 2021. "Intrusion Detection System Using Machine Learning for Vehicular Ad Hoc Networks Based on ToN-IoT Dataset." *IEEE Access* 9: 142206–17. https://doi.org/10.1109/ACCESS.2021.3120626.

Gandhi, Rishabh, and Yanyan Li. 2021. "Comparing Machine Learning and Deep Learning for IoT Botnet Detection." *2021 IEEE INTERNATIONAL CONFERENCE ON SMART COMPUTING (SMARTCOMP 2021)*. https://doi.org/10.1109/SMARTCOMP52413.2021.00053.

Guerra-Manzanares, Alejandro, Hayretdin Bahsi, and Sven Nõmm. 2019. "Hybrid Feature Selection Models for Machine Learning Based Botnet Detection in IoT Networks." In *2019 International Conference on Cyberworlds (CW)*, 324–27. https://doi.org/10.1109/CW.2019.00059.

Guerra-Manzanares, Alejandro, Sven Nõmm, and Hayretdin Bahsi. 2019. "Towards the Integration of a Post-Hoc Interpretation Step into the Machine Learning Workflow for IoT Botnet Detection." In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 1162–69. https://doi.org/10.1109/ICMLA.2019.00193.

Guo, G. 2021. "A Machine Learning Framework for Intrusion Detection System in IoT Networks Using an Ensemble Feature Selection Method." In *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2021*, edited by Paul R. Chakrabarti S., 593–99. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/IEMCON53756.2021.9623082.

Halim, Zahid, Muhammad Nadeem Yousaf, Muhammad Waqas, Muhammad Sulaiman, Ghulam Abbas, Masroor Hussain, Iftekhar Ahmad, and Muhammad Hanif. 2021. "An

Effective Genetic Algorithm-Based Feature Selection Method for Intrusion Detection Systems." *Computers & Security* 110 (November): 102448. https://doi.org/10.1016/j.cose.2021.102448.

Hassija, Vikas, Vinay Chamola, Vikas Saxena, Divyansh Jain, Pranav Goyal, and Biplab Sikdar. 2019. "A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures." *IEEE Access* 7: 82721–43. https://doi.org/10.1109/ACCESS.2019.2924045.

Hussain, Fatima, Rasheed Hussain, Syed Ali Hassan, and Ekram Hossain. 2020. "Machine Learning in IoT Security: Current Solutions and Future Challenges." *IEEE Communications Surveys & Tutorials* 22 (3): 1686–1721. https://doi.org/10.1109/COMST.2020.2986444.

Illy, Poulmanogo, Georges Kaddoum, Kuljeet Kaur, and Sahil Garg. 2022. "ML-Based IDPS Enhancement With Complementary Features for Home IoT Networks." *IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT*. https://doi.org/10.1109/TNSM.2022.3141942.

Ismail, Shereen, Tala Talaei Khoei, Ronald Marsh, and Naima Kaabouch. 2021. "A Comparative Study of Machine Learning Models for Cyber-Attacks Detection in Wireless Sensor Networks." Edited by R Paul. *2021 IEEE 12TH ANNUAL UBIQUITOUS COMPUTING, ELECTRONICS & MOBILE COMMUNICATION CONFERENCE (UEMCON)*. https://doi.org/10.1109/UEMCON53757.2021.9666581.

Jayalaxmi, P. L. S., Rahul Saha, Gulshan Kumar, and Tai-Hoon Kim. 2022. "Machine and Deep Learning Amalgamation for Feature Extraction in Industrial Internet-of-Things." *Computers & Electrical Engineering* 97: 107610. https://doi.org/10.1016/j.compeleceng.2021.107610.

Kareem, Saif S., Reham R. Mostafa, Fatma A. Hashim, and Hazem M. El-Bakry. 2022. "An Effective Feature Selection Model Using Hybrid Metaheuristic Algorithms for IoT Intrusion Detection." *Sensors* 22 (4): 1396. https://doi.org/10.3390/s22041396.

Keele, Staffs and others. 2007. "Guidelines for Performing Systematic Literature Reviews in Software Engineering." Technical report, ver. 2.3 ebse technical report. ebse.

Khater, Belal Sudqi, Ainuddin Wahid Abdul Wahab, Mohd Yamani Idna Idris, Mohammed Abdulla Hussain, Ashraf Ahmed Ibrahim, Mohammad Arif Amin, and Hisham A. Shehadeh. 2021. "Classifier Performance Evaluation for Lightweight IDS Using Fog Computing in IoT Security." *ELECTRONICS*. https://doi.org/10.3390/electronics10141633.

Kolias, Constantinos, Georgios Kambourakis, Angelos Stavrou, and Stefanos Gritzalis. 2016. "Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset." *IEEE Communications Surveys & Tutorials* 18 (1): 184–208. https://doi.org/10.1109/COMST.2015.2402161.

Koroniotis, Nickolaos, Nour Moustafa, Elena Sitnikova, and Benjamin Turnbull. 2019. "Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset." *Future Generation Computer Systems* 100 (November): 779–96. https://doi.org/10.1016/j.future.2019.05.041.

Koroniotis, Nickolaos, Nour Moustafa, and Jill Slay. 2022. "A New Intelligent Satellite Deep Learning Network Forensic Framework for Smart Satellite Networks." *Computers and Electrical Engineering* 99: 107745. https://doi.org/10.1016/j.compeleceng.2022.107745.

Kouicem, Djamel Eddine, Abdelmadjid Bouabdallah, and Hicham Lakhlef. 2018. "Internet of Things Security: A Top-down Survey." *Computer Networks* 141 (August): 199–221. https://doi.org/10.1016/j.comnet.2018.03.012.

Kumar, Ayush, Mrinalini Shridhar, Sahithya Swaminathan, and Teng Joon Lim. 2022. "Machine Learning-Based Early Detection of IoT Botnets Using Network-Edge Traffic." *Computers & Security* 117: 102693. https://doi.org/10.1016/j.cose.2022.102693.

Kumar, Prabhat, Govind P. Gupta, and Rakesh Tripathi. 2021. "Toward Design of an Intelligent Cyber Attack Detection System Using Hybrid Feature Reduced Approach for IoT Networks." *Arabian Journal for Science and Engineering* 46 (4): 3749–78. https://doi.org/10.1007/s13369-020-05181-3.

Kumar, Ravinder, Amita Malik, and Virender Ranga. 2022. "An Intellectual Intrusion Detection System Using Hybrid Hunger Games Search and Remora Optimization Algorithm for IoT Wireless Networks." *Knowledge-Based Systems* 256: 109762. https://doi.org/10.1016/j.knosys.2022.109762.

Leevy, J.L., J. Hancock, T.M. Khoshgoftaar, and J.M. Peterson. 2022. "IoT Information Theft Prediction Using Ensemble Feature Selection." *Journal of Big Data* 9 (1). https://doi.org/10.1186/s40537-021-00558-z.

Li, Taotao, Zhen Hong, and Li Yu. 2020. "Machine Learning-Based Intrusion Detection for IoT Devices in Smart Home." *2020 IEEE 16TH INTERNATIONAL CONFERENCE ON CONTROL & AUTOMATION (ICCA)*. IEEE International Conference on Control and Automation ICCA.

Mafarja, Majdi, Ali Asghar Heidari, Maria Habib, Hossam Faris, Thaer Thaher, and Ibrahim Aljarah. 2020. "Augmented Whale Feature Selection for IoT Attacks: Structure, Analysis and Applications." *Future Generation Computer Systems* 112: 18–40. https://doi.org/10.1016/j.future.2020.05.020.

Malik, Kainat, Faisal Rehman, Tahir Maqsood, Saad Mustafa, Osman Khalid, and Adnan Akhunzada. 2022. "Lightweight Internet of Things Botnet Detection Using One-Class Classification." *SENSORS*. https://doi.org/10.3390/s22103646.

Medjek, Faiza, Djamel Tandjaoui, Nabil Djedjig, and Imed Romdhani. 2021. "Fault-Tolerant AI-Driven Intrusion Detection System for the Internet of Things." *INTERNATIONAL JOURNAL OF CRITICAL INFRASTRUCTURE PROTECTION*. https://doi.org/10.1016/j.ijcip.2021.100436.

Meidan, Yair, Michael Bohadana, Yael Mathov, Yisroel Mirsky, Dominik Breitenbacher, Asaf Shabtai, and Yuval Elovici.

2018. "N-BaIoT: Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders." *IEEE Pervasive Computing* 17 (3): 12–22. https://doi.org/10.1109/MPRV.2018.03367731.

Moizuddin, M. D., and M. Victor Jose. 2022. "A Bio-Inspired Hybrid Deep Learning Model for Network Intrusion Detection." *KNOWLEDGE-BASED SYSTEMS*. https://doi.org/10.1016/j.knosys.2021.107894.

Motylinski, Michal, Aine MacDermott, Farkhund Iqbal, and Babar Shah. 2022. "A GPU-Based Machine Learning Approach for Detection of Botnet Attacks." *COMPUTERS & SECURITY*. https://doi.org/10.1016/j.cose.2022.102918.

Moustafa, Nour. 2021. "A New Distributed Architecture for Evaluating AI-Based Security Systems at the Edge: Network TON_IoT Datasets." *Sustainable Cities and Society* 72 (September): 102994. https://doi.org/10.1016/j.scs.2021.102994.

Nagaraja, Arun, and T. Satish Kumar. 2018. "An Extensive Survey on Intrusion Detection- Past, Present, Future." In *Proceedings of the Fourth International Conference on Engineering & MIS 2018*. ICEMIS '18. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3234698.3234743.

Nomm, Sven, and Hayretdin Bahsi. 2018. "Unsupervised Anomaly Based Botnet Detection in IoT Networks." In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1048–53. Orlando, FL: IEEE. https://doi.org/10.1109/ICMLA.2018.00171.

Ozer, Erman, Murat Iskefiyeli, and Jahongir Azimjonov. 2021. "Toward Lightweight Intrusion Detection Systems Using the Optimal and Efficient Feature Pairs of the Bot-IoT 2018 Dataset." *INTERNATIONAL JOURNAL OF DISTRIBUTED SENSOR NETWORKS*. https://doi.org/10.1177/15501477211052202.

Padmashree, A., and M. Krishnamoorthi. 2022. "Decision Tree with Pearson Correlation-Based Recursive Feature Elimination Model for Attack Detection in IoT Environment." *Information Technology and Control* 51 (4): 771–85. https://doi.org/10.5755/j01.itc.51.4.31818.

Page, Matthew J, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *BMJ* 372. https://doi.org/10.1136/bmj.n71.

Parker, Luke R., Paul D. Yoo, Taufiq A. Asyhari, Lounis Chermak, Yoonchan Jhi, and Kamal Taha. 2019. "DEMISe: Interpretable Deep Extraction and Mutual Information Selection Techniques for IoT Intrusion Detection." *14TH INTERNATIONAL CONFERENCE ON AVAILABILITY, RELIABILITY AND SECURITY (ARES 2019)*. https://doi.org/10.1145/3339252.3340497.

Prasad, Arvind, and Shalini Chandra. 2022. "VMFCVD: An Optimized Framework to Combat Volumetric DDoS Attacks Using Machine Learning." *ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING*. https://doi.org/10.1007/s13369-021-06484-9.

Rahman, Md Arafatur, A. Taufiq Asyhari, Ong Wei Wen, Husnul Ajra, Yussuf Ahmed, and Farhat Anwar. 2021. "Effective Combining of Feature Selection Techniques for Machine Learning-Enabled IoT Intrusion Detection." *Multimedia Tools and Applications* 80 (20): 31381–99. https://doi.org/10.1007/s11042-021-10567-y.

Ravi, Vinayakumar, Rajasekhar Chaganti, and Mamoun Alazab. 2022. "Recurrent Deep Learning-Based Feature Fusion Ensemble Meta-Classifier Approach for Intelligent Network Intrusion Detection System." *Computers and Electrical Engineering* 102: 108156. https://doi.org/10.1016/j.compeleceng.2022.108156.

Samdekar, Ramanand, S. M. Ghosh, and Konda Srinivas. 2021. "Efficiency Enhancement of Intrusion Detection in Iot Based on Machine Learning Through Bioinspire." In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 383–87. Tirunelveli, India: IEEE. https://doi.org/10.1109/ICICV50876.2021.9388392.

Saurabh, Kumar, Tanuj Kumar, Uphar Singh, O.P. Vyas, and Rahamatullah Khondoker. 2022. "NFDLM: A Lightweight Network Flow Based Deep Learning Model for DDoS Attack Detection in IoT Domains." In *2022 IEEE World AI IoT Congress (AIIoT)*, 736–42. https://doi.org/10.1109/AIIoT54504.2022.9817297.

Shafiq, M., Z. Tian, A.K. Bashir, X. Du, and M. Guizani. 2020. "IoT Malicious Traffic Identification Using Wrapper-Based Feature Selection Mechanisms." *Computers and Security* 94. https://doi.org/10.1016/j.cose.2020.101863.

Shafiq, Muhammad, Zhaoquan Gu, Shah Nazir, and Rahul Yadav. 2022. "Analyzing IoT Attack Feature Association with Threat Actors." *WIRELESS COMMUNICATIONS & MOBILE COMPUTING*. https://doi.org/10.1155/2022/7143054.

Shafiq, Muhammad, Zhihong Tian, Ali Kashif Bashir, Xiaojiang Du, and Mohsen Guizani. 2021. "CorrAUC: A Malicious Bot-IoT Traffic Detection Method in IoT Network Using Machine-Learning Techniques." *IEEE Internet of Things Journal* 8 (5): 3242–54. https://doi.org/10.1109/JIOT.2020.3002255.

Shi, Lingyun, Longfei Wu, and Zhitao Guan. 2021. "Three-Layer Hybrid Intrusion Detection Model for Smart Home Malicious Attacks." *Computers & Electrical Engineering* 96: 107536. https://doi.org/10.1016/j.compeleceng.2021.107536.

Siddiqi, Murtaza Ahmed, and Wooguil Pak. 2021. "An Agile Approach to Identify Single and Hybrid Normalization for Enhancing Machine Learning-Based Network Intrusion Detection." *IEEE ACCESS*. https://doi.org/10.1109/ACCESS.2021.3118361.

Soe, Yan Naung, Yaokai Feng, Paulus Insap Santosa, Rudy Hartanto, and Kouichi Sakurai. 2020. "Machine Learning-Based IoT-Botnet Attack Detection with Sequential Architecture Dagger." *SENSORS*. https://doi.org/10.3390/s20164372.

Soe, Y.N., Y. Feng, P.I. Santosa, R. Hartanto, and K. Sakurai. 2020. "Implementing Lightweight IoT-IDS on Raspberry Pi Using Correlation-Based Feature Selection and Its

Performance Evaluation." Edited by Xhafa F. Takizawa M. Enokido T., Barolli L. *Advances in Intelligent Systems and Computing* 926: 458–69. https://doi.org/10.1007/978-3-030-15032-7_39.

Statista. 2022. "Statista, 'Cybersecurity - Market Data Analysis & Forecasts.'" https://www.statista.com/study/124902/cybersecurity-report/.

Vigoya, Laura, Diego Fernandez, Victor Carneiro, and Francisco J. Novoa. 2021. "IoT Dataset Validation Using Machine Learning Techniques for Traffic Anomaly Detection." *ELECTRONICS*. https://doi.org/10.3390/electronics10222857.

Xu, Li Da, Wu He, and Shancang Li. 2014. "Internet of Things in Industries: A Survey." *IEEE Transactions on Industrial Informatics* 10 (4): 2233–43. https://doi.org/10.1109/TII.2014.2300753.

Zhou, L., Y. Zhu, T. Zong, and Y. Xiang. 2022. "A Feature Selection-Based Method for DDoS Attack Flow Classification." *Future Generation Computer Systems* 132: 67–79. https://doi.org/10.1016/j.future.2022.02.006.