Machine Learning Intrusion Detection as a Solution to Security and Privacy Issues in IoT: A Systematic Review

*Olufunke G. Darley, Adetokunbo A. Adenowo and Abayomi I. O. Yussuff
Department of Electronic & Computer Engineering, Lagos State University, Lagos, Nigeria
{funke_darley|adetokyom}@yahoo.com|abayomi.yussuff@lasu.edu.ng

Received: 17-FEB-2022; Reviewed: 15-APR-2022; Accepted: 07-MAY-2022
https://doi.org/10.46792/fuoyejet.v7i2.802

REVIEW ARTICLE

Abstract- Billions of IoT devices are in use worldwide and generate a humongous amount of data for the IoT system. This continuous stream of data is open to attack during its collection, transportation, processing, dissemination and storage cycle. Also, IoT devices themselves are points of system vulnerability through which the system can be attacked. Machine learning (ML), due to its ability to identify inherent patterns and behaviour in data, has been applied by many researchers to IoT data such that strange patterns or intrusions into IoT systems can be speedily detected and real-time decisions on security and privacy (S&P) protection implemented in a timely manner. Different ML techniques with their different algorithms have provided solutions in various scenarios such that security and privacy requirements for the IoT system can be met. In particular, ML has been successfully applied in intrusion detection and has been shown to perform better than traditional means in flagging new trends of attacks. This paper presents a systematic literature review on ML intrusion detection in IoT. Academic journals from 2011 to 2021 from two databases (IEEE and ProQuest) were explored using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. A review of the final selected papers revealed that data preprocessing, feature extraction, model training and deployment of ML-based Intrusion Detection Systems (IDS) increase computational complexity resulting in greater resource requirement (CPU, memory, and energy); enable ML to be used in the execution of adversarial attacks on IoT devices and networks (as seen with emerging attacks); give rise to scalability issues especially due to the heterogeneous nature of IoT networks; require trade-offs between detection accuracy and false-positive events; and highlight the superior performance of deep learning methods over traditional ML ones in anomaly detection. Generally, the changing nature of attacks makes it difficult for any particular IDS to be able to detect all attack types thus making the development of IDS a continuing project.

Keywords- Internet of Things, Intrusion Detection, Machine Learning, Security and Privacy, Systematic Review.

1 INTRODUCTION

The advent and growth of the Internet of Things (IoT) has simplified our lives as seen in its application in smart cities, smart health, smart homes, smart meters, smart supply chain and logistics, intelligent transportation, among others. However, this "ease of living", comes with associated challenges involving security risks and privacy concerns that must be addressed appropriately. This is due to the ubiquitous nature of IoT and its ability to impact many aspects of our lives positively or negatively.

Security and Privacy (S&P) issues in IoT can be presented based on IoT architecture which can be three-layered, four-layered (Zhang et al., 2017; (Rayes, A. and Salam, S., 2019), five-layered (Salman and Jain, 2017) or seven-layered (Salman and Jain, 2017) depending on the researcher's choice. Till date, there is no standard IoT architecture. In this study, the basic three-layered IoT architecture will be considered as shown in Figure 1.

IoT Application Layer: This layer consists of Cloud servers. These servers provide customized computational and storage services to both individuals and businesses.

IoT Network Layer: This is made up of different networks and devices. Networks include cellular networks, local area networks, and the Internet. Devices include hubs, routers, and gateways. All these are enabled by various communication technologies which include Bluetooth, LTE, Wi-Fi, and mobile networks.

IoT Perception Layer: This is made up of sensors. These collect data from the environment for onward transmission to the upper layers as well as actuators and controllers that monitor the process and take required action based on the outcome of the processed data.

The data collected from sensors are transferred to the application layer (Cloud) through the networks for further handling. Results/commands are then delivered to the end devices/actuators/users (Zhang and Tao, 2021; Lin et al, 2017). Whichever architecture is selected, data generated by IoT devices are vulnerable to attack during the data cycle – from the collection at the source, to transport from one layer to the other or between the elements of a layer, processing at the edge, fog or cloud (depending on architecture), data dissemination to user(s) and finally to storage. In addition, each IoT layer has its vulnerabilities and associated threats/attacks.

These attacks result in data breaches comprising unauthorized access to confidential and personal data that may be used for purposes of financial fraud, company data that may fall into the hands of a competing organization and threats to national security that may arise if vital government data falls into the hands of a hostile nation. For such reasons, there is a major need to ensure vulnerabilities and associated threats and attacks

*Corresponding Author

Section B- ELECTRICAL/ COMPUTER ENGINEERING & RELATED SCIENCES
Can be cited as:
Darley O.G., Adenowo A.A. and Yussuff. A.I.O. (2022): "Machine Learning Intrusion Detection as a Solution to Security and Privacy Issues in IoT: A Systematic Review", *FUOYE Journal of Engineering and Technology* (FUOYEJET), 7(2), 148-156. http://doi.org/10.46792/fuoyejet.v7i2.802

are identified and removed or minimized by deploying adequate protective solutions. This will remove users' fear of security breaches and privacy violations; inspire confidence; stimulate global acceptance and use and ultimately enable the commercialization of the IoT technology. An estimated value of about $163.2billion was expected in 2020 and a growth to about $493 billion was forecasted for the following five years (2021-2026) (Businesswire, 2021).

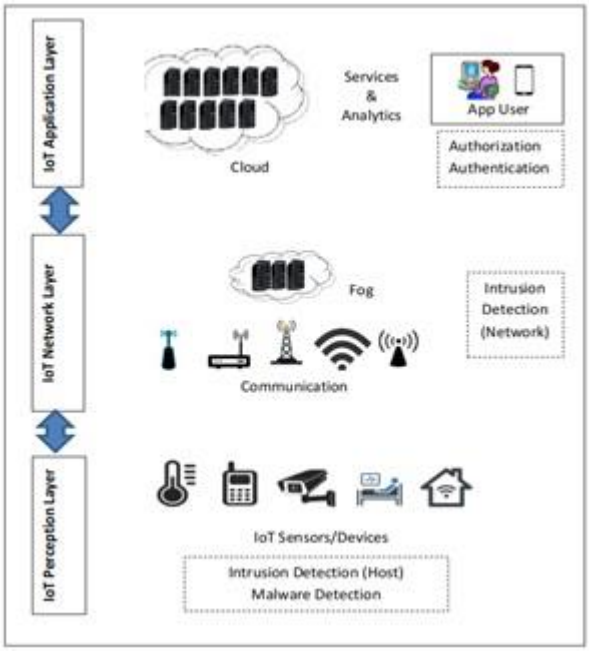

Fig. 1: Three-layered IoT Architecture with associated threats

The IoT system is particularly vulnerable at the device layer due to the millions of unsecured IoT devices already connected and with many more expected to join. Statista.com estimated that the sum of IoT-connected devices will be 30.9 billion units worldwide by 2025 (Statista, 2021). These devices are heterogeneous having different manufacturers, protocols, specifications, and command interfaces with no existing standard followed; thus opening up the entire IoT system to threats and attacks (Rani et al., 2021). While existing threats are easier to manage, emerging threats consist of different combinations of known threats and are therefore dynamic and more challenging. These threats and attacks leverage the huge data generated by the devices, the inadequate computing resources, and the limited storage capacity of IoT devices. The dynamic nature of threats calls for real-time responses of IoT devices to these threats. The time it takes for data to be transferred for computing in the Cloud, for threat to be identified and for meaningful information given such that appropriate mitigating actions can be deployed, determines how fast such threats can be shut down. Unfortunately, the route of device-cloud-device has high latency and inhibits the speed at which mitigating actions can be implemented. This challenge of high latency brought about the advent of Fog computing and Edge Computing (Bonomi, 2011; Dastjerdi et al, 2016; Zhang and Chiang, 2016; Ni et al, 2018) which extends the computing capabilities of the Cloud closer to the devices (at the edge) such that low latency can be achieved. Low latency ensures that real-time solutions to threats can be achieved and appropriate mitigating actions deployed if required.

In applying ML techniques for data protection or data preservation, some requirements (Rayes and Salam, 2019) have been determined to be necessary. These include Confidentiality, Integrity, Non-Repudiation, Authentication, Authorization, Availability and Freshness. Different applications call for different requirements to be met. One of the areas in which ML has been successfully applied to S&P is intrusion (anomaly, signature or hybrid) detection. ML can utilize the enormous amount of data generated by the IoT devices to identify inherent patterns and behaviours of the data (establish a norm in the system) and hence predict and detect vulnerabilities/threats/attacks (deviations from established norm/alien patterns and behaviour) in IoT-based systems; thereby flagging new trends of attacks (Hussain et al., 2020; Thamilarasu et al, 2020; Kasongo, 2021). Traditional means usually involve human intervention to set rules to which the system will live by to address threats that may arise. ML on the other hand, provides the means whereby the system does not need a predefined set of rules but will take decisions based on the training and learning received from the data that is generated by the IoT devices.

ML techniques can be classified under four categories – Supervised, Unsupervised, Semi-supervised and Reinforced Learning. Supervised learning utilizes input and output data that are labelled. The desired output is obtained by training an inferred model with input data. It includes two kinds - classification and regression (Kubat, 2021). Unsupervised learning utilizes unlabelled data and the machine classifies data by itself by detecting its characteristics of similarity. It includes two kinds – Clustering and Dimensionality Reduction. For semi-supervised learning, only some data are labelled. The computer will find features in the labelled data and apply to the unlabelled data for classification. The classification process from supervised learning will be used to identify data assets and the clustering process will be used to group them into unique parts. For reinforced learning, data is not labelled. The machine uses observations gathered from its interaction with the environment to adjust its classification based on the quality of feedback until it gets the correct result. A reward system is used whereby positive reward signifies that the performance of a particular sequence of actions be continued while negative reward penalizes for such actions. (Mohammed et al, 2016).

In this review, the use of ML techniques as a solution to security and privacy issues in IoT is presented with specific attention to intrusion detection; intrusion being a major challenge in IoT networks. Intrusion detection systems (IDS) can be device-based, network-based or hybrid (a combination of both). The network intrusion detection system (NIDS) utilizes network traffic in its analysis. On the other hand, the host-based intrusion

http://doi.org/10.46792/fuoyejet.v7i2.802 http://journal.engineering.fuoye.edu.ng/

detection (HIDS) utilizes the log data of sensors and devices (Meera et al, 2021; Singh and Singh 2014).

For anomalies (unknown-attacks), IDS seek to find patterns that are not consistent with the normal network traffic or sensor data patterns while for signatures (known-attacks), the patterns are compared with known patterns in the IDS database. Once an intrusion is flagged, required counter-measures will be deployed. Some examples include signature-based by (Eskandari, et al, 2020; Swarna et al, 2020), anomaly-based by (Ahmad et al, 2020; Eskandari et al., 2020; Sheikhan and Bostani, 2017) and for avoiding adversarial attacks on ML techniques by (Sagar et al, 2020; Miyato, et al, 2018; Zhang et al., 2019; Santana et al., 2021). The architecture of the IoT system – centralized or distributed is also considered.

Table 1. Some performance evaluation metrics for Machine Learning Models (Hameed et al, 2021)

Metrics	Description and Formula
Accuracy	Ratio of correctly predicted observations to total observations. It determines the performance of the model in recognizing all classes. $Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$
Specificity (Precision)	Ratio of correctly predicted positive observations to total predicted positive observations. It measures the exactness of the model. $Precision = \frac{TP}{TP + FP}$
Sensitivity (Recall/True Positive Rate (TPR))	Ratio of correctly predicted positive observation to all observations in actual class. It measures the completeness of model. $Sensitivity = \frac{TP}{TP + FN}$
False Positive Rate (FPR)	Measures the number of those normal network behaviours which are calculated as attacks. $FPR = \frac{FP}{FP + TN}$
F1-Score	Harmonic average of Precision and recall rates. $F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$
Area under curve (AUC)	True Positive Rate (TPR) is plotted against the False Positive Rate (FPR) of a given model. The area under the curve (AUC) always has a value between 0 and 1.
Mean Absolute Error (MAE)	Average of the difference between the original values and the predicted values. Gives a measure of how far the predictions were from the actual output. $MAE = \frac{1}{N}\sum_{j=1}^{N}\lvert y_j - \hat{y}_j \rvert$
Mean Squared Error (MSE)	Average of the square of the difference between the original values and the predicted values. Gives a measure of how far the predictions were from the actual output. $MSE = \frac{1}{N}\sum_{j=1}^{N}(y_j - \hat{y}_j)^2$

The performance of the ML technique applied determines its suitability for its intended purpose of intrusion detection. Some performance metrics are as listed in Table 1 while acronyms, description and meanings are listed in Table 2.

Table 2. Acronyms for performance metrics

Acronym	Meaning	Description
FN	False Negative	Network intrusions that are wrongly labelled as non-intrusive (normal).
FP	False Positive	Normal network traces that are labelled as intrusions.
TN	True Negative	Normal network traces that are correctly labelled as legitimate.
TP	True Positive	Intrusions that are correctly labelled as attacks.

This study is unique because it reviewed some existing ML-based IDS and identified new techniques (e.g., blockchain, multi-layer, and federated learning) that have been explored in combination with ML techniques in a bid to improve privacy preservation in IoT systems. The following sections are organized thus: Section 2 highlights related work, Section 3 presents the research methodology, Section 4 presents review findings, performance metrics used and identifies gaps/limitations in the various studies. Section 5 concludes with the main points of the study and suggests future work that can be carried out.

2 Related Work

Several surveys/reviews have been employed on utilizing ML to provide solutions to security and privacy concerns in IoT in general and intrusion detection in particular. Some are presented in this section.

Amiri-Zarandi et al. (2020) reviewed several studies that have applied machine learning (ML) to address privacy issues in IoT including scalability, interoperability, and resource (computation, storage and energy) limitations. Data generated in the different IoT layers were categorized and ML applied for robust privacy management. Chaabouni et al. (2019) reviewed traditional and ML-based NIDS by comparing IoT architectures selected, detection methodologies and validation strategies applied, threats/attacks identified and algorithms deployed as solutions. Results showed that the latter has an advantage of increased detection accuracy and decreased false positive alarms.

Hameed et al. (2021) focused on Implantable Medical Devices (IMDs) because of the serious impacts on the health and life of patients. For these devices, security solutions consisted of device anomaly detection, authentication, and access control as well as network layer security. Results showed that though traditional ML techniques were found to be effective, proper consideration must be given to resource capability, time complexity, and energy usage to ensure their effectiveness. Thamilarasu et al. (2020) also presented an IDS for IoT sensor and network data in connected medical devices. Results showed that the model had high detection accuracy with minimal computational resource required. (Iwendi et al., 2021) sought to detect network-based attacks and privacy violations in smart health care by using ML methods (Random Forest (RF)) and a feature optimization method) which resulted in a high detection

http://doi.org/10.46792/fuoyejet.v7i2.802 http://journal.engineering.fuoye.edu.ng/

rate with lower false alarm rate. Verma et al. (2021) also applied ML techniques as an effective solution for intrusion detection in Internet of Medical Things and privacy preservation of patients' medical records.

Liang et al. (2019) presented the various aspects (benefits, vulnerabilities and trends) of using ML for cybersecurity in "cyber-physical systems" (CPS) with CPS defined as "collections of physical and computer components that are integrated with each other to operate a process safely and efficiently" (Munirathinam, 2020). Examples of CPS are industrial control systems and smart grids and the benefits of applying ML techniques in to them include improved intrusion detection and decision accuracy. However, the growing trend of ML being used in execution of cyberattacks and intrusions was a major concern.

Skowron and Janicki (2020) presented a new method by which traffic fingerprinting attack can be detected. It also highlighted the vulnerability of ML being used adversarially against IoT devices and proposed countermeasures by which traffic analysis attacks can be mitigated were proposed. Sharma and Liu (2021) analysed six supervised learning algorithms (SVM, K-NN, RF, Naïve Bayes, Ensemble-V and Ensemble-B) for detecting misbehaviour in Internet of Vehicles (IoV). Despite the good performance obtained, it was seen that some attacks may escape detection as a result of the close similarity between the normal and abnormal data. Weng and Liu (2019) proposed anomaly detection for mobile service computing; the aim being to prevent hacking during data transfer to the Cloud. Wei et al. (2017), Lei et al.. (2019) and Zhu et al. (2020) on the other hand, sought to improve the efficiency and effectiveness of malware detection on Android phones by using ML techniques

Hassan et al. (2020), Shahid et al. (2020), Kasongo (2021) and Liu et al. (2020) presented various ML and deep learning (DL) models that are reliable in countering cyber-attacks and preserving the integrity of industrial internet of things (IIoT) networks. This is of particular importance as industries have become more defenceless against new threats/intrusions because of the nature of their networks. Nie et al. (2021) proposed a DL-based IDS for Social IoT. It was first implemented for a single attack and thereafter, several models were combined to handle multiple attacks. Simulation result showed that this method significantly improved the accuracy of intrusion detection. Liu and Lang (2019) presented a taxonomy of IDS that take data objects to classify and summarize ML- and DL-based IDS literature. The paper highlighted the advantage of DL over ML techniques and the constraint of the former requiring more computing resources. Asharf et al. (2020) and Chaabouni et al. (2019) presented comprehensive reviews on the use of ML in IDS and highlighted the various system architecture, protocols, detection methods, validation strategies, threats and algorithm utilizations as well as DL techniques considered.

Furthermore, this review has shown that in a bid to ensure privacy preservation, many researchers have veered from traditional ML techniques via centralized or edge learning methods, to newer ones such as federated learning (FL), blockchain, multi-layer/deep learning. FL is a privacy-preserving decentralized method. Raw data is kept on devices for local computation thereby denying hackers access to raw data. Many researchers (Nguyen et al. 2021; Dinesh et al, 2021; Ferrage and Friha, 2021; Pei et al, 2020; AbdulRahman et al, 2020) have used this technique in key applications such as smart cities, smart healthcare, smart industry, smart transportation and smart vehicles. Hassija et al. (2019) proposed using Blockchain, fog computing, edge computing, and ML to enhance IoT security. Alkadi et al. (2021) presented a Deep Blockchain Framework (DBF) which combined distributed intrusion detection and blockchain approaches to enable migration of data in a timely, reliable and secure manner. Lu et al. (2020) proposed the combination of blockchain and FL in order to enable secure and intelligent data sharing with high efficiency and utility. Chai et al. (2021) and Iftikhar et al. (2021) presented a hierarchical blockchain-enabled FL algorithm such that knowledge can be safely shared in Internet of Vehicles applications for traffic control, accident prevention and critical message sharing among vehicles. Ibrahim et al. (2022) presented a new blockchain-enabled protocol (BEP) and Software Defined Networking (SDN) architectures to boost IoT security against Denial of Service (DoS) and Distributed DoS (DDoS) attacks.

Multi-layer ML / DL made use of multiple layers of ML techniques (multiple deep networks) to boost detection accuracy and improve performance of IDS. (Lee et al, 2020) developed a lightweight ML-based IDS using deep auto-encoder and feature extraction. This enabled effective intrusion detection in the resource-constrained IoT devices. Khater et al. (2019) also proposed a lightweight IDS using a Multilayer Perceptron (MLP) model. These confirmed that with the right choice of techniques, intrusion detection can be carried out on resource constrained IoT devices. Qaddoura et al. (2021) combined two stages of detection to ensure better quality of classification results. Gassais et al. (2020), Bica et al. (2019) and Pajooh et al. (2021) implemented several ML and DL algorithms to achieve high detection capabilities, with little computational overhead on IoT devices (Hameed et al, 2021) proposed a combined (host and network) IDS which is used for the detection of malicious sensor and network data simultaneously. A summary of some of the studies reviewed are presented in Table 3.

3 RESEARCH METHODOLOGY

In this study, the focus was on the use of ML techniques in intrusion detection; intrusion being a major challenge in IoT networks. The study undertook a systematic literature review (SLR) utilizing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Azevedo et al, 2017; Guelph-Humber, 2021) for the review process. PRISMA is a recognized standard and demonstrates the procedure and details of document identification, screening,

http://doi.org/10.46792/fuoyejet.v7i2.802 http://journal.engineering.fuoye.edu.ng/

inclusion/exclusion such that the accuracy of reviews and meta-analysis reports can be improved and review methods can be replicated, if necessary. Two databases, the IEEE and ProQuest (comprising 4 databases) were searched for relevant papers. The following keywords were used: Internet of Things, IoT, Intrusion Detection, Machine Learning, Security and Privacy and Systematic Review. Papers listed were thereafter screened using specific inclusion/exclusion criteria. Eligible papers were then further scrutinized based on title, content of abstract, etc., and reviewed.

4 RESULTS AND DISCUSSIONS

Using PRISMA with relevant key words in an advanced search, 6,128 papers were obtained from both the ProQuest and IEEE databases. Advance Search was implemented as follows: Internet of Things AND IoT AND Security and Privacy AND Machine learning AND Intrusion Detection with the search process presented in Fig. 2. Several filters such as type of publication (journal, conferences, or book chapter), date of publication, document type (article, literature review, review, etc.) were applied. With inclusion/exclusion criteria also applied, the number of potential papers for review was reduced to 99 for ProQuest and 55 for IEEE databases. Further selection was made based on abstracts' contents and relevance. The final selection of papers was reduced to 60; 30 papers from each database, for the purpose of equity.

Table 3. Summary of some related works

Paper Details	Contributions	Limitations/Gaps
(Chaabouni et al. 2019)	Compared current defense techniques in traditional and ML NIDS in terms of architecture, detection methodologies, validation strategies, treated threats, and algorithm deployments; with the latter providing increased detection accuracy and decreased false positive alarms.	Standard public benchmark dataset is required for improved validation strategy such that a clear, practical and convenient comparison of the different NIDS can be carried out.
(Hassija et al. 2019)	Identified threats at the various layers of IoT architecture and proposed solutions using blockchain, fog computing, edge computing, and machine learning; the aim being to enhance the level of security of the system.	Blockchain: Leakage of private information of users. Scalability and availability issues as number of miners increase. Fog & Edge Computing: data security and user privacy due to leakage and misuse of a user's private data. ML: Effectiveness and accuracy of IDS is dependent on choice of algorithm and dataset.
(H. Liu and Lang 2019)	Proposed a taxonomy of IDS that takes data objects as the main dimension to classify and summarize machine learning-based and deep learning-based IDS literature.	Lack of available datasets, low detection accuracy in actual environments, low efficiency due to complicated models and extensive data preprocessing methods required between effectiveness and efficiency in order for the IDSs to detect attacks in real time.
(Liang, et al, 2019)	Proposed mechanisms to enhance IDS accuracy in "cyber-physical systems (CPS)" using various ML algorithms.	Ability of attackers to use ML techniques in the execution of cyberattacks and intrusions and vulnerabilities at all stages of the data life-cycle.
(AbdulRahman et al., 2020)	Presented how ML techniques can be utilized for S&P issues and resource management in IoT using FL.	Being a new research direction, more studies need to be carried out to develop more robust FL systems.
(Amiri-Zarandi et al, 2020)	Identified merits and demerits of utilizing data in ML-based solutions for privacy in IoT.	Identified limitations include lack of standard data practices and policies, interoperability of devices and regulatory compliance. Need for new ideas such as using Blockchain and ML techniques to be further investigated.
(Asharf et al. 2020)	Presented various aspects of IoT systems in terms of architecture, protocols, technologies, and emerging threats from compromised IoT devices. An overview of intrusion detection models using ML and DL techniques for attack detection was also presented.	Lack of suitable datasets, assumption of real-time IDS that there is no attack traffic during the learning phase which sometimes leads to false alarms, resource limitation of IoT devices, heterogeneity of IoT system and scalability issues.
(Hussain et al. 2020)	Identified requirements for IoT network security, attack types and current solutions proposed using existing ML and DL solutions.	Shortcomings in MD/DL techniques such as computational complexity, learning efficiencies and parameter tuning strategies.
(Leevy and Khoshgoftaar 2020)	Presented and analyzed IDS based on the CSE CICIDS2018 dataset; this being the most recent intrusion detection dataset that is big data, publicly available and covers a wide range of attack types.	Unusually high performance scores which may be a consequence of overfitting, class imbalance of dataset and lack of the data cleaning to enhance dataset usability.
(Hameed, et al, 2021)	Provided security solutions (sensor anomaly detection, device authentication, access control and network layer security) to Implantable Medical Devices (IMDs) by applying ML techniques.	Effectiveness is determined by resource capacity, time complexity and energy usage which must be given proper consideration.
(Nguyen et al. 2021)	Used federated learning (FL) in IoT networks to enhance privacy preservation.	Resource requirements are not always met.

http://doi.org/10.46792/fuoyejet.v7i2.802 http://journal.engineering.fuoye.edu.ng/

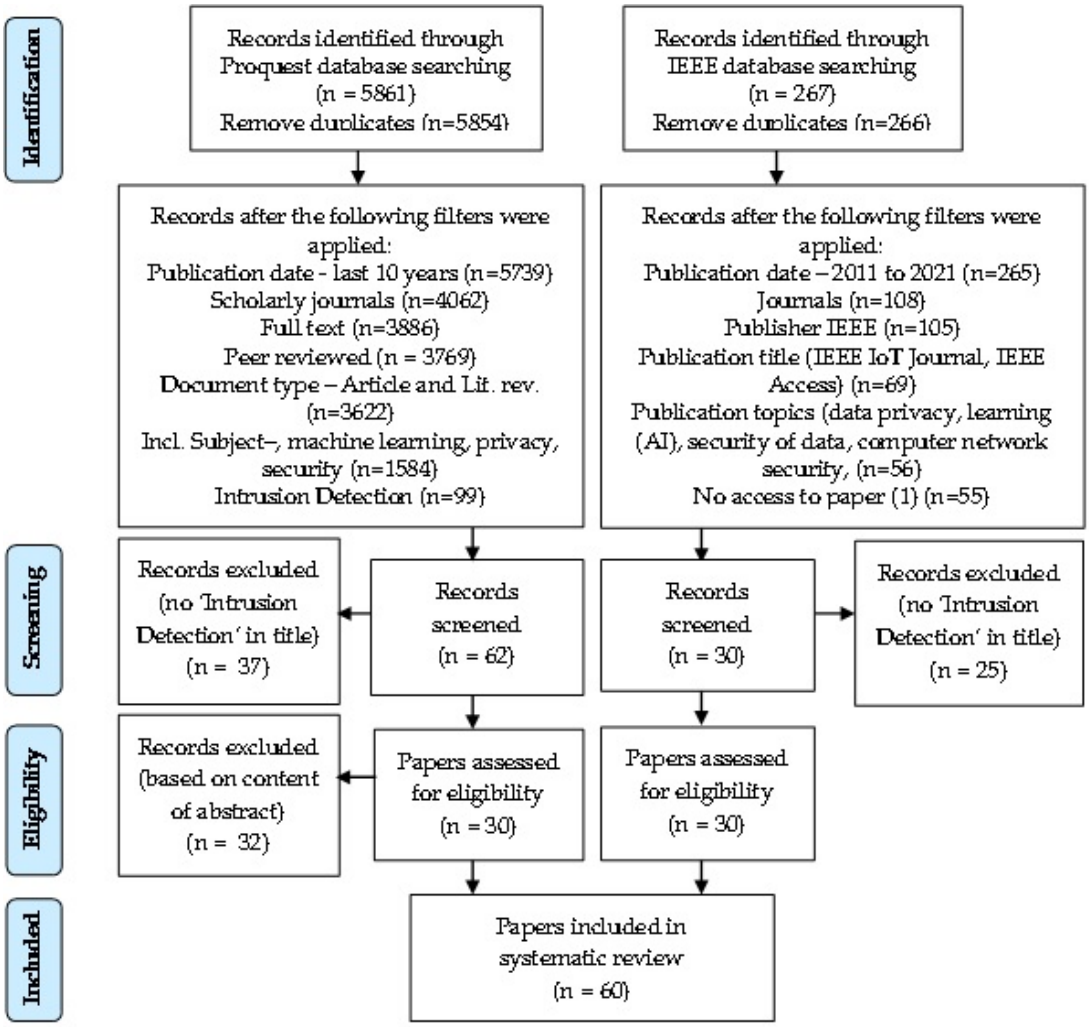

Fig. 2: PRISMA Flow Diagram
(Based on Advance Search: Internet of Things AND IoT AND Security and Privacy AND Machine learning).

Table 4. Summary of some studies on using ML/DL Techniques in IDS

Paper	DR and FS*	ML/DL Classifiers	Dataset	Metrics			Limitations / Gaps
Liu et al, 2020	DSSTE	RF SVM XGBoost LSTM Mini VGGNet AlexNet	NSL-KDD CSE-CIC-IDS2018		Accuracy	F1-Score	The preprocessed dataset used is suitable for ML but not so for DL which performs better on original network traffic data.
				DSSTE +RF	96.92%	96.98%	
				DSSTE +SVM	94.88%	94.63%	
				DSSTE +XGBoost:	96.02%	96.11%	
				DSSTE +LSTM	96.38%	96.50%	
				DSSTE +Mini VGGNet	96.99%	97.04%	
				DSSTE +AlexNet	96.53%	96.49%	
(Mezina et al, 2021)	Focal loss function (with temporal CNN with LSTM)	CNN, RNN, Autoencoder Fully connected network	KDD99 CSE-CIC-IDS2018		KDD 99	CIC-IDS2018	Overfitting of models trained on KDD99 dataset, poorer performance when used on ICIDS2018.
				U-net Accuracy	93.03%	94.65%	
				Temporal CNN with LSTM Accuracy	92.05%	97.77%	
(Aleesa et al, 2021)	Nil	LSTM (RNN) ANN DNN	UNSW-NB15	Accuracy	Binary	Multiclass	Limited hardware capability which limited the number of hidden layers and neurons that could be used.
				RNN-LSTM	85.42%	85.38 %	
				ANN	99.26 %	97.89 %	
				DNN	99.22 %	99.59 %	
(Zhou et al, 2021)	Autoencoder Neural Network	Variational LSTM	UNSW-NB15	Precision (86%), Recall (97.8%), F1-Score (90.7%), AUC (0.895)			Use of other DL methods for improved performance.
(Manimurugan et al, 2021)	Nil	Deep Belief Network (DBN)	CIC-IDS2017	Accuracy (%) : Normal Class (99.37), Botnet (97.93), Brute Force (97.71), DoS/DDoS (96.67), Infiltration (96.37), PortScan (97.71),Web attack (98.37)			More recent database to be used so that IDS can detect newer attacks.
(Maithem and Al-sultany, 2021)	Z-Score normalization for numerical values One hot encoder for text values	Multi-Layer Perceptron (MLP)	KDD Cup1999		Binary:	Multiclass	Only 4 attack types (DoS, R2L, U2R and Probe) are categorized ignoring infiltration and web attacks; Overfitting of model
				Accuracy	99.98%	99.98%	
				Precision	99.99%	%	
				Recall	96.3%	%	
				F1-Score	79.7%	%	
(Wu et al, 2020)	Nil	DNN (Combination of CNN and RNN)	NSL-KDD UNSW-NB15	Accuracy: NSL-KDD (99.21%), UNSW-NB15 (86.64%)			Requires more experiments to improve performance
(Gao et al, 2019)	Adaptive Principal Component (APAC)	Incremental Extreme Learning Machine (IELM)	NSL-KDD UNSW-NB15		NSL-KDD	UNSW-NB15	More research required to adapt to industrial control systems (ICS).
				Accuracy	81.22%	70.51%	
(Vinayakumar, R., Alazab, Soman, K., Poornachandran, & Al-nemrat, A. and Venkatraman, S., 2019)	Nil	DNN – Binary Modelling DNN – Multiclass Modelling	UNSW-NB15		Binary	Multiclass	Execution time to be reduced; more complex DNN not trained for perfor-mance enhancement due to high computational costs.
				Accuracy	76.1%	65.1%	
				Precision	95.1%	59.7%	
				Recall	96.3%	65.1%	
				F1-Score	79.7%	75.6%	
(Hanifa et al, 2019)	Nil	ANN	UNSW-NB15	Precision (84%)			No feature selection.

Note: *DR and FS - Dimensionality Reduction and Feature Selection

http://doi.org/10.46792/fuoyejet.v7i2.802 http://journal.engineering.fuoye.edu.ng/

Overall, the purpose of intrusion detection in IoT is to implement an IDS that is effective, reliable and accurate in detecting and preventing various attacks (both anomalous and signature) against the system. In this quest, various ML and DL methods have been used – in singular form or in hybrid form (combination of two or more techniques). Efforts have also been made to ensure the dataset is suitable and usable for learning. Various data transformation techniques and feature selection/extraction methods have been utilized with the aim of balancing the dataset and reducing features, respectively. The type of a dataset (balanced or imbalanced; old or current) and the type of feature selection/extraction method employed play key roles in the performance of IDS (Kurniabudi, et al, 2020; Adekunle et al, 2019; Ayogu et al, 2019) as they impact the ability of the IDS to accurately detect attacks. This is known as the data preprocessing stage and the new dataset obtained will be used to train the IDS by using ML/DL classifiers. Despite the numerous studies that have been carried out in this regard, many gaps still exist with each study having its own peculiarities. Results of a few of the papers reviewed are presented in Table 4 with gaps pertaining to each study highlighted.

From this review, some gaps and limitations of the various methods used in the design and development of IDS have been identified. These include:

- The choice of methods of dimensionality reduction and feature selection in the preprocessing stage of IDS development impact the capability of the IDS to accurately detect attacks.
- Data preprocessing and feature reduction, model training and deployment of ML- and DL-based NIDS, increase computational complexity. Increase in resource requirement (CPU, storage, energy) in implementing ML/DL techniques is a major challenge for resource-constrained IoT devices in particular and hinders the provision of adequate protection for them. This highlights the need for an efficient NIDS with light computational requirements.
- The need for a trade-off between detection accuracy and false positive events such that mitigating actions are not unnecessarily deployed which may lead to disruption of services.
- The use of ML in the execution of adversarial attacks on IoT devices and networks as seen with emerging attacks.
- Inability of IDS to detect some attacks especially those hidden in network datasets due to the imbalanced nature of the datasets. This is because normal data are much larger than attack data thus resulting in bias towards the larger dataset; leading to high false-positive rates.
- Inability of IDS to capture all possible normal observations in the network, particularly in IoT network where data is sent from heterogeneous devices, resulting in high false-negative rates.
- Use of ML/DL techniques in IoT networks (which are large and distributed) raises scalability issues especially due to the diverse (heterogeneous) nature of IoT networks.
- DL models are more efficient in attack detection than ML but have the disadvantage of longer running times making them sometimes unsuitable to meet the real-time requirement of IDS.

The dynamic nature of attacks makes it difficult for any particular IDS to detect all types of attacks. New and emerging attacks come in various combinations of existing attacks, keep evolving and increase in complexity. Thus, the development of IDS is a continuing project.

5 CONCLUSIONS

The role of ML in security and privacy issues in IoT has been investigated by many researchers with ML techniques significantly utilized for security in both the device and network layers. In this study, the systematic review of various ML techniques applied for intrusion detection was carried out. This review examined relevant studies that were published in the last ten years (2011-2021) from two databases (IEEE and ProQuest). From the final set of papers (60) selected, it was observed that DL methods were better suited than traditional ML methods for anomaly detection in IoT. This was based on the various performance metrics utilized in the studies. To further enhance IoT system performance, it was observed that researchers explored combinations of various techniques such as federated learning, blockchain and multi-layer classification approaches. Models that combined various ML techniques were also observed to present improvement in performance. Despite the success of these efforts, some gaps as listed above exist. Future work should explore further enhancement of system performance and efficiency by improving dataset balancing and feature selection methods. Efficient IDS will contribute to the larger picture of acceptability and commercialization of IoT technology which is estimated to reach $493 billion by 2026.

REFERENCES

AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C. and Guizani, M. (2020). A Survey on Federated Learning: The Journey from Centralized to Distributed On-Site Learning and Beyond. *IEEE Internet of Things Journal*, *0*(0). https://doi.org/10.1109/JIOT.2020.3030072

Adekunle, B. O., Akinyemi, J. B., Aladesanmi, T. A., Aderounmu, A. G. and Kamagate, B. H. (2019). An Improved Anomalous Intrusion Detection Model. *FUOYE Journal of Engineering and Technology*, *4*(2). https://doi.org/10.46792/fuoyejet.v4i2.418

Ahmad, I., Yousaf, M., Yousaf, S. and Ahmad, M. O. (2020). Research Article Fake News Detection Using Machine Learning Ensemble Methods. *Wiley Hindawi*. Retrieved from https://doi.org/10.1155/2020/8885861

Alkadi, O., Moustafa, N., Turnbull, B., & Choo, K. K. R. (2021). A Deep Blockchain Framework-Enabled Collaborative Intrusion Detection for Protecting IoT and Cloud Networks. *IEEE Internet of Things Journal*, *8*(12), 9463–9472. https://doi.org/10.1109/JIOT.2020.2996590

Amiri-Zarandi, M., Dara, R. A. and Fraser, E. (2020). A survey of machine learning-based solutions to protect privacy in the

Internet of Things. *Elsevier Computers & Security*. https://doi.org/10.1016/j.cose.2020.101921

Ayogu, B. A., Adetunmbi, A. O. and Ayogu, I. I. (2019). A Comparative Analysis of Decision Tree and Bayesian Model for Network Intrusion Detection System. *FUOYE Journal of Engineering and Technology*, *4*(2). https://doi.org/10.46792/fuoyejet.v4i2.362

Bica, I., Chifor, B. S., Arseni, S. and Matei, I. (2019). Multi-Layer IoT Security Framework for Ambient Intelligence Environments. *MDPI Sensors*, *19*. https://doi.org/10.3390/s19184038

Businesswire. (2021). Global IoT Connectivity Market Analysis and Forecast Report 2021. In *Businesswire*. Retrieved from https://www.businesswire.com/news/home/20211015005387/en/Global-IoT-connectivity-Market-Analysis-and -Forecast-Report-2021/

Chaabouni, N., Mosbah, M., Zemmari, A., & Sauvignac, C. and Faruki, P. (2019). Network Intrusion Detection for IoT Security Based on Learning Techniques. *IEEE - Communications Surveys and Tutorials*, *21*(3), 2671–2701.

Chai, H., Leng, S., Chen, Y., & Zhang, K. (2021). A Hierarchical Blockchain-Enabled Federated Learning Algorithm for Knowledge Sharing in Internet of Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 3975–3986. https://doi.org/10.1109/TITS.2020.3002712

Dastjerdi, A. V., Gupta, H., Calheiros, R., Ghosh, S.K. and Buyya, R. (2016). *Fog Computing: Principles, Architectures, and Applications*. https://doi.org/10.1016/B978-0-12-805395-9.00004-6

Eskandari, M. M., Janjua, Z. H., Vecchio, M. and Antonell, F. (2020). An Intelligent Anomaly Based Intrusion Detection System for IoT Edge Devices M. *IEEE Internet of Things Journal*. https://doi.org/OI 10.1109/JIOT.2020.2970501,

Ferrage, M. A., & Friha, O. (2021). Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis. *IEEE Access*, *9*(MI), 138509–138542. https://doi.org/10.1109/ACCESS.2021.3118642

Gassais, R., Ezzati-jivan, N., Fernandez, J. M., Aloise, D. and Dagenais, M. R. (2020). Multi-level host-based intrusion detection system for Internet of things. *Journal of Cloud Computing*, *9*(1). https://doi.org/10.1186/s13677-020-00206-6

Guedes, A.L.A., Alvarenga, J.C., Goulart, M.S.S., Rodriguez, M.V.R. and Pereira Soares, C. A. P. (2018). Smart cities: The main drivers for increasing the intelligence of cities. *MDPI - Sustainability*, *10*(9). https://doi.org/10.3390/su10093121

Hameed, S. S., Hassan, W. H., Abdul Latiff, L. and Ghabban, F. (2021). A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. *PeerJ Comput. Sci.*, *7*. https://doi.org/10.7717/peerj-cs.414

Hameed, S. S., Selamat, A., Latiff, L. A., Razak, S.A., Krejcar, O., Fujita, H., Sharif, A. and Omatu, S. (2021). A Hybrid Lightweight System for Early Attack Detection in the IoMT Fog. *MDPI - Sensors*, *21*(24). https://doi.org/10.3390/s21248289

Hassan, M. M., Gumaei, A., Huda, S. and Almogren, A. (2020). Increasing the Trustworthiness in the Industrial IoT Networks Through a Reliable Cyber-Attack Detection Model. *IEEE Transactions on Industrial Informatics*, *16*(9), 6154–6162. https://doi.org/10.1109/TII.2020.2970074

Hassija, V., Chamola, V., Saxena, V. and Jain, D. (2019). A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Access*, *7*, 82721–82743. https://doi.org/10.1109/ACCESS.2019.2924045

Hongyu Liu, H. and Lang, B. (2019). Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. *Appl. Sci.*, *9*. https://doi.org/10.3390/app9204396

Huijuan Zhu, Yang Li, Ruidong Li, Jianqiang Li, Zhuhong You, and H. S. (2021). Empowering Things with Intelligence: A Survey of the Progress, Challenges, and Opportunities in Artificial Intelligence of Things. *IEEE Internet of Things Journal*, *8*(10), 7789–7817.

Hussain, F., Hussain, R., Hassan, S. A. and Hossain, E. (2020). Machine Learning in IoT Security: Current Solutions and Future Challenges. *IEEE - Communications Surveys and Tutorials*, *22*(3), 1686–1721. https://doi.org/10.1109/COMST.2020.2986444

Ibrahim, M., Hanif, M. Ahmad, S., Jamil, F., Sehar, T., Lee, Y. and Kim, D. (2022). SDN Based DDos Mitigating Approach Using Traffic Entropy for IoT Network. *ResearchGate - Computers, Materials and Continua*. https://doi.org/10.32604/cmc.2022.017772

Iftikhar, Z., Javed, Y., Zaidi, S. Y. A., Shah, M. A., Khan, Z. I., Mussadiq, S. and Abbasi, K. (2021). Privacy Preservation in Resource-Constrained IoT Devices Survey Using Blockchain - A Survey. *MDPI- Electronics*, *10*(14). https://doi.org/10.3390/electronics10141732

Iwendi, C., Anajemba, J. H., Biamba, C. and Ngabo, D. (2021). Security of Things Intrusion Detection System for Smart Healthcare. *MDPI- Electronics*, *10*(12), 1–28. https://doi.org/10.3390/electronics10121375

J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Z. (2017). A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications. *IEEE Internet Things Journal*, *4*(5), 1125–1142.

Kasongo, S. M. (2021). An Advanced Intrusion Detection System for IIoT Based on GA and Tree Based Algorithms. *IEEE Access*, *9*, 113199–113212. https://doi.org/10.1109/ACCESS.2021.3104113

Khater, B. S., AbdulWahab, A. W., Idris, M. Y. I., Hussain, M. A. and Ibrahim, A. A. (2019). A Lightweight Perceptron-Based Intrusion Detection System for Fog Computing. *MDPI - Applied Sciences*, *9*. https://doi.org/10.3390/app9010178

Kubat, M. (2021). *An Introduction to Machine Learning* (3rd ed.). Springer.

Kurniabudi, Stiawan, D., Darmawijoyo, Bin Idris, M. Y., Bamhdi, A. M. and Budiarto, R. (2020). CICIDS-2017 Dataset Feature Analysis with Information Gain for Anomaly Detection. *IEEE Access*, *8*, 132911–132921. Retrieved from 10.1109/ACCESS.2020.3009843

Lee, S. J., Yoo, P. D., Asyhari, A. T., Jhi, Y., Chermak, L, Yeu, C. Y. and Taha, K. (2020). IMPACT: Impersonation Attack Detection via Edge Computing Using Deep Autoencoder and Feature Abstraction. *IEEE Access*, *8*, 65520–65529. https://doi.org/10.1109/ACCESS.2020.2985089

Lei, T., Qin, Z., Wang, Z., Li, Q. and Ye, D. (2019). EveDroid: Event-Aware Android Malware Detection Against Model Degrading for IoT Devices. *IEEE Internet of Things Journal*, *6*(4), 6668–6680.

Liang, F. W., Hatcher, W. G., Liao, W., Gao, W. and Yu, W. (2019). Machine Learning for Security and the Internet of Things: The Good, the Bad, and the Ugly. *IEEE Access Special Section on Security and Privacy in Emerging Decentralized Communication Environments*, *7*, 158126–158147. https://doi.org/10.1109/ACCESS.2019.2948912

Lin, J., Yu, W., Zhang, N., Yang, X. and Zhao, W. (2017). A survey on Internet of Things: architecture, enabling technologies, security and privacy and applications. *IEEE Internet Things Journal*, *4*(5), 1125–1142.

Liu, Y., Garg, S., Nie, J., Zhang, Y., Xiong, Z., Kang, J. and Hossain, M. S. (2020). Deep Anomaly Detection for Time-series Data in Industrial IoT: A Communication-Efficient On-device Federated Learning Approach. *IEEE Internet Things Journal*, *8*(8), 6348–6358. https://doi.org/10.1109/JIOT.2020.3011726

Lu, Y., Huang, X., Dai, Y., Maharjan, S. and Zhang, Y. (2020). Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT. *IEEE Transactions on Industrial Informatics*, *16*(6), 4177–4186. https://doi.org/10.1109/TII.2019.2942190

Ma, L., Pei, Q., Zhou, L., Zhu, H., Wang, L. and Ji, Y. (2020). Federated Data Cleaning: Collaborative and Privacy-Preserving

Data Cleaning for Edge Intelligence. *IEEE Internet of Things Journal*, *4662*(c), 1–15. https://doi.org/10.1109/JIOT.2020.3027980

Meera, A. J., Kantipudi, M.V.V.P. and Aluvalu, R. (2021). Intrusion Detection System for the IoT: A Comprehensive Review. *Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR). Advances in Intelligent Systems and Computing.*, *1182*, 235–243. https://doi.org/doi:10.1007/978-3-030-49345-5_25

Miyato, T., Maeda, S., Koyama, M. and Ishii, S. (2018). Virtual Adversarial Training: A Regularization Method for Supervised and Semi-Supervised Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/DOI: 10.1109/TPAMI.2018.2858821

Mohammed, M., Khan, M. B. and Bashier, E. B. M. (2016). Machine Learning: Algorithms and Applications. In *Boca Raton, FL, USA: CRC Press*.

Munirathinam, S. (2020). Chapter Six - Industry 4.0: Industrial Internet of Things (IIoT). *Advances in Computers*, *117*(1), 129–164. https://doi.org/10.101s.adcom.2019.10.0106

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J. and Poor, H. V. (2021). Federated Learning for Internet of Things: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, *23*(3), 1622–1658. https://doi.org/10.1109/COMST.2021.3075439

Ni, J., Zhang, K., Lin, X. and Shen, X. S. (2018). Securing fog computing for Internet of Things applications: Challenges and solutions. *IEEE - Communications Surveys and Tutorials*, *20*(1), 601–628.

Nie, L., Wu, Y., Wang, X., Guo, L., Wang, G., Gao, X. and Li, S. (2021). Intrusion Detection for Secure Social Internet of Things Based on Collaborative Edge Computing: A Generative Adversarial Network-Based Approach. *IEEE Transactions on Computational Social Systems*, *9*(1), 134–145. https://doi.org/10.1109/TCSS.2021.3063538

Pajooh, H. H., Rashid, M. and Alam, F. (2021). Multi-Layer Blockchain-Based Security Architecture for. *MDPI - Sensors*, *21*(3), 1–27. https://doi.org/10.3390/s21030772

Qaddoura, R., Al-zoubi, A. M., & Faris, H., Almomani, I. (2021). A Multi-Layer Classification Approach for Intrusion Detection in IoT Networks Based on Deep Learning. *MDPI - Sensors*, *21*(9). https://doi.org/10.3390/s21092987

Rani, S., Kataria, A., Sharma, V., Ghosh, S., Karar, V., Lee, K. and Choi, C. (2021). Threats and Corrective Measures for IoT Security with Observance of Cybercrime: A Survey. *Hindawi - Wireless Communications and Mobile Computing*, *2021*. https://doi.org/10.1155/2021/5579148

Rayes, A. and Salam, S. (2019). *Internet of Things from Hype to Reality: The Road to Digitization* (2nd ed.). https://doi.org/10.1007/978-3-319-99516-8

Rayes, Ammar and Salam, S. (2019). *Internet of Things from Hype to Reality the Road to Digitization* (Second). Retrieved from https://doi.org/10.1007/978-3-319-99516-8

Research Guides: Systematic Reviews: PRISMA Diagram & Checklist. (n.d.). Retrieved January 20, 2011, from Guelph-Humber Library Services website: https://guelphhumber.libguides.com/c.php?g=213266&p=1406923

Sagar, R., Jhaveri, R. and Borrego, C. (2020). Applications in Security and Evasions in Machine Learning: A Survey. *MDPI- Electronics*, *9*(1). https://doi.org/10.3390/electronics9010097

Salman, T. and Jain, R. (2017). Chapter 13: Networking Protocols and Standards for Internet of Things. In *Internet of Things and Data Analytics Handbook*. https://doi.org/10.1002/9781119173601.ch13

Santana, E.J., Silva, R.P., Zarpelão, B.B. and Barbon, S. (2021). Detecting and Mitigating Adversarial Examples in Regression Tasks: A Photovoltaic Power Generation Forecasting Case Study. *MDPI - Information*, *12*(10). https://doi.org/10.3390/info12100394

Shahid, L., Zhuo, Z., Zeba, I. and Jawad, A. (2020). A Novel Attack Detection Scheme for the Industrial Internet of Things using a Lightweight Random Neural Network. *IEEE Access*, *4*, 1–14. https://doi.org/10.1109/ACCESS.2020.2994079

Sharma, P. and Liu, H. (2021). A Machine-Learning-Based Data-Centric Misbehavior Detection Model for Internet of Vehicles. *IEEE Internet of Things Journal*, *8*(6), 4991–4999.

Sheikhan, M. and Bostani, H. (2017). A Security Mechanism for Detecting Intrusions in Internet of Things Using Selected Features Based on MI-BGSA. *International Journal of Information and Communication Technology Research, IJICTR*, *9*(2). Retrieved from https://www.researchgate.net/publication/320691235

Singh, A.P and Singh, M. D. (2014). Analysis of Host-Based and Network-Based Intrusion Detection System. *International Journal of Computer Network and Information Security*, *6*(8), 41–47. https://doi.org/10.5815/ijcnis.2014.08.06

Skowron, M., Janicki, A. and Mazurczyk, W. (2020). Traffic Fingerprinting Attacks on Internet of Things Using Machine Learning. *IEEE Access Special Section on Data Mining for Internet of Things*, *8*, 20386–20400.

Statista. (2021). IoT number of connected devices worldwide. *Statista*.

Swarna P.R.M., Praveen, K.R.M., Parimala, M., Srinivas, K., & Thippa, R.G., Chiranji L.C. and Mamoun, A. (2020). An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. *Computer Communications*, *160*, 139–149. https://doi.org/http://dx.doi.org/10.1016/j.comcom.2020.05.048

Thamilarasu, G., Odesile, A. and Hoang, A. (2020). An Intrusion Detection System for Internet of Medical Things. *IEEE Access*, *8*, 181560–181576. https://doi.org/10.1109/ACCESS.2020.3026260

Verma, G., Pathak, N. and Sharma, N. (2021). A Secure Framework for Health Record Management Using Blockchain in Cloud Environment. *Journal of Physics Conference Series*, *1998*(1). https://doi.org/10.1088/1742-6596/1998/1/012019

Wei, L., Luo, W., Weng, J., Zhong, Y., Zhang, X. and Yan, Z. (2017). Machine Learning-Based Malicious Application Detection of Android. *IEEE Access Special Section on Internet-of-Things (IOT) Big Data Trust Management*, *5*, 25591–25601.

Weng, Y. and Liu, L. (2019). A collective Anomaly Detection Approach for Multidimensional Streams in Mobile Service Security. *IEEE Access Special Section on Mobile Service Computing with Internet of Things*, *7*, 49157–49168. https://doi.org/10.1109/ACCESS.2019.2909750

Zhang, H., Chen, H., Song, Z., Boning, D., Dhilon, I. and Hsieh, C. (2019). The Limitations of Adversarial Training and the Blind-Spot Attack. *International Conference on Learning Representation (ICLR)*. Retrieved from https://arxiv.org/abs/1901.04684

Zhang, M. and Chiang, T. (2016). Fog and IoT: An overview of research opportunities. *IEEE Internet Things Journal*, *3*(6), 854–864.

Zhu, H., Li, Y., Li, R., Li, J., You, Z. and Song, H. (2020). SEDMDroid: An enhanced stacking ensemble framework for Android malware detection. *IEEE Transactions on Network Science and Engineering*, *4697*(c), 1–12. https://doi.org/10.1109/TNSE.2020.2996379