Received: 24 May 2021 | Revised: 28 September 2021 | Accepted: 15 November 2021

DOI: 10.1002/cpe.6894

WILEY

RESEARCH ARTICLE

Artificial intelligence empowered threat detection in the Internet of Things: A systematic review

Nasim Soltani[1] | Amir Masoud Rahmani[2] | Mahdi Bohlouli[3,4,5] | Mehdi Hosseinzadeh[6,7]

[1] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Future Technology Research Center, National Yunlin University of Science and Technology, Yunlin, Taiwan

[3] Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

[4] Research Center for Basic Sciences in Modern Technologies, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

[5] Research and Innovation Department, Petanux GmbH, Bonn, Germany

[6] Computer Science, University of Human Development, Sulaymaniyah, Iraq

[7] Mental Health Research Center, Psychosocial Health Research Institute, Iran University of Medical Sciences, Tehran, Iran

Correspondence

Amir Masoud Rahmani, Future Technology Research Center, National Yunlin University of Science and Technology, University Road, Section 3, Douliou, Yunlin 64002, Taiwan.
Email: rahmania@yuntech.edu.tw

Summary

Internet of Things (IoT) is a new phenomenon that proposes novel business opportunities. IoT allows the world to be programmable and might provide several benefits for organizations. Based on the IoT survey, cyber-security issues are among the most extensive and complicated challenges faced by IoT devices. Threat detection is considered a preventive measure against malware threats, ransomware, and attacks, which become more serious each year because of the dramatic rise in malware attacks. This article investigates threat detection techniques that fall into three categories: malware detection, attack detection, and ransomware detection, published from 2017 to August 2021. We examine solutions, techniques, features, classifiers, and tools proposed by IoT researchers. Some questions are proposed, and answering the questions may help the researchers suggest a more efficient solution in future works. Furthermore, the achievement and disadvantages of each study are discussed. Finally, based on the reviewed studies, some open challenges and practical measures to future directions are suggested, worth further studying and researching threat detection techniques in the IoT.

KEYWORDS

artificial intelligence approach, cyber-security, Internet of Things intrusion detection, malware attacks

1 | INTRODUCTION

Internet of Things (IoT) directs to the interconnection of homogeneous and heterogeneous intelligent devices such as sensors, cameras, smart bulbs, personal monitoring devices, smart farming, smart locks, and medical devices based on standard communication protocols at any time and any place. IoT has been specified as one of the tremendous paradigms in information technology through Google search trends during the last 10 years and is forecasted to get over the world in the future.[1,2] The Internet-connected things will increase from 0.3 million in 1990 to 75 million in 2025.[3,4] Smart things connect to gateway devices through IoT communication protocols, and then the gateway will then be joined to an application or network server. The network server/application provides promising opportunities to its consumers, such as helpful information, smart home services, business intelligence, industrial automation, environmental monitoring, commercial services, health care services, and smart city services. However, along with the advantages, this ecosystem brings several technical challenges, encompassing bounded storage, power, and computational capabilities, but IoT is attracted significant attention by cyber-security threats, which is a critical concern for industries.[5-8] Vital infrastructures such as IoT are vulnerable to malware, ransomware, and attacks due to their connection and communication via several intelligent devices. It is essential to have lightweight security preventive measures due to limited resources, bounded memory, low computing power, power consumption, small disk space, lack of operating system, and different wireless communication in IoT compared to the traditional network.[9-11] In general, data confidentiality, authentication, access control, and privacy policies are underway projects to build up IoT security because multiple attacks disrupt the network.

https://doi.org/10.1002/cpe.6894
© 2022 John Wiley & Sons, Ltd.

1.1 | An evolutionary of major IoT malware from 2001 to August 2021

Throughout recent years, cyber-attacks against IoT devices have become more popular among hackers for varied purposes. Therefore, the importance of attack detection will be prominent. Figure 1 illustrates the growth of different malware, ransomware, and attacks against IoT devices over 20 years, starting in 2001. It is axiomatic that by technology growth, new malware comes looking from an overall perspective. By referring to the information in more detail, between 2001 and 2008, three main malware, including RBOT family,[12] Hydra, and DNSChanger, had bad implications on IoT devices. After 2009–2013, Darlloz or Zollard, Chuck Norris, Stuxnet ransomware,[13] Aidra or LightAidra, or Zendran had detrimental effects. In addition, BASHLITE,[14,15] XorDDoS, AESDdos, Gafgyt, SynoLocker, TheMoon, and SynoLocker were major malware from 2014 to 2015. In 2016, enormous attacks such as Mirai, Nyadrop self-updating, Hajime Vigilante, and closed-circuit television (CCTV) botnet infected IoT devices to access data. From 2017 to 2018, a group of attacks (Reaper, Amnesia, Persirai, Shishiga, BrickerBot, and Gr1n) and from 2019 to 2021, another group (Silex Linux, Sweyntooth family, Dark Nexus, Kaiji) were bold during this period to attack connected devices such as cars, cameras on TV. Xor, Stuxnet, and BrickerBot try to manipulated or destroy connected devices, while Mirai and Gafgyt are used to maintain a foothold. Also, both Shishiga and SynoLocker gain permissions on a device. Raspberry and Gr1n control and remote IoT on a local network. Hence, the less security configuration in IoT devices, the more attack they receive.[16-18]

1.2 | An evolutionary of major IoT threats from 2005 to August 2021 toward reputed industry

Figure 2 illustrates major attacks into the prominent industry during 16 years starting in 2005. The first major attack was by controlling implantable cardiac defibrillators to access medical information and manipulating power consumption of electrical utility in 2008 and 2009, respectively. In addition, in 2011, researchers found a special way to remotely change the setting of any IoT device to pump excessive quantities of insulin into the blood, which has severe consequences in hospitals. Michigan University's water utility system and traffic light were attacked from their access point in 2011 and 2014, respectively. By the year 2014, hackers sent phishing to the German steel network to control its systems setting. Also, between

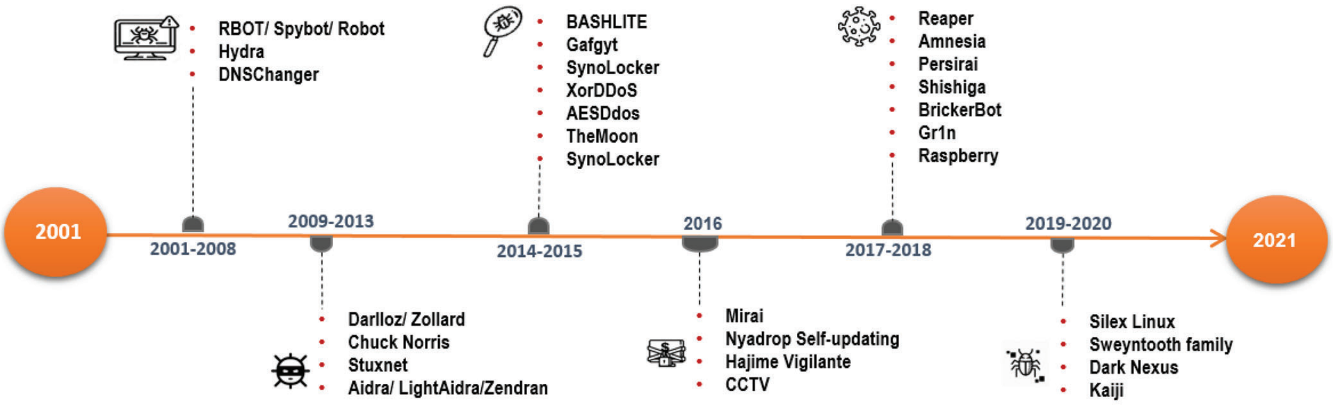
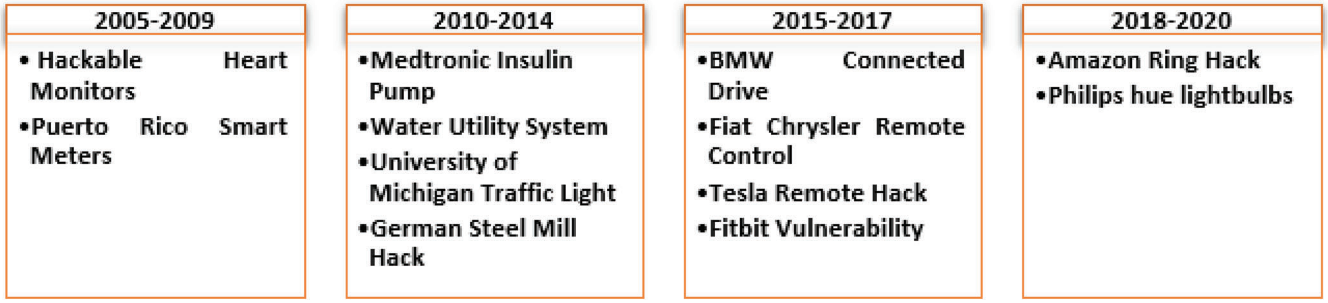


FIGURE 1 A chronological representation of the evolution of IoT malware from 2001 to 2021 in different malware family[16-18]



2005-2009	2010-2014	2015-2017	2018-2020
• Hackable Heart Monitors •Puerto Rico Smart Meters	•Medtronic Insulin Pump •Water Utility System •University of Michigan Traffic Light •German Steel Mill Hack	•BMW Connected Drive •Fiat Chrysler Remote Control •Tesla Remote Hack •Fitbit Vulnerability	•Amazon Ring Hack •Philips hue lightbulbs

FIGURE 2 A chronological representation of the evolution of IoT attacks into the major organization from 2005 to August 2021 in different industrial corporations[18]

2017 and 2017, connected cars and driverless cars were exploited to modify dashboard displays, doors, signals, and trucks. During this period, Fitbit products were vulnerable, and their messages were transmitted between trackers and servers, which hackers could share their data with others after achieving. Finally, from 2018 till the present, big industries such as Amazon[19] and Philips[20] were infected by hacking a ring camera and a single light bulb.[18]

The primary preventive measure is to have a threat detection mechanism to recognize various threats, while it is not straightforward because of code obfuscation to avoid pattern-matching detection. However, threat detection techniques are essential to protect IoT users, investigate threats,[21,22] and generate responses in real-time with different communication protocols. This article presents a Systematic Literature Review (SLR) of the threat detection techniques in IoT. The main contributions of this study are highlighted as follows:

- Mapping a technical taxonomy to classify the various IoT threat detection techniques.
- Providing a comprehensive discussion of the research challenges about threat detection techniques in the IoT environment.
- Introducing open issues in IoT threat detection techniques.

The rest of this SLR is coordinated as follows: The related works are presented in Section 2. Section 3 provides a research methodology and motivation regarding the SLR process. Also, Section 4 outlines the threat detection techniques with a technical taxonomy and compares the techniques for selected papers. In Section 5, a professional discussion is provided on the diverse IoT classifiers, tools, data resources, and techniques that have not been analyzed comprehensively up to now. In addition, Section 6 illustrates some major challenges in threat detection and their practical solutions. Finally, Section 7 demonstrates the conclusion along with the article restrictions.

2 | RELATED SURVEYS

Threat detection in IoT has been gaining colossal research attention due to the requirement for more security. This section explains the survey studies about threat detection in IoT, which are divided by year.

2.1 | 2017

Yaqoob et al.[23] presented a review of ransomware and other security challenges in the IoT environment in which IoT malware is more destructive than traditional malware. First, this research described the basic information of ransomware and introduced widespread types. The author categorized current ransomware into three groups of crypto-ransomware, locker-ransomware, hybrid-ransomware. Furthermore, ransomware challenges and their remedies were discussed. The advantage of this study is presenting new requirements for securing IoT and explaining new open research concerns and future challenges on IoT security include data integrity, lightweight security mechanisms, trust, and privacy. Nevertheless, this survey excludes published papers from 2018 to 2021, and other IoT threats have not been included.

Additionally, Zarpelão et al.[24] introduced a survey related to intrusion detection (ID) in IoT environments. In this survey, various types of placement strategies (distributed, centralized, and hybrid), diverse types of detection methods (signature-based, anomaly-based, signification based, and hybrid), various types of security threats (conventional attack, routing attack, man-in-the-middle, and denial of service (DoS) attack), validation strategies (hypothetical, empirical, simulation, and theoretical) were offered. However, other types of threats such as malware and ransomware and intelligent techniques such as machine learning (ML) techniques did not analyze.

2.2 | 2018

Makhdoom et al.[7] surveyed threats to the IoT to clarify how threats can affect the IoT. This study's strong point is to present a classification of various threats in diverse layers include the semantic layer, application layer, network layer, and perception layer. Furthermore, this article describes the differences between IoT and the traditional network to clarify that existing measures are unsuitable for IoT. Also, this article classifies different attacks based on their impacts on deployment. Another field that goes over by this research was about analyzing malware and its behavior in IoT.

Moreover, risk assessment, defense-in-depth, and preventive measures are introduced to improve security in IoT. However, malware detection techniques based on intelligent methods such as ML were not considered. Besides, Sharmeen et al.[25] provided a survey of malware detection techniques in mobile IoT networks. First, researchers categorized malware detection techniques into three groups include static, dynamic, and hybrid. Then, they discussed diverse datasets that were used in three categories and how researchers created datasets.

Another point to consider is that feature extractor, feature selection, detection technique, and detection technique among mobile IoT networks were analyzed. Nevertheless, this survey excludes published papers between 2019 and 2021. This survey just focused on IoT mobile IoT networks, and other operating systems have not been analyzed.

Elrawy et al.[26] presented a comprehensive survey for intrusion detection systems (IDS) based on IoT innovative environments. Researchers introduced IoT paradigms, IoT system architectures, security challenges in IoT-based intelligent environments. The following step described diverse IDS, various methods, and techniques such as data mining, ML, statistical model, rule model, payload model, protocol model, and signal processing model). In the final step, various techniques, systems, features are compared with each other. However, all of the published papers from 2019 and 2021 and malware and ransomware publications were not considered.

2.3 | 2019

A survey on IoT malware was presented by Vignau et al..[16] First, this review introduced 16 common malware, their behavior, and their performance and then classified malware based on behavioral features. Then, the authors found that Telnet and Secure Shell (SSH) protocols are mainly used among infected IoT devices. However, ransomware and intrusion were not considered a big dataset of malware with various types was not analyzed. Moreover, Clincy and Shahriar[27] analyzed three mechanisms of malware detection techniques: blockchain technology as a ledger system, ML to train and detect malware, recognizing based on image processing, and using OllyDBG as static to extract application programming interface (API) and to show malware behavior. Nevertheless, this survey was not investigated systematically, and parametric comparisons on papers were not analyzed deeply. Either open issue challenges were not provided.

Also, Neshenko et al.[28] provided a comprehensive survey about vulnerabilities and exploitations in IoT environments. First, this survey analyzed vulnerabilities in different layers, and then impacts of security include integrity, availability, and accountability confidentially. In the next step, diverse attacks against confidentially, data integrity, and data availability were extracted. However, malware and ransomware did not consider in their category.

Furthermore, Chaabouni et al.[29] presented a complete survey related to IoT IDS based on learning techniques. In this article, various IoT threats were classified by challenges and tools analyzed to detect malicious behavior. Free datasets, free network traffic sniffers, and open-source IDS compared with each other. Learning techniques include data mining, ML, and deep learning (DL), besides their algorithms, which were thoroughly analyzed in the following step. Nevertheless, all of the published papers from 2019 to 2021 and all malware and ransomware publications were not considered.

2.4 | 2020

Dange and Chatterjee[30] discussed human to thing and things to things communications that refer to intelligent devices. First, this article provides recent attacks and their IoT mechanism and predicts that the IoT would be an essential target and attack for years. Furthermore, evolution, architecture, the botnet's life cycle in IoT, and their differences against traditional botnets are provided. Another point in this survey is related to botnet detection techniques that contain host-based, network-based, signature-based, DNS-based, anomaly-based, and mining-based. Lastly, existing techniques were analyzed, and network detection techniques were introduced preferably. However, there are diverse types of malware and ransomware, but this survey has not been considered.

2.5 | 2021

In addition, Mishra and Pandya[31] provided a systematic review based on security challenges, attacks, and intrusion detection in IoT. The future vision of IoT is also considered in their analyses. In the first step of this study, various IoT applications are introduced, including smart vehicles, smart cities, smart grids, and other types of applications. Also, the evolution of IoT technologies over 21 years ending in 2021 is scrutinized by authors. Another point to consider is that research challenges concentrating on security and related attacks in that domain from 2013 to 2020 are presented. Finally, the performance of DDoS attacks and attackers to access IoT devices and possible solutions to have a future perspective are illustrated in this survey. However, malware, ransomware, and other types of attacks should be taken into consideration.

A summary of the above review studies on the IoT threat issues concerning the SLR is presented in Table 1. The main subject of each survey, such as ransomware, malware, intrusion, botnet, threat, and vulnerabilities, advantages, and disadvantages by considering publication year for each study are investigated in this table. Additionally, all analyzed papered are from 2017 to August 2021.

TABLE 1 Comparison of existing surveys about IoT threat detection techniques from 2017 to August 2021 in IoT

Year/ref.	Main topic	Advantages	Disadvantages
2017[23]	IoT ransomware	Categorizing different types of ransomware and providing future and open issue challenges	Excluded threats diversity
2017[24]	IoT intrusion	Evaluating attributes and detection methods and DS placement strategy and validation strategy	Excluded malware and ransomware detection
2018[7]	IoT threats	Providing numerous threats at various layers	Lack of intelligent malware detection techniques
2018[25]	Mobile IoT malware	Analyzing feature selection algorithms and investigating diverse detection techniques	Excluded multiple systems
2018[26]	IoT intrusion	Comparison of various IoT intrusion techniques	Excluded multiple threats
2019[16]	IoT malware	Different features and various malware	Small analyzed dataset
2019[27]	IoT malware	Introducing malware detection techniques	Lack of open issues and challenges and without comparison on detection techniques
2019[28]	IoT vulnerability	Categorizing IoT vulnerability and exploitations	Excluded malware and ransomware detection
2019[29]	IoT intrusion	Comprehensive analysis on IoT IDS	Excluded multiple threats
2020[30]	IoT botnet	Comparison of diverse malware detection techniques	Not considering different types of threats
2021[31]	IoT attacks	Focusing on DDos attacks in IoT layers	Not scrutinizing other types of attacks

Based on the above surveys and to cover up the following weaknesses, an exhaustive study is investigated:

- The general subject of the existing studies does not portray an investigative classification and comparison for threat detection techniques.
- The organization of the existing studies was about security challenges in IoT, not threat detection techniques.
- The existing surveys are not systematic on three categories of malware, ransomware, and attacks.

3 | RESEARCH METHODOLOGY

We plan an exhaustive study based on SLR for threat detection techniques in IoT. This systematic study aims to provide a broad and complete summary of the recent literature related to researching fields published from 2017 to August 2021. Contrary to non-structured review, the SLR follows a defined and exact order of research literature's methodological steps.

3.1 | Definition of research questions

Collecting and analyzing all logical and efficient studies related to IoT threat detection techniques is considered in this study. Research questions originate from our motivation to perform this study. In other words, a clear vision related to threat detection techniques will be created by answering these questions. Critical factors and methods of papers are extracted, and their specifications are elaborated. This SLR answers the following research questions (RQ) regarding the aim of this research.

In IoT threat detection, the main research questions are as follow:

- RQ1: What strategies are provided?

  The answer to this question allows researchers to know about widespread strategies among scholars and which techniques would be more productive. Therefore, researchers can bank on this systematic review to find the best strategy rather than dealing with several papers.
- RQ2: What strategies are provided?

  The answer to this question illustrates an excellent perspective to scholars about valid solutions and all methods applied until August 2021.
- RQ3: What types of features are applied?

  One of the challenging questions between cyber-security specialists is finding the main features from infected files. Hence, the answer to this question will help researchers focus on prominent features through artificial intelligence techniques rather than a wide array of features to achieve more accuracy.

- RQ4: Which classification algorithms are appropriate?

  Classification algorithms in artificial intelligence approaches are many. The answer to this question will provide a fantastic outlook to scholars, which classifiers would have more accuracy than others by considering their data and techniques.

- RQ5: What types of tools are used?

  The answer to this question helps researchers to use an exact tool to simulate or implement their novel ideas.

- RQ6: Which data resources are suitable?

  Simulating proposed ideas by prepared data is indispensable to evaluate their proposed work by scholars. So, the answer to this question will help to identify various datasets in IoT threat intrusion detection.

3.2 | Selecting research studies

Seven online scientific databases include ACM,[*] ScienceDirect,[†] Springer,[‡] Sage,[§] IEEE,[¶] Wiley,[#] Taylor and Francis[‖] were used based on search terms, considering the following search strings. Table 2 depicts the search terms used in this article.

Figure 3A shows the distribution of the research studies considered in this SLR based on their titles, abstracts, keywords, and conclusions. The highest portion of selected papers accounted for Elsevier journals which are 35% compared to other online databases. Also, as shown in Figure 3A, 31% of all papers were related to IEEE, followed by springer at 24%. The lowest percentage of this SLR is 3% related to ACM, 7% to Wiley, and 0% to sage, Taylor & Francis. Selected research studies were published from 2017 up to August 2021. Furthermore, Figure 3B shows the research studies distribution by year in this systematic review. Looking from an overall perspective, it is clear that security issues in IoT in the late years are considered more attention than before 2017. In this figure, 3% of papers are accounted for 2017, 28% to 2018, 28% and 24% to 2019 and 2020, respectively. Finally, 17% of all papers belonged to 2021.

3.3 | Inclusion and exclusion criteria

After providing the research questions, the inclusion/exclusion principles were applied for the ultimate research selection, depicted in Figure 4. The methodology of this SLR is shown in this article. Initially, we obtained 124 records, all of them identified through electronic databases and research engines. These results were reduced to 72 after excluding duplicates. Second, eligibility criteria based on the title and abstract were applied to these 54. This elimination round reduced our results to 45; further eligibility criteria based on the full text finally enabled us to obtain 31 results. These 31 papers will be analyzed in depth to extract the results presented in Section 4. In the following step, all filtered studies were selected to analyze accurately.

Finally, as shown in Figure 4, 31 peer-reviewed papers were analyzed to answer the mentioned research questions. The prominent inclusion and exclusion criteria are considered as follows:

Inclusion criteria are considered as follows:

- The research studies are available online.
- The research studies are from 2017 up to August 2021.

TABLE 2 Search terms used through surfing nine online databases to find related research works

Search terms
(IoT or "Internet of Things") and botnet and detection
(IoT or "Internet of Things") and malware and detection
(IoT or "Internet of Things") and ransomware and detection
(IoT or "Internet of Things") and threat and detection
(IoT or "Internet of Things") and intrusion and detection
(IoT or "Internet of Things") and worm and detection
(IoT or εInternet of Thingsε) and vulnerability and detection
(IoT or εInternet of Thingsε) and attack and detection

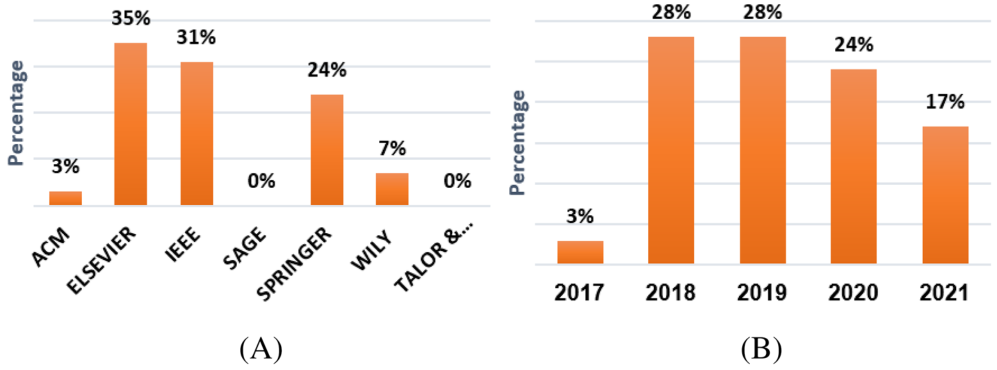
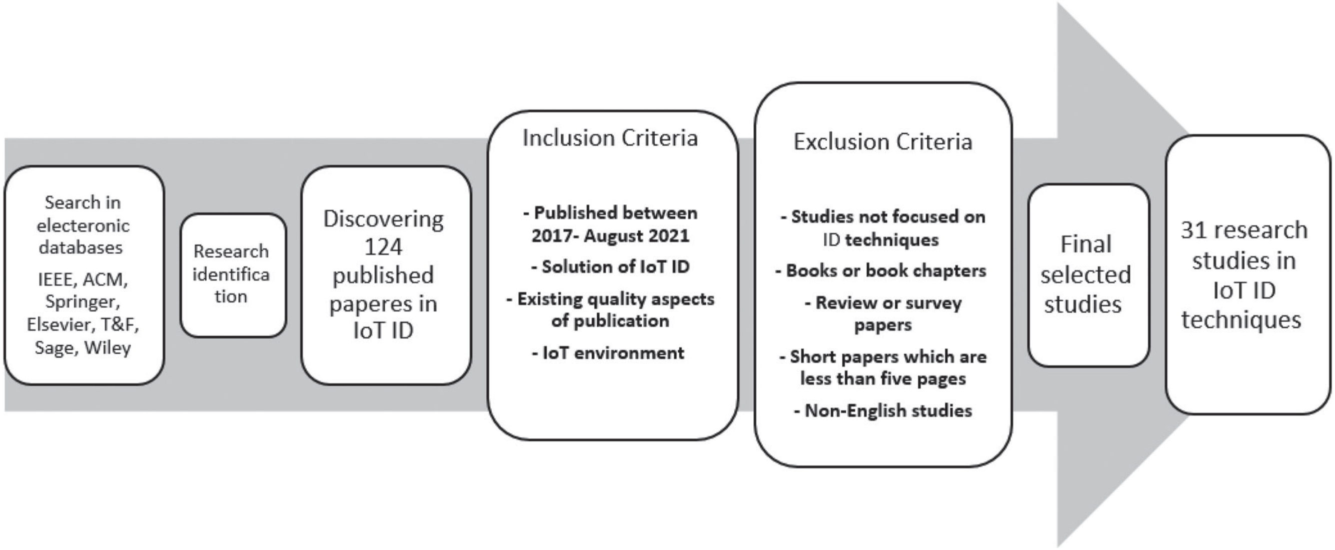


FIGURE 3 The distribution portion of research papers by varied publishers and year in this SLR from 2017 to August 2021. (A) Division by publishers. (B) Division by year


Search in electeronic databases

IEEE, ACM, Springer, Elsevier, T&F, Sage, Wiley

Research identification

Discovering 124 published paperes in IoT ID

Inclusion Criteria

- Published between 2017- August 2021
- Solution of IoT ID
- Existing quality aspects of publication
- IoT environment

Exclusion Criteria

- Studies not focused on ID techniques
- Books or book chapters
- Review or survey papers
- Short papers which are less than five pages
- Non-English studies

Final selected studies

31 research studies in IoT ID techniques

FIGURE 4 The selection criteria and evaluation chart of research studies

- The research studies are in an IoT environment.
- The research studies with a practical quality structure.
- The research studies using existing quality aspects.

Exclusion criteria are considered as follows:

- The research studies providing review and survey papers.
- The research studies not written in the English language.
- The research studies not focused on threat detection techniques.
- The research studies are a book or book chapter.
- The research studies are less than five pages.

4 | ORGANIZATION OF IOT THREAT DETECTION TECHNIQUES

A technical review from selected papers of threat detection techniques is provided in this section, according to the applied SLR process.

Figure 5 shows a comprehensive taxonomy of IoT's threat detection techniques based on related works, including malware, ransomware, and attack detection techniques. The malware detection category includes ML, DL, signature-based, and providing sandbox. The attack detection

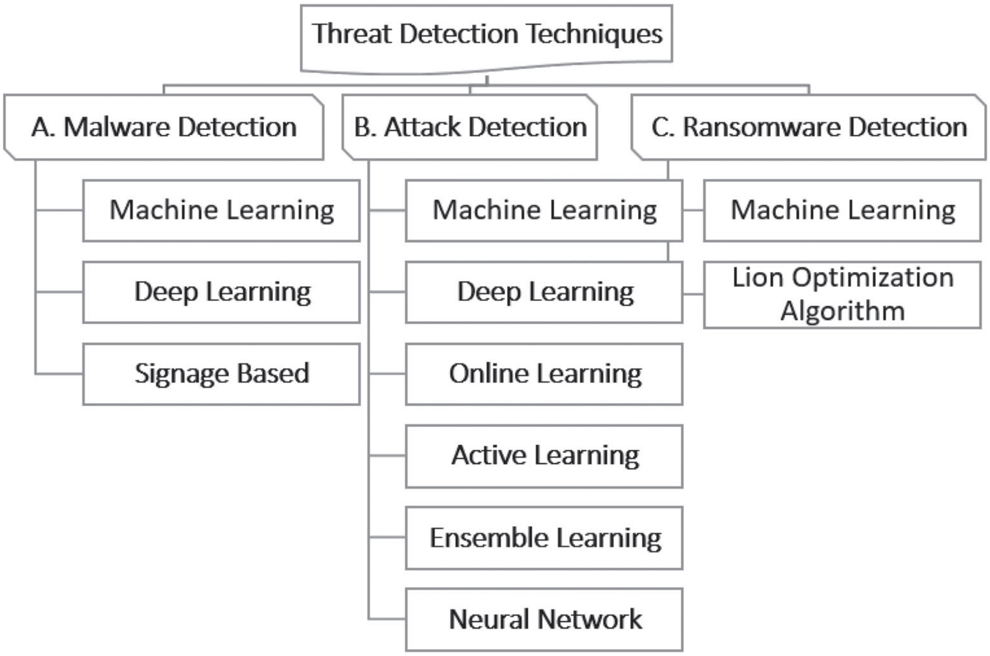


FIGURE 5 Proposed taxonomy of threat detection techniques and their solutions to detect anomalies

category contains ML, DL, online learning (OL), active learning (AL), ensemble learning (EL), and neural network (NN). Finally, the ransomware detection category composes the ML technique and lion optimization algorithm (LOA) to detect threats. Additionally, more analyses are examined in Sections 4.1–4.3.

4.1 Overview of malware detection techniques in IoT

Malware (for instance, viruses, trojans, spyware, rootkits, worms, and bots) is a malicious program that has been designed to achieve the detrimental goals of attackers on the infected systems. These goals included discovering security vulnerabilities, disturbing system operations, access to computational and network resources, obtaining users' sensitive information, modifying integrity, and monetary incentives.[32-34] System calls, API calls, binary codes, and operation codes (Op-Code) are widespread IoT malware features considering analysis. The selected IoT malware detection techniques and their summary have been analyzed in this section.

Abbas and Srikanthan[35] presented an offline detection technique based on signatures. First, this research study system calls from 70 malware and benign were extracted. Second, unique sequences of system calls were referenced to identify a group of malware. Finally, results portray that this proposed approach achieved a 100% detection rate compared to anti-virus detection. Nevertheless, this research does not compare with other intelligent methods such as ML to analyze accuracy. Also, it needs updating signature lists based on new malware files.

Azmoodeh et al.[36] provided a deep Eigensapce method to detect malware, consisting of two steps. The Op-Code sequences graph is generated from the executable file and the edge value in the first step. In the second step, due to complicated data structure, eigenvectors and eigenvalue are applied to convert a graph's adjacency matrix into a vector space. Vectors are considered as a value for classification in the convolutional network of DL. Accuracy, recall, and F-measure were evaluated, which researchers found 98.37% accuracy. However, only 128 malware instances for ARM-based IoT applications from VirusTotal** are small as a dataset.

HadadPajouh et al.[37] introduced a novel IoT malware detection approach, including three stages. Op-Codes were extracted from 280 malware samples and 271 benign samples through an object dump tool in the first step. In the second step, features were extracted through text mining to create feature vectors. Researchers used long short term memory (LSTM) to create a deep structure and Google TensorFlow as the backend structure in the last step. Also, they have used bidirectional neural networks (BNNs) rather than the recurrent neural network (RNN) neuron structure. Finally, 98.18% of accuracy was received against other classifiers such as random forest (RF), support-vector machine (SVM), naive Bayes (NB), multi-layer perceptron (MLP), K-nearest neighbors (KNN), AdaBoost, and decision tree (DT). Nevertheless, the dataset is too small to deal with the real world.

Nguyen et al.[38] introduced a graph-based approached to detect botnet. This study's aim was based on printable string information (PSI), which contains six steps. First, analyzing the botnet life cycle and attaining some footprints such as IP address, username, password patterns, and malicious code. Then, producing function call by using IDA-Pro or UPX and creating PSI-Graph. In the last step, data are preprocessed and classified through convolution neural network (CNN). Finally, researchers achieved 98% accuracy in classifying botnet and benign samples. Nevertheless, creating PSI-Graph has high overheads.

Darabian et al.[39] provided a new method to distinguish malware from benign in IoT. Regarding some Op-Codes frequency, this research counts the Op-Code replication portion in the malware to figure out the frequent Op-Codes. In the second step, ML was analyzed to classify malware and benign-ware, and then Op-Code was converted to numerical sequence as a vector. Researchers also evaluate the accuracy and F-measure of different classifiers containing DT, KNN, RF, MLP, SVM, and AdaBoost that all of the classification methods had more than 98% accuracy. Finally, researchers suggest maximal frequent patterns (MFP) of Op-Code sequences to detect malware, which achieved 99% accuracy. However, the newer classification of ML techniques has not been evaluated.

Dovom et al.[40] introduced a modern detection of IoT malware that contains two steps. First, features were extracted from binary code, and then the control flow graph (CFG) was generated. Second, the fast fuzzy tree was applied to classification detect malware families. Also, four datasets were considered in this research. The Vx-Heaven dataset[††] includes 22,000 samples, the ransomware dataset contains 555 instances, the IoT dataset includes 1207 samples, and the Kaggle dataset[‡‡]consists of 9601 instances. However, this study is compared against other classifiers; the RF had low accuracy.

Karanja et al.[41] proposed an approach to classify malware and benign-ware with ML using the Haralick image texture feature. First, the IoT malware binary code has been converted into 8-bit vectors then transformed into a grey image. The second phase of this gray level co-occurrence matrix (GLCM) from each image is computed, and texture features relying on the GLCM are estimated. In the multi-class classification RF, NB, and KNN are applied as classifiers, which RF had more accuracy against other classifiers. However, new classifiers have not been applied to this research.

Tien et al.[42] used executable and linkable format (ELF) and Op-Code as features to distinguish malware from benign. First, the features of all malware and benign files were extracted. In the following step, the labeling dataset was done and saved in the detection dataset. After preprocessing malware, 30,146 IoT malware were collected by the IoT Honey-Pot project and then classified by ML techniques through KNN, ANN, and SVM classification. However, family types of malware have not been detected. In addition, a dynamic analysis for IoT malware detection (DAIMD) was proposed by Jeon et al.[43] Malware is executed to extract diverse features such as memory, network, process, and system calls in this research. Then, the extracted features are preprocessed for feature selection and classification. Finally, the behavior of malware is converted into an image and trained through CNN. In other words, CNN can detect anomalies through differences in malware and benign images. However, the hybrid method includes a dynamic and static analysis that could be selected.

Radhakrishnan et al.[44] proposed a LSTM technique to detect IoT malware. First and foremost, all data gathered from VirusTotal Website and their Op-Code were extracted through a Linux bash code. The following step extracted Acid, Addfs, Ter, and Rox features by text mining approach. Next, BNN was used to detect malware, which is RNN modification. Finally, compared with varied ML classifiers. Nevertheless, this article could not detect malware by their families, and multi-class detection should be considered.

4.1.1 | The review of malware detection techniques

Table 3 shows a side-by-side evaluation of the abovementioned research studies and the influential aspects of evaluating malware detection techniques in the IoT environment. Essential features in the malware detection techniques are based on the Op-Code, API call, system call, binary code, and ELF characteristics. Additionally, KNN, ANN, MLP, RF, CNN, LSTM, SVM, NB, and fuzzy patterns are selected as popular classification techniques.

Table 4 portrays datasets of the above research studies and malware, benign-ware, and resource, including each dataset's differences. Additionally, the number of malware samples, the number of benign samples, the dataset's link, labeled, and publicity are elaborated in Table 4. The primary resources of malware datasets are the VirusTotal website and the VirusShare website.[§§]

4.2 | Overview of attack detection techniques in IoT

Widespread attacks and anomalies in the IoT environment which bring about failure are DoS attack, data type probing, malicious control, man-in-the-middle (MITM), scan, spying, user-to-root (U2R), malicious operation, remote-to-local (R2L), and wrong setup. Anomalies in IoT packets would detect through widespread network traffic features such as IP address, protocols, source IP, file type, frame length, frame number, packet forwarding, packet dropping, and so forth. The selected IoT attack detection techniques and their summary have been analyzed through their network traffic data.

Yang et al.[48] aimed to detect anomalies through an AL approach for wireless attack detection. Its dataset includes attack and normal data that has been trained and then updated regularly to detect new attacks. Researchers applied three AL phases: supervised learning, label selection, and labeling by a specialist. This learning continues until it reaches a threshold. In the last step, they have found more than 99% in recall and precision. However, its comparison with the random-select algorithm was not efficient.

TABLE 3 Malware detection techniques in IoT of varied papers from 2017 to August 2021

Year/ref.	Technique	Main idea	Advantage	Disadvantage	Features	Classifier	Accuracy	Tool	Disassembler
2017[35]	Signature-based	Extracting unique signatures	Low complexity	Inefficient comparison	System call	–	100%	Linux sandbox	Sandbox
2018[36]	DL	Extraction Op-Code sequences	High accuracy	Small dataset	Op-Code	CNN	98.37%	Matlab	Obj-dump
2018[37]	DL	Extraction Op-Code	High accuracy	Small dataset	Op-Code	LSTM	98%	–	Object-Dump
2019[38]	DL	Printable string information	High accuracy	Overhead	Function call	CNN	98%	Python	IDS-Pro, UPX
2019[39]	Pattern mining and ML	Extracting frequent Op-Codes	High accuracy	Traditional ML	Op-Code	KNN, SVM, MLP, DT, AdaBoost	99%	–	IDA-Pro
2019[40]	ML	Classify by fast fuzzy tree	Diverse malware	Low accuracy	Op-Code	SVM, KNN, RF, DT, fuzzy, and fast	93%	Python & Matlab	–
2020[41]	ML	Extracting Haralick image texture	High accuracy	Excluded data diversity	Binary code	RF, KNN, NB	95%	–	–
2020[42]	ML	–	Varied dataset	Low features	Op-Code& ELF	ANN, CNN, and SVM	98%	Python	Radare2
2020[43]	ML	Considering diverse behavior	High accuracy	Lack of static features	Behavior	CNN	95%	–	IDA-Pro
2021[44]	DL	Text mining and classification	High accuracy	Small dataset	Op-Code	BNN	99.08%	Weka	Linux

TABLE 4 Malware IoT datasets from different references and their detail

Paper	# malware samples	# benign samples	Resources	Labeled	Public
Abbas and Srikanthan[35]	70	–	VirusShare	Yes	Yes
Azmoodeh et al.36, [36]	128	1078	VirusTotal	Yes	Yes
HadadPajouh et al.37, [37]	280	271	VirusTotal	Yes	Yes
Nguyen et al.[38]	6165	345	VirusShare and T-Pot	Yes	Yes
Darabian et al.[39]	247	269	VirusTotal	Yes	Yes
Dovom et al.[40]	12,000[a]	NA	Kaggle[45,46]	Yes	Yes
Karanja et al.[41]	258	–	47	No	No
Tien et al.[42]	6251	2157	VirusTotal	Yes	No
Jeon et al.[43]	1000	401	47	NA	NA
Radhakrishnan et al.[44]	282	271	VirusTotal	Yes	No

[a] Include four different databases.

Meidan et al.[49] proposed deep auto-encoders to detect anomalous network traffic by extracting behavioral snapshots. In this study, first, researchers have generated data by using nine IoT devices and infected them with two botnets (Mirai and BASHLITE). After data collection, features were extracted when a packet arrived. Then, deep auto-encoders were used to train an anomalies detector. In the final step, packets have been observed regularly for anomaly detection. Nevertheless, traffic predictability is a new idea that needs more investigation on this subject.

Rathore and Park[50] introduced a robust framework based on the fuzzy C-Means method with an extreme ML classifier to support qualified attack detection in IoT. First, the fuzzy C-means algorithm was used to group unlabeled data. Also, to execute training of single hidden-layer feedforward neural network was provided. Ultimately, researchers used fuzzy C-means for attack detection with a distributed framework while not needing labeled data. Most importantly, it shows distributed architecture is faster than centralized architecture. However, the random assignment in the extreme ML algorithm may lead to lower performance.

Moustafa et al.[51] introduced a threat detection technique based on EL in the IoT environment due to vulnerabilities in application protocols such as DNS, HTTP, and MQTT. In the first step of EL, researchers designed and developed a tool to extract traffic flow in an IoT network. Alongside their proposed module, TCP-dump to capture network traffic and Bro-IDS to analyze PCAP files and expand flow features. After feature selection, data were divided into three subsets, then several classifications by DT, ANN, and NB were applied to separate normal and attack data. Two datasets, namely the UNSW-NB15 dataset[52,53] and the NIMS dataset,[54] were considered to evaluate this research study. Finally, the ensemble has achieved approximately 98% of data sources in different protocols. However, it needs to find relevant features of other IoT protocols.

Diro and Chilamkurti[55] presented a new approach to detect attacks. In this distributed detection technique, fog nodes are considered hosts, and the master node shares parameters. The master node updates neighbors' parameters, and the centralized training could be extended into a distributed network. Finally, a DL algorithm detects normal and attack networks through parameters submitted from the master node. Nevertheless, it needs to compare with other ML algorithms.

Hasan et al.[56] proposed an ML framework with several independent processes. In the first step of this research, data were collected and preprocessed on the dataset. Data preprocessing was included as cleaning data, visualization data, engineering feature, and vectorization. In the next step, training and testing sets were built based on feature vectors. No feature selection approach has been selected, but feature types such as discrete, continuous, ordinal, and nominal in the feature engineering phase were necessary. Finally, researchers found that the RF technique has had more accuracy than other classification techniques through ML, while LR, SVM, DT, and ANN were considered. However, a new problem may happen in terms of significant and real data.

A threat detection system based on DL to secure IoT environments was introduced by Otoum et al..[57] First, the original dataset was preprocessed to replace missing values with nearest neighbor computation and remove redundant data. The spider monkey optimization (SMO) algorithm was applied to extract relevant features and decrease feature learning time in the next step. Stacked-deep polynomial network was adapted to develop a supervised deep NN algorithm and layers learning in the classification phase. In the final step, accuracy (99.02%) and precision (99.38%) were considered comparison metrics. Nevertheless, this method needs to evaluate various datasets and different classifiers.

Zolanvari et al.[58] proposed ML-based network attacks in industrial IoT. The authors implemented an Industrial IoT (IIoT) testbed that supervised the water storage tank's water level. They recorded cyber-attacks, including SQL injection, command injection, backdoor attacks, and Wireshark's normal traffic in their testbed to generate a real dataset. In IDS, 23 features were extracted, and its method was designed as two classes sample in attack or normal. After collecting data, ML was used to detect anomalies with seven different classifiers such as SVM, KNN, NB, RF, DT, LR, and ANN. In the final step, accuracy and sensitivity were analyzed as performance metrics. However, multiple algorithms and preprocessing data may read to better performance.

Jan et al.[59] presented a slight attack detection through an ML strategy based on SVM into the IoT network. The suggested method contains two steps, the training, and the evaluation step. In the training step, a labeled dataset was given, and features were extracted from samples. So, the result of extracted features with their labels was used to train the classifier. Second, the same features were extracted from the test dataset. In the last step, the trained classifier distinguishes benign and anomaly. Also, packet arrival was the only attribute to the node that was considered. Thirty-nine features were selected from the dataset, including destination port, flow duration, total backward packets, and other features. CICIDS2017 dataset[¶¶]was used in this research, which is available online to access and apply as IDS and prevention systems. However, just DDoS attacks were evaluated in this research, while other types of attacks were not considered.

Anthi et al.[60] provided an original supervised method to detect attacks based on IoT network information by three main steps. First, the IDS system scans the network to collect and classify IoT devices connected to the Internet. Second, the detection of infected packets as malicious or normal. Finally, it classifies the type of attacks that includes 12 attacks in 4 principal categories of DoS (TCP/UDP/hello flood), MITM (ettercap, ARP)/spoofing, reconnaissance, and replay attacks. The correlation evaluation filter and gain ratio were selected to extract features in the ML approach. Evaluation filter provided in Weka. Its empirical dataset was captured by TCP-dump for three weeks and saved in a PCAP format. Nevertheless, this research needs to evaluate more complex and widespread attacks.

A multi-objective particle swarm optimization (MOPSO) to detect harmful behaviors was developed by Habib et al.[61] This research transformed feature selection problems into the binary search space of MOPSO. The error rate (ER) and several selected features are used in this algorithm's fitness function to find the optimized selected feature subsets from share space. In the final step, 5-NN evaluated the answer to distinguish malicious and benign traffic data. Its dataset was collected from nine real IoT datasets from the UCI repository with diverse attacks includes UDP flooding, spam data, UDP and TCP flooding, Scan for vulnerable devices, SYN flooding, ACK flooding, and UDP flooding with fewer options. However, it has a high computational cost to detect attacks.

Al Shorman et al.[62] introduced a botnet detection technique based on the one-class support vector machine (OCSVM) and grey wolf optimization algorithm (GWO). In this research study, the principal step of this botnet detection includes five phases. In the first step, problem understanding was considered to disconnect IoT devices from the network and negatively affect compromised IoT devices. Second, N-BaIoT real dataset was applied from an ML repository . In the third phase, data is cleaned, integrated, normalized, and reduced. Also, in the model development phase, the advantages of GWO and OCSVM were taken to execute feature selection. Finally, the evaluated model was compared with OCSVM, isolation forest (IF), and local outlier factor (LOF). Nevertheless, the IoT devices in their dataset were limited, and researchers could not detect multi-class attacks.

Li et al.[63] introduced a novel method based on an online learning (OL) algorithm to detect malicious nodes. In this research study, the OL algorithm was used to discover each node's trust and then trained the model with the K-means algorithm to classify nodes into malicious and benign nodes. In general, this research includes twelve steps, first, injection packets into benign nodes, then constructing the detection equation set. In the third step, the OL algorithm was trained to model the truncated gradient and online gradient descent with a detection equation. The following step was evaluating the trust of nodes and k-means clustered IoT nodes. Then, the result was analyzed and constructed a new detection equation set, training the model, accumulating the model's parameter, and compute all nodes' trust. In the final step, K-means clusters all nodes into two benign and malicious groups. But just malicious nodes were detected, and different types of attacks have not been discovered.

Almiani et al.[64] proposed an innovative method to detect intrusion in IoT with a deep proportional recursive network structure. The intelligence intrusion technique has been implemented in the fog computing layer, which contains two significant engines. First, traffic connection has been recorded and processed into an acceptable format for processing by the deep NN. In another engine, normal and abnormal are selected for connections. However, diverse types of attacks have not been detected.

Keserwani et al.[65] introduced both smart and efficient intrusion detection approaches based on the grey wolf optimization (GWO), particle swarm optimization (PSO), and RF. In this research study, data were collected, preprocessed trained from various datasets in the first step. In the next step, feature selection techniques were applied to reduce overfitting and improve accuracy. Next, rules were taken into consideration for reduction and making decisions. Finally, a testing model was used to analyze the model performance. However, it was not implemented in the real world.

Qureshi et al.[66] presented a combination of deep neural network (DNN) and LSTM to detect multi-class malware families. First, data were used from available sources in different formats from nine diverse devices and preprocessed. Next, the second step is training hybrid classifiers and developing them. In the final phase, the proposed algorithms were applied to predict attacks by traffic. Nevertheless, the proposed approach did not apply in real environment.

Jeong et al.[67] provided a new approach to detecting five active network anomalies. First, a BroadR-Reach testbed has been built to network packets through online learning in PCAP format. In the following step, a feature generation and a CNN-based model contain two hidden sub-layers. Two dense layers are considered to identity intrusion by continuous stream changes with high accuracy. However, this research only considered one type of stream, and the result could not detect various attacks as a multi-class detection.

Laghrissi et al.[68] introduced intrusion detection systems for IoT devices based on neural networks and LSTM to detect anomalies. Also, some reduction algorithms were used to decrease the number of features, which helped faster detection with acceptable accuracy. Finally, only through 2-features and roughly 99% accuracy, the system could identify an attack. On the other hand, the proposed model only compared with other LTSM algorithms, and its data were not real.

4.2.1 | The review of attack detection techniques

Table 5 shows a side-by-side evaluation of the abovementioned research studies and the important aspects of evaluating the IoT environment's attack detection techniques. The most important features of malware detection techniques are based on network traffic such as internet protocol address, protocols, frame length, frame number, host number, and so forth.

Table 6 portrays datasets of the above research studies and the number of instances (attack and normal traffic), file format (CSP/PCAP), publicity, recording environment (emulate/real), labeled, and resources of each paper that can download them their labels. In general, emulated datasets include normal and abnormal samples, which are widespread datasets that researchers use.

4.3 | Overview of ransomware detection techniques in IoT

Azmoodeh et al.[78] proposed ransomware detection in IoT by ML algorithms. In this research, the energy consumption of ransomware's fingerprint was considered a detection technique that used the Power-Tutor tool to track and record power usage. After monitoring and recording power consumption, the correct label was assigned to them through the Power-Tutor tool from three android devices. Finally, the ML algorithm was applied to distinguish ransomware from benign applications in MATLAB. In this research study, classification algorithms were KNN, NN, SVM, and RF. Nevertheless, it needs to evaluate in the real world with more devices.

Additionally, Hirano and Kobayashi[79] introduced an automated ransomware detection by ML from a live-forensic hypervisor. First, storage access patterns are obtained from ransomware samples and non-malicious applications by a forensic hypervisor called WaybackVisor of storage devices that contain all read and write requests. In the following step, five-dimensional features (entropy, the total amount of reading, variance of the

TABLE 5 Attack detection techniques in IoT during a 5-year period

		Detection				Accuracy		Compared	Classification	Security	
Year	Ref.	technique	Advantage	Disadvantage	Classifier	rate	Tool	with	type	threat	Supervised
2018	Yang et al.[48]	AL	High accuracy	Inefficient comparison	XGBoost	–	–	Random select	Binary	–	Yes
	Meidan et al.[49]	DL	High accuracy	Inefficient comparison	IF, SVM, LOF	–	–	Other classifiers	Binary	BASHLITE, Mirai	No
	Rathore and Park[50]	ML	Distributed architecture	Low performance	SVM, LR, KNN, DT, RF, BN	86.53%	–	Traditional ML	Binary	–	Yes
	Moustafa et al.[51]	EL	High accuracy	Lack of variant data protocols	DT, NB, ANN	98%	–	ML	Binary	Botnet	NA
	Diro and Chilamkurti[55]	DL	Acceptable time detection	Inefficient comparison	–	Over 96%	–	Shallow learning algorithm	Binary	DoS, Probe, R2L, U2R	NA
2019	Hasan et al.[56]	ML	Strong data preprocess	Weak detect algorithm	LR, SVM, ANN, RF, DT	99.4%	Python	Other classifiers	Multi-class	Botnet	NA
	Otoum et al.[57]	DL	Reduce training time	Lack of different classifiers	KNN, DT, SVM	99.02%	Python	Deep feature embedding learning and distributed DL	Multi-class	DoS, U2R, probe, R2L	Yes
	Zolanvari et al.[58]	ML	High accuracy	Small attack dataset	RF, DT, KNN, ANN, SVM, LR, NB	99.99%	Python	Other classifiers	Binary	SQL injection, backdoor, command injection	NA
	Jan et al.[59]	ML	High accuracy	Lack of variety of attacks	SVM, NN, KNN, DT	98%	MATLAB	59-61	Binary	DDoS	Yes
	Anthi et al.[60]	ML	Detect different types of attack	Limited features	NB, BN, J48, Zero R, OneR, simple logistic, SVM, multi-layer perceptron, RF	–	Python	Other classifiers	Multi-class	DoS, MITM, spoofing, reconnaissance, replay	Yes
2020	Habib et al.[61]	ML	Big and variant dataset	High computational cost	K-NN	–	MATLAB	Traditional ML, NSGA-II	Binary	Mirai botnet, Gafgyt botnet	Yes
	Al Shorman et al.[62]	ML and GWO	Lowest detection time	Limited IoT devices	SVM	–	Python	IF, LOF, OCSVM	Binary	DDoS	No
	Li et al.[63]	OL	High accuracy	Not discovering diverse attacks	TG, OGD	–	Python	Common and enhanced TG, OGD algorithm	Binary	NA	No
	Almiani et al.[64]	Multi-layered deep recursive NN	Fast execution time	Lack of multi-class detection	ANN	92%	MATLAB	61-63	Binary	DoS, probe, R2L, and U2R	Yes
2021	Keserwani et al.[65]	GWO-PSO-RF	High accuracy	Lack of real data and real implementation	RF	99.66%	Python	Other classifiers	Multi-class	NA	Yes
	Qureshi et al.[66]	DNN & LSTM	Multi-class detection with high accuracy	Lack of real data	DNN	99.96%	Python	–	Multi-class	Mirai, Gafgyt	Yes
	Jeong et al.[67]	OL	Real data	Lack of detect various attack	CNN	0.995%	Python	Five-fold cross-validation	Binary	DDOS	Yes
	Laghrissi et al.[68]	ML	Limited features	Lack of real data	LSTM	99.49%	Python	69-72	Multi-class	DOS, R2L	Yes

TABLE 6 Network-based dataset of varied international studies

Ref.	Attack data	Normal data	Resource	Labeled	Recording	# attributes	Format	Public	# instances
48	Yes	Yes	KDD 1999[73]	Yes	Emulate	42	CSV	Yes	4,000,000
48	Yes	Yes	AWID trained set[74]	Yes	Emulate	155	CSV	Yes	162,375,247
48	Yes	Yes	AWID test set	Yes	Emulate	155	CSV	Yes	48,524,866
49, 62, 66	Yes	Yes	N_BaIoT	Yes	Real	115	PCAP	Yes	7,062,606
50, 55, 57, 64, 65, 68	Yes	Yes	NSL-KDD trained set	Yes	Emulate	42	CSV	Yes	4,898,431
50, 55, 57, 64, 65, 68	Yes	Yes	NSL-KDD test set	Yes	Emulate	42	CSV	Yes	311,027
51, 75	Yes	Yes	UNSW-NB15[54]	Yes	Emulate	49	PCAP & CSV	Yes	668,000
51, 75	Yes	Yes	NIMS[55]	Yes	Emulate	23	PCAP	Yes	713,851
47	Yes	Yes	Kaggle	Yes	Emulate	13	CSV	Yes	357,952
49	Yes	Yes	-	Yes	Real	23	CSV	Yes	225,745
50, 65	Yes	Yes	CICIDS2017	Yes	Emulate	78	CSV	Yes	225,745
52	Yes	Yes	76	Yes	Emulate	121	PCAP	Yes	220,785
53	Yes	Yes	77	Yes	Real	–	–	Yes	6,993,107

TABLE 7 Comparison of ransomware detection techniques in IoT

Year/Ref.	Techniques	Advantage	Disadvantage	Feature	Classification	Tool	Extractor	Type
2018[78]	ML	High accuracy	Small dataset and devices	Power consumption	KNN, NN, SVM, RF	MATLAB	Power-Tour	Binary
2019[79]	ML	High accuracy	ineffective comparison	Entropy, variance, read and write requests	RF, KNN, SVM	Python	Wayback Visor	Binary
2020[75]	LOA	Meta-heuristic algorithm	Lack of diversity in the dataset	CPU & memory consumption, read bytes, write bytes, process active time	–	Snort and Honey-Pot	Python	Binary

logical block address (write), the total amount of write and variance of the logical block address (read)) from actual ransomware and a benign program's input and output logs were extracted. Finally, diverse ML models include RF, SVM, and KNN, were used to distinguish benign and ransomware logs. However, a small dataset and ineffective comparison are its disadvantages.

Chakkaravarthy et al.[75] presented a powerful ID Honey-Pot to detect ransomware, including three subsystems (honey folder, audit watch, and event processing). First, events from various sources were collected to feature extraction. In the second level, normal and abnormal were distinguished through a range of values. In the following level, correlation level 1 and level 2 were considered by logical, spatial, and temporal constructs from the complex event. Nevertheless, other types of ransomware and malware were not analyzed.

Table 7 summarizes the studies mentioned above, describing the detection technique, advantages, disadvantages, main features to extract from files to detect anomalies, classifier algorithms, the selected tool to simulate the proposed method, the tool to extract features from ransomware and non-malicious files, and the classification type of the detection include binary and multi-class.

5 | RESULTS AND DISCUSSION

The review of the selected threat detection techniques in IoT was elaborated in former sections, which are defined in three main categories: malware detection, attack detection, and ransomware detection techniques. Section 4 presents a statistical analysis of threat detection techniques in IoT. Besides, planned research questions in Section 3 will be analyzed with analytical reports.

5.1 | RQ1: What strategies are provided in IoT?

The percentage of diverse threat detection techniques percentage up to now is provided in Figure 6, refer to the taxonomy in Section 4 shows that attack detection techniques have the highest percentage of threat detection methods by 58% of the literature. Of course, malware detection techniques have 32%, and ransomware detection techniques have 10% usage in threat detection methods in the IoT environment.

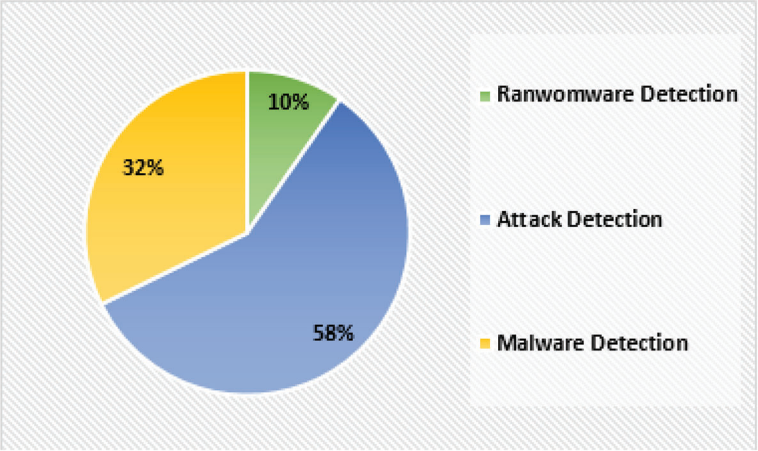


FIGURE 6 The percentage of diverse distribution of existing strategies in IoT

5.2 | RQ2: What types of solutions are applied to detect the threat in IoT?

The used solutions in different approaches are depicted in Figure 7A considering RQ2. Figure 7A distinguishes diverse applied solutions into three categories: ransomware detection, malware detection, and attack detection. The highest percentage of the applied solution is on ML algorithms in all groups. Also, DL algorithms in attack detection and malware detection techniques have a significant percentage. Furthermore, attack detection techniques have diverse solutions, including AL, EL, OL, and deep NN.

In general, the leading solutions for threat detection techniques are shown in Figure 7B. We observed that ML techniques have the most usage with 31 studies. After ML, DL techniques have the highest rate, among other techniques. Looking at the pie chart in more detail, it is evident that machine learning techniques presented 58%, following by deep learning at 23%. Also, 7% of total studies belong to online learning. In addition, other solutions (such as AL, DNN, LOA, EL, and signature base) have a similar percentage (3%).

5.3 | RQ3: What types of features are applied to detect the threat in IoT?

As claimed by Figure 8A, it can be discovered that, in 58% of the research studies, features were used to form network traffic data for attack detection. Also, 20% and 10% of papers in the IoT environment used Op-Code and power consumption to detect malware, respectively. Finally, each

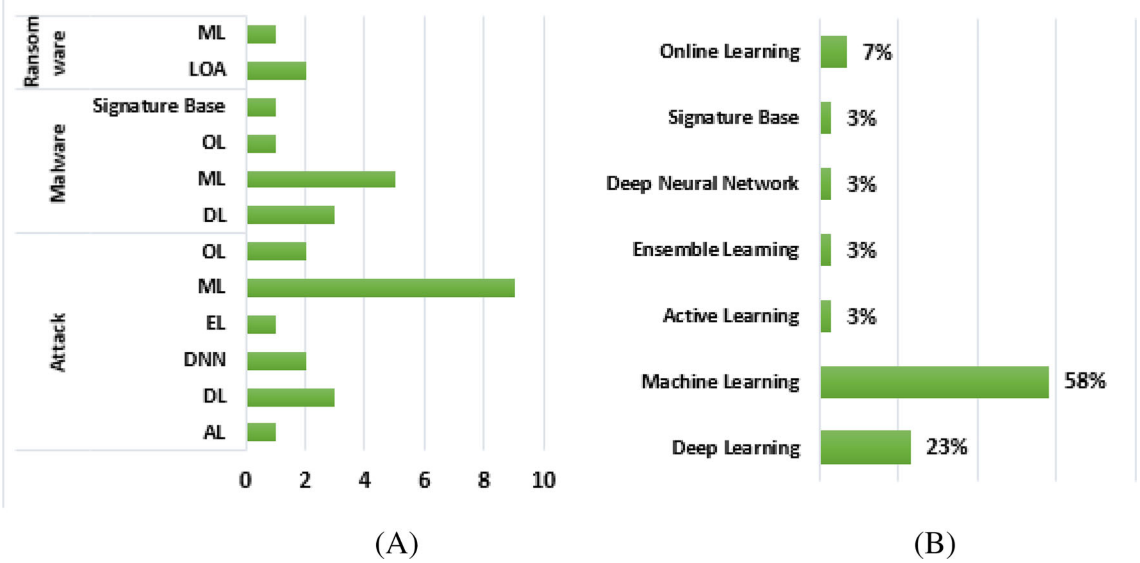


FIGURE 7 The number of widespread solutions and their importance among researchers in order to detect malware, ransomware, and intrusion. (A) Divided by detection method. (B) Divided by solutions

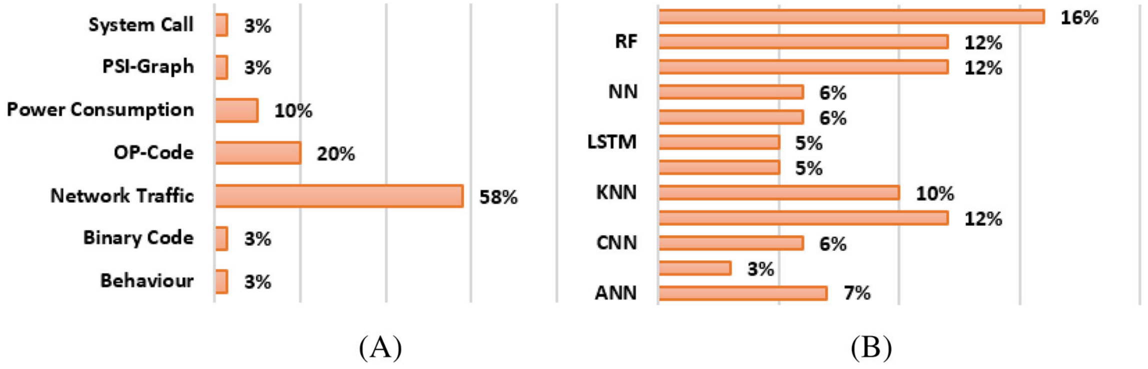


FIGURE 8 The percentage of using various features and classifiers to detect threats. (A) Features. (B) Classifiers

feature includes a binary code of malware, malware's behavior, system calls, and PSI-Graph covered 3% of the studies' evaluation. It can be concluded that network traffic data has the highest rate among other features to detect threats.

5.4 | RQ4: Which classification algorithms are appropriate in threat detection techniques in IoT?

Regarding RQ4, most research studies have focused on the following classifier algorithms: SVM, KNN, DT, RF, and ANN. The classification algorithms and their statistical percentage classifiers used in threat detection techniques based on the learning method to classify intrusion and benign data are depicted in Figure 8B. The statistical percentage of the classification algorithms shows that SVM has the highest usage in the classification algorithms by 16%. Both DT and RF have 12%, while ANN has only 7%. Other classifiers (XGBoost, autoencoder, isolation forest, LOF, AdaBoost, OGD, fuzzy pattern, TG, NB, BN, LR) have less than 6%. Therefore, SVM, DT, and RF are the most accurate algorithms to distinguish normal and malicious features.

5.5 | RQ5: What types of tools are used to support threat detection techniques in IoT?

Figure 9 clarifies the statistical portion of various simulation environments and tools. Looking at the bar chart in more detail, it is evident that Python accounted for the highest proportion of tools to simulate IoT threat detection solutions, followed by MATLAB, which are supported more than other tools (16 and 4, respectively). Nevertheless, 9 of the researchers did not introduce their tools, while 2 studies were used Linux Sandbox to simulate their ideas. Hence, using Python is one of the straightforward simulations among researchers to implement their approaches.

5.6 | RQ6: Which data resources are suitable to use in IoT threat detection techniques?

According to Figure 10A, the VirusTotal website is the most common resource to collect malware in the malware detection category, in which 5 research studies in this SLR were used. The VirusShare website, Honey-Pot, and collecting data from real IoT devices are the other methods to gather data, which accounted for 2, 2, and 1 research papers through this SLR.

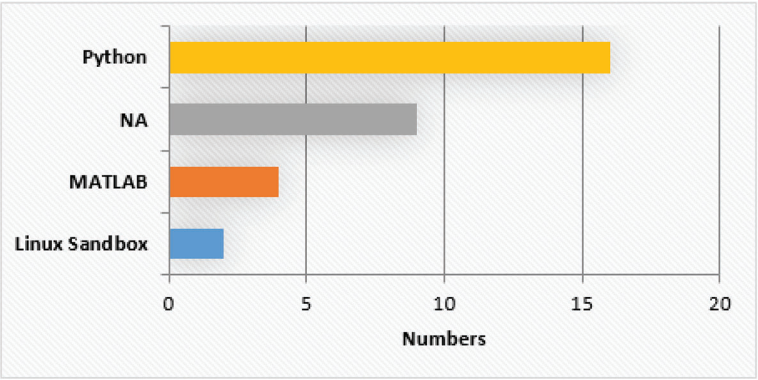


FIGURE 9 The percentage of standard tools to simulate IoT threat detection solutions

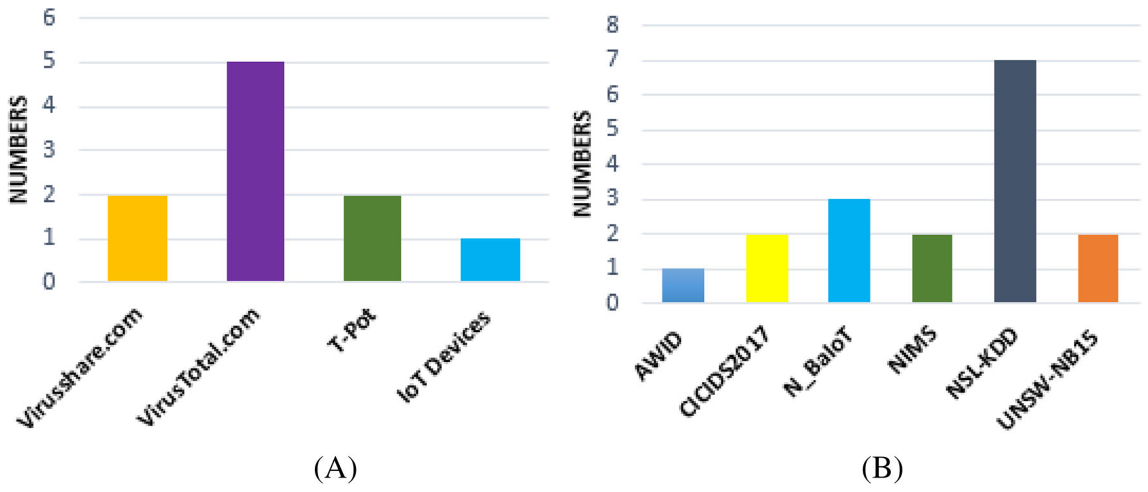


FIGURE 10 The percentage of using different online dataset to download malware and ransomware, and major attack datasets. (A) Malware detection data resources. (B) Attack detection datasets

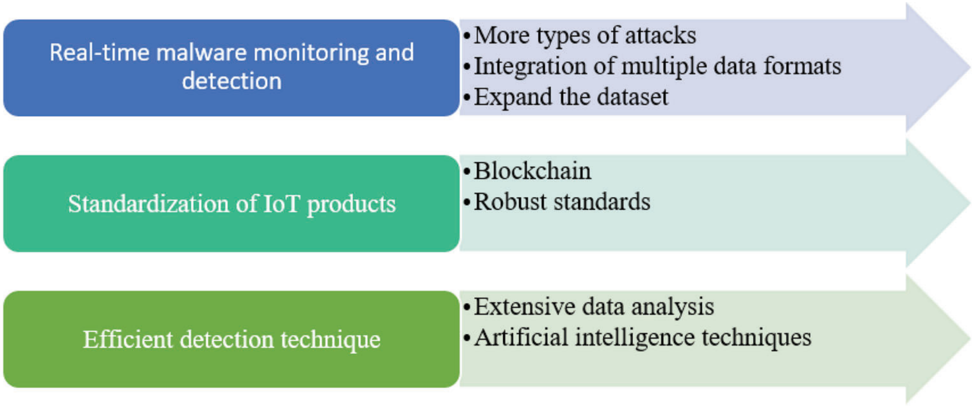
Also, according to Figure 10B, NSL-KDD was selected as a widespread dataset attack detection techniques, which 7 studies used in their researches, following by N-BaIoT, presented at 3 studies. However, AWID, UNSW, and CICIDS are prevalent among researchers.

With the access of threat detection techniques in IoT, three new paradigms appear ID, malware detection, and ransomware detection. Despite advanced research on the study assortment of detection techniques in the IoT environment, some open issues are still significant challenges. Suggested solutions do not treat the strength and weaknesses of each threat detection method entirely. Besides, some attacks and IoT technology have been concentrated in researches. So, we portray a detailed view of those three categories. Future detection methods for IoT need to expand the threat detection scope and view the intended application requirements.

6 | THREAT DETECTION CHALLENGES AND PROPOSED MEASURES

Based on the above evaluation, the IoT sector has serious challenges in threat detection, which need to be solved for the IoT future progress. Figure 11 shows a detailed analysis of various research challenges, research gaps, and possible solutions in IoT threat detection, including real-time detection, lack of sufficient dataset, a wide array of various attacks and standards in IoT products.

- Challenge 1: Real-time malware monitoring and detection is still a big challenge. The majority of advanced research use available datasets and simulation environments to test the validity of the proposed algorithm, which are not appropriate for real-time analysis, while malware detection techniques require running in the real world.


Real-time malware monitoring and detection
- More types of attacks
- Integration of multiple data formats
- Expand the dataset

Standardization of IoT products
- Blockchain
- Robust standards

Efficient detection technique
- Extensive data analysis
- Artificial intelligence techniques

FIGURE 11 A logical mapping of comprehensive challenges and possible solutions in thereat detection

TABLE 8 Abbreviations

IoT, Internet of Things	DNS, domain name system	SLR, systematic literature review	ID, intrusion detection
DoS, denial of service	ML, machine learning	IDS, intrusion detection systems	DL, deep learning
DDoS, distributed denial-of-service	RQ, research questions	OL, online learning	AL, active learning
EL, ensemble learning	NN, neural network	LOA, lion optimization algorithm	Op, Code-Operation Codes
LSTM, long short term memory	BNN, bidirectional neural networks	RNN, recurrent neural network	RF, random forest
SVM, support vector machine	NB, naive Bayes	MLP, multi-layer perceptron	KNN, K nearest neighbors
DT, decision tree	PSI, printable string information	CNN, convolution neural network	MFP, maximal frequent patterns
SSH, Secure Shell	API, application programming interface	CCTV, closed-circuit television	

- Proposed measure A: Expand the dataset near real-time systems is a significant solution in threat detection techniques in IoT. Nevertheless, more researchers are using an available dataset that does not contain new attacks. Also, the number of malware samples is not equal to benign samples in extant datasets. Therefore, publishing a real dataset will help scholars to have real analysis.
- Proposed measure B: Regarding detection techniques, few attacks are focused on advanced research by authors, which considering and analyzing more types of attacks would be future research directions. So, increasing the threat detection range will help authors to work near the real world.
- Proposed measure C: Integration of multiple data formats would be another solution for threat detection near real-time due to different data formats in incoming frames. However, it does not consider the study assortment of threat detection or even malware diversity in available datasets.
- Challenge 2: Standardization of IoT products can be an open challenge in the IoT environment for future research.[80] IoT devices suffer from different threats, such as threats, malware, ransomware, and attacks, because they do not have a valid standard. Some requirements of various integrated security measures in IoT products are necessary.
- Proposed measure A: Using blockchain to secure a communication environment can also be an open issue for future works because of decentralized, efficient, and transparent blockchain performance.
- Proposed measure B: Introducing some robust standards based on the IoT would effectively reduce several attacks between IoT devices through diverse layers.
- Challenge 3: The efficient detection technique is one of the essential factors in IoT.[81] IoT devices have less computation power, low memory, and short battery life. Researchers cannot use heavy DL algorithms, ML algorithms, and other heavy intelligent algorithms. It requires lightweight ML algorithms as a detection or prevention method in IoT.
- Proposed measure A: In the reviewed paper,[82] it is observed that extensive data analysis to combat attacks and threats in IoT will be an exciting direction for future works. Precise evaluation would increase productivity to detect diverse attacks.
- Proposed measure B: Artificial intelligence approaches such as learning techniques would help threat detection significantly. Such as ML, DL, AL, and other learning techniques.

All of the commonly used words of abbreviations are listed in Table 8.

7 | CONCLUSION AND LIMITATION

This SLR's objective was to produce an overview of the advanced threat detection techniques in the IoT environment. Threat detection techniques and IoT were investigated in three categories: malware, ransomware, and threat detection techniques in IoT. We discussed 31 principal research studies among the 124 research studies from our search query that were chosen between 2017 and August 2021. Regarding the systematic review, we classified 31 selected papers into three main categories as follows: malware detection techniques (32%), attack detection techniques (58%), and ransomware detection techniques (10%), according to RQ1. We also determined the main idea, advantage, disadvantage, tool type, classifies, features, evaluation type, and dataset for each IoT threat detection technique. The majority of research studies try to enhance the accuracy rate with new algorithms. However, real-time environment, lightweight solution, short time detection, varied dataset, diverse formats of data, and actual data are rarely used. As a considerable effort is made to plan a systematic review, the ML approach has the highest percentage of the threat detection solutions in this literature, besides other learning techniques and signature-based solutions, according to RQ2. According to RQ3, we observed that network traffic features have the most percentage with 14 studies, binary code, Op-Code malware behavior, PSI-Graph, system call, and power consumption were other observed features in other studies. For the RQ4, we observed that all papers evaluated various classification algorithms, but SVM (16%) was the most popular classifier in 31 studies. Python and MATLAB are used in most of the selected studies to simulate proposed approaches based on RQ5. With respect to RQ6, VirusShare and VirusTotal websites are the most widespread resources to support data resources. However, some restrictions of this SLR are as follows:

- Study scope: This SLR covered only academic international reputed conferences and journals. Also, national journals, national conferences, survey papers, books, non-English papers, editorials papers, short research studies, and non-peer-reviewed articles have been omitted.
- Publication bias: This SLR, regarded online article repositories were supported. These online data repositories introduce valid papers. Furthermore, the writers can claim that all selected papers are related to IoT.
- Study taxonomy: Malware detection, attack detection, and ransomware detection classify threat detection in this research. However, other categories might be possible.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ENDNOTES

[*] http://dl.acm.org/dl.cfm
[†] http://www.sciencedirect.com/
[‡] https://www.springer.com/
[§] https://journals.sagepub.com/
[¶] http://ieeexplore.ieee.org/Xplore/home.jsp
[#] https://onlinelibrary.wiley.com/
[||] https://taylorandfrancis.com/journals/
[**] https://www.virustotal.com
[††] https://archive.org/details/vxheaven-windows-virus-collection
[‡‡] https://www.kaggle.com/c/malware-classification
[§§] https://virusshare.com
[¶¶] https://www.unb.ca/cic/datasets/ids-2017.html

ORCID

Nasim Soltani https://orcid.org/0000-0001-8133-523X
Amir Masoud Rahmani https://orcid.org/0000-0001-8641-6119
Mehdi Hosseinzadeh https://orcid.org/0000-0003-1088-4551

REFERENCES

1. Asghari P, Rahmani AM, Javadi HH. Internet of Things applications: a systematic review. *Comput Netw*. 2019;148:241-261.
2. Rajkumar B, Dastjerdi AV. *Internet of Things: Principles and Paradigms*. Elsevier; 2016.
3. Wazid M, Das AK, Rodrigues JJ, Shetty S, Park Y. IoMT malware detection approaches: analysis and research challenges. *IEEE Access*. 2019;7:182459-182476.
4. Kambourakis G, Kolias C, Geneiatakis D, Karopoulos G, Makrakis GM, Kounelis I. A state-of-the-art review on the security of mainstream IoT wireless PAN protocol stacks. *Symmetry*. 2020;12(4):579.
5. Demirkol CM, Josep P, Marc C, Wasiu P, Harald H. Powering the Internet of Things through light communication. *IEEE Commun Mag* 2019;57(6):107-113.
6. Lee I, Lee K. The Internet of Things (IoT): applications, investments, and challenges for enterprises. *Bus Horizons*. 2015;58(4):431-440.
7. Makhdoom I, Mehran A, Justin L, Ren PL, Wei N. Anatomy of threats to the Internet of Things. *IEEE Commun Surv Tutor*. 2018;21(2):1636-1675.
8. Mahmoud A, Russello G, Crispo B. Internet of Things: a survey on the security of IoT frameworks. *J Inf Secur Appl*. 2018;38:8-27.
9. Xu LD, He W, Li S. Internet of Things in industries: a survey. *IEEE Trans Ind Inform*. 2014;10(4):2233-2243.
10. Brous P, Janssen M, Herder P. The dual effects of the Internet of Things (IoT): a systematic review of the benefits and risks of IoT adoption by organizations. *Int J Inf Manag*. 2020;51(4):101952.
11. Anagnostopoulos M, Spathoulas G, Viaño B, Augusto-Gonzalez J. Tracing your smart-home devices conversations: a real world IoT traffic data-set. *Sensors*. 2020;20(22):6600.
12. Basole S, Stamp M. Tracing your smart-home devices conversations: a real world IoT traffic data-set. In: Stamp M, Alazab M, Shalaginov A, eds. *Malware Analysis Using Artificial Intelligence and Deep Learning*; Springer International Publishing; 2021:361-379.
13. Sembiring Z. Stuxnet threat analysis in SCADA (Supervisory Control and Data Acquisition) and PLC (Programmable Logic Controller) Systems. *J Comput Sci Inf Technol Telecommun Eng*. 2020;1(2):96-103.
14. Sriram S, Vinayakumar R, Alazab M, Soman K. Network flow based IoT botnet attack detection using deep learning. Proceedings of the IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); 2020:189-194.
15. Marzano A, Alexander D, Fonseca O, et al. The evolution of Bashlite and Mirai IoT botnets. Proceedings of the IEEE Symposium on Computers and Communications (ISCC); 2018:813-818.
16. Vignau B, Khoury, R. & Hallé, S. 10 years of IoT malware: a feature-based taxonomy. Proceedings of the IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C); 2019:458-465.

17. Wang H, Zhang W, He H, et al. An evolutionary study of IoT malware. *IEEE Internet Things J*. 2021.
18. Sectigo evolution of IoT attacks: an interactive infographic; 2021. https://sectigo.com/uploads/resources/Evolution-of-IoT-Attacks-Interactive-IG_May2020.pdf
19. Agnihotri A, Bhattacharya S. *Ring: The New Amazon Subsidiary and the Social, Privacy, and Security Issues It Generates*. SAGE Publications, SAGE Business Cases Originals; 2021.
20. Davis BD, Mason JC, Anwar M. Vulnerability studies and security postures of IoT devices: a smart home case study. *IEEE Internet Things J*. 2020;7(10):10102-10110.
21. Kimani K, Oduol V, Langat K. Cyber security challenges for IoT-based smart grid networks. *Int J Critical Infrastr Protect*. 2019;25:36-49.
22. Khan MA, Salah K. IoT security: review, blockchain solutions, and open challenges. *Future Gener Comput Syst*. 2018;82:395-411.
23. Yaqoob I, Ahmed E, ur Rehman MH, et al. The rise of ransomware and emerging security challenges in the Internet of Things. *Comput Netw*. 2017;129:444-458.
24. Zarpelão BB, Miani RS, Kawakani CT, de Alvarenga SC. A survey of intrusion detection in Internet of Things. *J Netw Comput Appl*. 2017;84:25-37.
25. Sharmeen S, Huda S, Abawajy JH, Ismail WN, Hassan MM. Malware threats and detection for industrial mobile-IoT networks. *IEEE Access*. 2018;6:15941-15957.
26. Elrawy MF, Awad AI, Hamed HF. Intrusion detection systems for IoT-based smart environments: a survey. *J Cloud Comput*. 2018;7(1):1-20.
27. Clincy V, Shahriar H. IoT malware analysis. Proceedings of the IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC 2019); Vol. 1, 2019:920-921.
28. Neshenko N, Bou-Harb EJ, Crichigno GK, Ghani N. Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. *IEEE Commun Surv Tutor*. 2019;21(3):2702-2733.
29. Chaabouni N, Mosbah M, Zemmari A, Sauvignac C, , Faruki P. Network intrusion detection for IoT security based on learning techniques. *IEEE Commun Surv Tutor* 2019; 21(3): 2671-2701.
30. Dange S, Chatterjee M. IoT botnet: the largest threat to the IoT network. In: Jain L, Tsihrintzis G, Balas V, Kumar Sharma D, eds. *Data Communication and Networks*. Springer; 2020:137-157.
31. Mishra N, Pandya S. Internet of Things applications, security challenges, attacks, intrusion detection, and future visions: a systematic review. *IEEE Access*. 2021;9:59353-59377.
32. Ye Y, Li T, Adjeroh D, Iyengar SS. A survey on malware detection using data mining techniques. *ACM Comput Surv (CSUR)*. 2017;50(3):1-40.
33. Dai SY, Fyodor Y, Wu MW, Huang Y, Kuo SY. Holography: a behavior-based profiler for malware analysis. *Softw Pract Exper*. 2012;42(9):1107-1136.
34. de Rey AM. Mathematical modeling of the propagation of malware: a review. *Secur Commun Netw*. 2015;8(15):2561-2579.
35. Abbas MFB, Srikanthan T. Low-complexity signature-based malware detection for IoT devices. *Int Conf Appl Techn Inf Secur*. 2017;8(15):181-189.
36. Azmoodeh A, Dehghantanha A, Choo KKR. Robust malware detection for internet of (battlefield) things devices using deep eigenspace learning. *IEEE Trans Sustain Comput*. 2018;4(1):88-95.
37. HaddadPajouh H, Dehghantanha A, Khayami R, Choo KKR. A deep recurrent neural network based approach for Internet of Things malware threat hunting. *Future Gener Comput Syst*. 2018;85:88-96.
38. Nguyen HT, Ngo QD, Le VH. A novel graph-based approach for IoT botnet detection. *Int J Inf Secur*. 2019;1-11.
39. Darabian H, Dehghantanha AS, Hashemi SH, Choo KKR. An opcode-based technique for polymorphic Internet of Things malware detection. *Concurr Comput Pract Exper*. 2020;32(6):5173.
40. Dovom EM, Azmoodeh A, Dehghantanha A, Newton DE, Parizi RM, Karimipour H. Fuzzy pattern tree for edge malware detection and categorization in IoT. *Concurr Comput Pract Exper*. 2019;97:1-7.
41. Karanja EM, Masupe S, Jeffrey MG. Analysis of Internet of Things malware using image texture features and machine learning techniques. *IoT*. 2020;9:100153.
42. Tien CW, Chen SW, Ban T, Kuo SY. Machine learning framework to analyze IoT malware using ELF and opcode features. *Digital Threats Res Pract*. 2020;1(1):1-19.
43. Jeon J, Park JH, Jeong YS. Dynamic analysis for IOT malware detection with convolution neural network model. *IEEE Access*. 2020;8:96899-96911.
44. Radhakrishnan G, Srinivasan K, Maheswaran S, Mohanasundaram K, Palanikkumar D, Vidyarthi A. A deep-RNN and meta-heuristic feature selection approach for IoT malware detection. *Mater Today Proc*. 2021.
45. Azmoodeh A. IoT malware detection dataset organization; 2017. https://github.com/azmoodeh/IoTMalwareDetection
46. Homayoun S, Dehghantanha A, Ahmadzadeh M, Hashemi S, Khayami R. Know abnormal, find evil: frequent pattern mining for ransomware threat hunting and intelligence. *IEEE Trans Emerg Top Comput*. 2017.
47. Pa YMP, Suzuki S, Yoshioka K, Matsumoto T, Kasama T, Rossow C. IoTPOT: a novel honeypot for revealing current IoT threats. *J Inf Process*. 2016;24(3):522-533.
48. Yang K, Ren J, Zhu Y, , Zhang W. Active learning for wireless IoT intrusion detection. *Wirel Commun* 2018; 25(6): 19–25.
49. Meidan Y, Bohadana M, Mathov Y, et al. N-baiot–network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervas Comput*. 2018;17(3):12-22.
50. Rathore S, Park JH. Semi-supervised learning based distributed attack detection framework for IoT. *Appl Soft Comput*. 2018;72:79-89.
51. Moustafa N, Turnbull B, Choo KKR. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things. *IEEE Internet Things J*. 2018;6(3):4815-4830.
52. Moustafa N. The UNSW-NB15 dataset; 2015. The Australian Research Data Commons (ARDC). 10.26190/5d7ac5b1e8485.
53. Moustafa N, Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS); 2015:1-6; Canberra, ACT, Australia. 10.1109/MilCIS.2015.7348942
54. The-NIMS-dataset; 2018 [Online]. https://projects.cs.dal.ca/projectx/Download.html
55. Diro AA, Chilamkurti N. Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Gener Comput Syst*. 2018;82:761-768.
56. Hasan M, Islam MM, Zarif MII, Hashem M. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. *IoT*. 2019;7:100059.
57. Otoum Y, Liu D, Nayak A. DL-IDS: a deep learning–based intrusion detection framework for securing IoT. *Trans Emerg Telecommun Technol*. 2019;3803.

58. Zolanvari M, Teixeira MA, Gupta L, Khan KM, Jain R. Machine learning-based network vulnerability analysis of industrial Internet of Things. *IEEE Internet Things J*. 2015;6(4):6822-6834.
59. Jan SU, Ahmed S, Shakhov V, Koo I. Toward a lightweight intrusion detection system for the Internet of Things. *IEEE Access*. 2019;7:42450-42471.
60. Anthi E, Williams L, Słowińska M, Theodorakopoulos G, Burnap P. A supervised intrusion detection system for smart home IoT devices. *IEEE Internet Things J*. 2019;6(5):9042-9053.
61. Habib M, Aljarah I, Faris H, Mirjalili S. Multi-objective particle swarm optimization for botnet detection in Internet of Things. In: Mirjalili S, Faris H, Aljarah I, eds. *Evolutionary Machine Learning Techniques: Algorithms and Applications*. Springer; 2020:203-229.
62. Shorman AA, Faris H, Aljarah I. Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection. *J Ambient Intell Humanized Comput*. 2020;11(7):2809-2825.
63. Li B, Ye R, Gu G, Liang R, Liu W, Cai K. A detection mechanism on malicious nodes in IoT. *Comput Commun*. 2020;151:51-59.
64. Almiani M, AbuGhazleh A, Al-Rahayfeh A, Atiewi S, Razaque A. Deep recurrent neural network for IoT intrusion detection system. *Simul Model Pract Theory*. 2020;101:102031.
65. Keserwani PK, Govil MC, Pilli ES, , Govil P. A smart anomaly-based intrusion detection system for the Internet of Things (IoT) network using GWO-PSO-RF model. *J Reliable Intell Environ* 2021; 7(1): 3-21.
66. Qureshi S, He J, Tunio S, et al. A hybrid DL-based detection mechanism for cyber threats in secure networks. *IEEE Access*. 2021;9:73938-73947.
67. Jeong S, Jeon B, Chung B, Kim HK. Convolutional neural network-based intrusion detection system for AVTP streams in automotive Ethernet-based networks. *Veh Commun*. 2021;29:100338.
68. Laghrissi F, Douzi S, Douzi K, Hssina B. Intrusion detection systems using long short-term memory (LSTM). *J Big Data*. 2021;8(1):1-16.
69. Staudemeyer RC. Applying long short-term memory recurrent neural networks to intrusion detection. *South African Comput J*. 2015;56(1):136-154.
70. Kim J, Kim J, Thu HLT, Kim H. Long short term memory recurrent neural network classifier for intrusion detection. Proceedings of the 2016 International Conference on Platform Technology and Service (PlatCon); 2016:1-5.
71. Shen Y, Zheng K, Wu C, Zhang M, Niu X, Yang Y. An ensemble method based on selection using bat algorithm for intrusion detection. *Comput J*. 2015;61(4):526-538.
72. Khan RU, Zhang X, Alazab M, Kumar R. An improved convolutional neural network model for intrusion detection in networks. Proceedings of the 2019 Cybersecurity and Cyberforensics Conference (CCC); Vol. 61, 2019:74-77; IEEE.
73. Kolias GKC, Stavrou A, Gritzalis S. The UCI KDD archive of large data sets for data mining research and experimentation. *SIGKDD Explor*. 2000.
74. Khan RU, Zhang X, Alazab M, Kumar R. Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset. *IEEE Commun Surv Tutor*. 2016;18(1):184-208.
75. Chakkaravarthy SS, Sangeetha D, Cruz MV, Vaidehi V, Raman B. Design of intrusion detection honeypot using social leopard algorithm to detect IoT ransomware attacks. *IEEE Access*. 2020;8:169944-169956.
76. To access dataset, send a request by contacting the corresponding author. anthies@cardiff.ac.uk
77. Dua G. UCI machine learning repositor; 2017. http://archive.ics.uci.edu/ml
78. Azmoodeh A, Dehghantanha A, Conti M, , Choo KKR. Detecting crypto-ransomware in IoT networks based on energy consumption footprint. *J Ambient Intell Humanized Comput* 2018; 9(4): 1141-1152.
79. Hirano M, Kobayashi R. Machine learning based ransomware detection using storage access patterns obtained from live-forensic hypervisor. Proceedings of the 2019 6th International Conference on Internet of Things: Systems, Management and Security (IOTSMS); Vol. 9, 2019:1-6; IEEE.
80. Hajjaji Y, Boulila W, Farah IR, Romdhani I, , Hussain A. Big data and IoT-based applications in smart environments: a systematic review. *Comput Sci Rev* 2021; 39: 100318.
81. Khraisat A, Alazab A. A critical review of intrusion detection systems in the Internet of Things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity*. 2021;4(1):1-27.
82. Yaqoob I, Hashem IAT, Ahmed A, Kazmi SA, Hong CS. Internet of Things forensics: recent advances, taxonomy, requirements, and open challenges. *Future Gener Comput Syst*. 2019;92:265-275.

How to cite this article: Soltani N, Rahmani AM, Bohlouli M, Hosseinzadeh M. Artificial intelligence empowered threat detection in the Internet of Things: A systematic review. *Concurrency Computat Pract Exper*. 2022;e6894. doi: 10.1002/cpe.6894