Using Machine Learning for Security Issues in Cognitive IoT

Wafa Alzuabi
College of Information Technology
University of Bahrain
Sakhir, Bahrain
202100005@stu.uob.edu.bh

Wael Elmedany
College of Information Technology
University of Bahrain
Sakhir, Bahrain
welmedany@uob.edu.bh

Mhd Saeed Sharif
Intelligent Technologies Research Group, CDT,
UEL, London, E16 2RD, UK.
s.sharif@uel.ac.uk

Abstract*—Cognitive learning is progressively prospering in the field of Internet of Things (IoT). With the advancement in IoT, data generation rate has also increased, whereas issues like performance, attacks on the data, security of the data, and inadequate data resources are yet to be resolved. Recent studies are mostly focusing on the security of the data which can be handled by machine learning. Security and privacy of devices intrusion detection their success in achieving classification accuracy, machine deep learning with intrusion detection systems have greatly increased popularity. However, the need to store communication centralized server compromise privacy and security. Contrast, Federated Learning (FL) fits appropriately as a privacy-preserving decentralized learning technique that trains locally transfer the parameters the centralized instead of purpose current research provide thorough and application FL intrusion detection systems. Machine Learning (ML) and Deep Learning (DL) approaches, which may embed intelligence in IoT devices and networks, can help to overcome a variety of security challenges. The research includes a detailed overview of the application of FL in several anomaly detection domains. In addition, it increases understanding of ML and its application to the field of the Cognitive Internet of Things (CIoT). This endeavour also includes something crucial . The relevant FL implementation issues are also noted, revealing potential areas for further research. The researcher emphasised the flaws in current security remedies, which call for ML and DL methods. The report goes into great detail on how ML and DL are now being utilised to help handle various security issues that IoT networks are facing. Random Neural Networks that have been trained using data retrieved by Cognitive Packets make the routing decisions. A number of potential future directions for ML and DL-based IoT security research are also included in the study. The report concludes by outlining workable responses to the problem. The paper closes by offering a beginning point for future study, describing workable answers to the problem of FL-based intrusion detection system implementation.

***Index Terms*—Machine Learning, Federated Learning, Intrusion detection,system.**

I. Introduction

A huge network of physically connected objects that communicate with other online systems and devices is known as the ”internet of things” (IoT). If its full potential can be realised, the IoT is poised to become one of the most important technical advancements of our time. The IoT is ”a worldwide infrastructure that uses cutting-edge services to connect (physical and virtual) things based on existing and developing interoperable information and communication technologies” [1]. Even though it pertains to actual items, The term ”Internet of Things” is used to describe a highly dispersed network that combines connection with sensors and lightweight applications that are embedded into tools and objects [2]. These exchange data with a variety of devices, including connected electricity grids, smart plugs, connected medical equipment, and linked automobiles. However, due to the rise of the IoT system and rapid growth in the rise of the network, the security challenges for IoT have also increased. IoT security is a broad phrase that describes the methods, apparatuses, frameworks, procedures, and tactics employed to safeguard all facets of the internet of things [3]. IoT ecosystem integrity, availability, and confidentiality must be guaranteed for all physical components, applications, data, and network connections. By preventing unauthorized access to IoT devices, it is possible to avoid their leaking crucial data or acting as a gateway into other regions of the network.IoT was listed as one of the six crucial civil technologies that might potentially alter US power by the US National Intelligence Council (NIC) in a 2008 assessment. The Mark Weiser-envisioned ubiquitous computing is made possible through IoT. IoT, which connects the physical world to the digital world and is altering how the researcher perceive our surroundings, is no longer a technological buzzword as defined in . Due to a lack of technology and other restrictions on the worldwide scenario, IoT is currently only partially adopted [4]. Wearables, smart household appliances, smart grids, and cars are just a few of the IoT devices that have security issues. Researchers have discovered security holes in cameras that made it easier for hackers to enter into networks as well as in smartwatches with location tracking and the ability to listen in on conversations [5]. Due to the vast majority of security problems that are frequently found in IoT devices, there are numerous security challenges. IoT component hardening, monitoring, firmware updates, access control, threat detection, and vulnerability patching are only a few of the protection measures covered by comprehensive IoT security. Since these systems are widely used, insecure, and a common target for attackers, IoT security is crucial. Considering the importance of security issues facing IoT, this article aims to examine the use of Machine Learning (ML) in

solving the security issues of cognitive.
The organization of this paper is as the following : In Section II we review the relevant background within the literature. In section III we describe the cognitive IoT .In section IV we discuss the use of ML and FL within the CIoT. In Section V we briefly describe the intrusion detection and the most types of (IDS) that used for intrusion detection. In Section VI we describe and discuss the methodology are used in this paper .In Section VII and VIII discuss the findings of the paper. Finally, the paper is concluded along with open research direction in section IX.

II. Related Work

Forecasts predict that there will be 64 billion IoT devices online by 2025. Undoubtedly, the widespread use of these gadgets is creating a world that is extremely linked. The IoT paradigm is enabling new business prospects and application scenarios, Along with new 5G and Beyond 5G (B5G) network technology, these include Smart Cities and Industries 4.0. [6]. However, the quantity and diversity of cyberattacks have increased recently along with the development of new technologies, quickly rendering the present security measures ineffective. Through its pervasive sensing and processing capabilities, the Internet of Things (IoT) aims to facilitate consumer services and applications by linking a variety of devices and things to the network. Intelligent IoT systems are now possible because to the widespread usage of AI techniques like ML and DL, which have the ability to learn from IoT data and gather insights for the development of several smart applications [7]. The studies have found that the massive availability of data on the IoT networks enables a fast and enhance the ML process, which leads to the detection of anomalies more effectively. In most cases, it has been observed that the anomalous activity is an intruder trying to break into the network. Due to the use of ML models, this anomalous activity is detected in minimal time, protecting the customers and service providers from any losses. Federated learning (FL), a distributed ML technique, also employs the local data of dispersed devices to simplify the training of an overall model. Building local models and transmitting the model parameters to the server rather than uploading the raw data has shown promising results in maintaining the privacy of the participants [8]. The ongoing issue of safeguarding privacy in smart cities stems from the expansion of privacy regulations. ML models are frequently trained using a fraction of the data that smart city sensors collect. With this in mind, FL is meant to be used for sensing in smart cities. The participants will be able to add more levels of privacy protection to their data. The performance gains made as a result of the participants' local data contributions can be used to measure the quality of the data that was collected from them. Transmission of model parameters rather than raw data can save communication costs [9]. Additionally, FL-assisted smart city sensing may boost other privacy-preserving technologies like blockchain and various uses of differential privacy. Another study looked at the Industrial Internet of Things' use of blockchain and federated learning (IIoT). The significant growth in the volume of data generated by connected devices in the IIoT paradigm has created a new opportunity to enhance the quality of service for developing applications through data sharing [10]. However, security and privacy concerns are significant impediments to data providers exchanging their data over wireless networks (such as data leaking). Beyond causing the providers to lose money, the disclosure of sensitive information might cause major problems. In this study, the authors proposed a dispersed multiple parties' federated learning-permissioned blockchain-based data-sharing technique for industrial IoT applications. The numerical examples demonstrated how blockchain-enabled data-sharing strategy improves security without relying on centralized trust [11]. Additionally, the effectiveness of the data-sharing system as well as the use of computer resources were enhanced by introducing FL into the permissioned blockchain's consensus process. For IoT, a federated machine learning-based intrusion detection technique can also be utilised, which leaves data generated by devices on them and trains their own models to protect identified data and retain privacy. When a training cycle in FL is conducted across several devices, a server aggregates the changes locally calculated in order to benefit from peers' models [12]. The experimental analysis revealed that the aggregated models, which refer to a centralized model trained over the complete dataset, had an accuracy oscillating of about 83.09 percent after the last round of FL. Additionally, the authors contrasted the FL settings in a self-learning environment with and without sharing model updates with a server. The data revealed that FL outperforms self-learning across all training rounds and for all examined application cases [12]. As a consequence, the study's authors found that federated intrusion detection might achieve equivalent accuracy to centralized intrusion detection, which has a thorough awareness of the whole system. Furthermore, knowledge aggregation in federated intrusion detection enabled the latter to regularly beat the self-learning technique. Further review of the literature highlighted various limitations of FL. There are still many difficulties encountered in competitive contexts since FL is still in the phase of development. To assure the security of FL-enabled smart city sensing, more research on defence mechanisms is required. Problems like statistical heterogeneity of the dispersed data and hardware heterogeneity of participating devices exist because of the nature of a FL environment. [13]. The several recommended fixes have helped to reduce these issues significantly, but they are still present. Last but not least, a thorough analysis of node placement and coverage area is a challenge that needs to be solved before FL can be widely employed for smart city sensing. Furthermore, various frameworks for FL are now actively being developed, and it is anticipated that they will soon have new features and properties. The fact that they do, however, currently permit FL performance in the simulation mode, makes it viable to begin using this technology in industrial systems. One may begin developing neural network-based models on all the frameworks under consideration that will be put to use in production in the

upcoming year [14]. PFL is the FL framework that is the most prepared for commercial application, according to the evaluation of its features and experiments. However, there is a limited development community and little documentation. The usage of this framework's unpopular, proprietary DL PaddlePaddle platform is another drawback. Additionally, PFL requires the most training time. The FATE architecture also employs the federated mode. It has restrictions on the neural network layers that may be employed, yet it works well for Deep Learning (DL) [15]. Decision trees and linear models are also implemented, however they are not yet functional. According to the writers, these flaws will soon be fixed, allowing FATE to be fully utilized in production. Of course, the Google TensorFlow Federated framework must be taken into account. Its primary flaw is that it solely employs DL methods, which have a number of drawbacks including a lack of explainability and a lengthy training period. The 2019 book by Da Costa and de Albuquerque, which focuses on the most recent, in-depth research on ML methods used in the Internet of Things and intrusion detection for computer network security, has been published by Springer. The study aims to undertake a complete and contemporary review of key works that deal with various intelligent techniques and their applicable intrusion detection structures in computer networks, with a focus on the IoT and ML. The researcher examined more than 95 papers on a range of subjects related to security flaws in IoT devices [16]. In order to provide resource optimization, Khan et al. [17] propose a unique Dispersed Federated Learning (DFL) system, where distributed learning gives resilience. they create a challenge for linear integer optimization to lower the overall price of FL for the DFL architecture. The association and resource allocation challenges are first divided into two more manageable problems by the researcher. Second, by relaxing the resource allocation and association subproblems, the researcher turns them into convex optimization problems. The researcher then derives binary association and resource allocation variables using the rounding procedure. They proposed algorithm operates iteratively by resolving one issue variable (for instance, association) and computing the other (for example, resource allocation). Up until the specified cost optimization problem converges, the iterative approach is used. The researcher compare the proposed DFL with two alternative hypotheses, namely random resource distribution and random association. Numerical statistics demonstrate the superiority of the suggested DFL system [17], [18]. According to Shah et al. [19] the key aspects of smart buildings are discussed, along with several ML techniques that may be used in conjunction with IoT technology to increase the efficiency of smart buildings. The use of IoT devices in smart buildings is highlighted via. This assessment emphasizes the IoT devices platform and its elements. This evaluation also discusses security issues with IoT and smart buildings. The Internet of Things (IoTs) and an artificial intelligence system, according to Pathik et al. [20] can be used to create an intelligent accident detection and rescue system that replicates the cognitive processes of the human mind (AI). The creation of an Internet of Things (IoT) kit allows for the detection of accidents as well as the collection and transmission to the cloud of all accident-related data, including position, pressure, gravitational force, speed, and others. Once the accident is discovered, a DL model is employed in the cloud to evaluate the results of the IoT module and turn on the rescue module. All nearby emergency services, including the hospital, police station, mechanics, etc., are alerted as soon as the DL module detects the accident. The false detection rate is reduced through ensemble transfer learning with dynamic weights.The absence of the dataset forces the generation of a tailored dataset from the many online movies. An evaluation of ResNet and InceptionResnetV2 is used to compare the suggested technique to the existing one. The experiment's findings demonstrate that InceptionResnetV2 outperforms ResNet with training, validation, and test accuracy of 98 percentage each. The suggested method is validated on the toy automobile in order to gauge how well it performs in the real world [20].

Wadhwa et al. [21] claim that a hetero federated learning technique is used to apply cognitive learning to data generated by Internet of Things devices. Security for cognitive IoT data is provided by the blockchain's Proof of Work consensus algorithm. To assess the efficacy of our proposed framework, the researcher has done a number of simulations using the blockchain over hetero FL technique. When assessing performance, variables like the effect of the size of the data sample on accuracy dependent on learning rates and the impact of the number of blocks on memory usage are taken into consideration [21]. According to Liu et al. [1] a ML tool's training data is attacked using a combination of conventional interference techniques and data poisoning. The researcher suggesting a new adversarial method to decrease the sensing precision in DL-based spectrum sensing systems and presenting a brand-new jamming waveform design whose interference capability is boosted by data poisoning. According to the simulation results, compared to conventional white-box assault techniques, substantial performance improvement and better mobility may be attained [1]. Providing thorough basic knowledge of ML algorithms and examine the role of ML, Deep Reinforcement Learning (DRL), and Artificial Intelligence (AI) in the creation of the smart city was the primary goal of Mehta [22]. the researcher provides a thorough introduction of the smart city idea and concentrates on various privacy options in the smart city. Highlights the importance of ML in a number of smart city applications, including supply chain management, smart grids, healthcare, and intelligent transportation systems. The article concludes by listing a few potential paths for further research to lead future developments in the field [22]. ML-enabled IoT literature will be categorised and examined from the angles of data, application, and industry. Bzai et al. [23] illustrate how ML and IoT work together to dramatically improve our surrounds through an investigation of around 300 published sources and the implementation of many cutting-edge techniques and applications. They also looked at pandemic management, networked autonomous cars, edge and fog computing, and

lightweight DL, among other cutting-edge IoT technologies. They also divide IoT-related problems into four categories: social, business, personal, and technological. Their research will aid in utilising the IoT's promise and challenges to create and sustain civilization [23]. Sengan et al. [24] demonstrated the combined influence of hunched challenges along the path, issues at the medium get-right-of-area to impact layer, or pack catastrophes precipitated by the remote control moving off course. In this study, the AODV routing MANET protocol is utilized, and the method is built and assessed using support vector machines (SVM) to identify malicious network nodes [24]. Using their optimisation paradigms for cognitive radio networks, Kaur [4] offers a comprehensive categorization and overview of several ML algorithms for intelligent spectrum management. as well as new avenues and unresolved problems for the scientific community to concentrate on in CR networks. By categorizing them into appropriate groups, Ahmed et al. [25] try to investigate the ML spam filtering strategies used in email and IoT systems. Based on accuracy, precision, recall, etc., a thorough comparison of different methods is also made. Detailed conclusions and potential future study areas are also covered in the conclusion [25].

In order to address security and privacy concerns in the context of the Internet of Things, Waheed, et al. [26] provide a synopsis of current research activities from 2008 to 2019 that use ML algorithms and BC technologies. Many of the security and privacy issues that have been brought up in the IoT industry during the past twelve years are listed and categorised in the article's first section. The research of IoT security and privacy initiatives based on ML algorithms and BC technologies is then divided into groups. Last but not least, it discusses and draws attention to a range of difficulties involved in implementing ML algorithms and BC technologies, from security and privacy concerns in the IoT industry to long-term research objectives [26]. According to another studyin 2022, by building different network models to address the lack of the necessary dataset, selecting the best features to enhance model performance, and using a light gradient boosting machine-based algorithm that is optimized for multiclass classification-based attack detection, a novel dataset was created. Multiple metrics, like as the confusion matrix, accuracy, precision, and recall, are used to illustrate the outcomes of long trials. The model was further assessed using multiclass-specific metrics, such as cross-entropy, Cohn's kappa, and Matthews correlation coefficient, in order to assess performance further and eliminate any bias, and then results were compared to previous studies [18]. In an effort to provide a thorough review of the security needs, attack surfaces, and available security solutions for Internet of Things networks, Hussain et.al [27] was mentioned. The shortcomings of these ML and DL based security solutions are then exposed. provides a thorough explanation of the current ML and DL techniques for resolving a number of security challenges in IoT networks. A thorough analysis of the existing literature-based solutions is followed by the provision of fresh research paths for ML and DL-based IoT security [27]. In IoT networks, ML-based Darknet traffic detection systems (DTDS) are being developed, examined, and evaluated by Al-Haija et al. [28]. Use six supervised ML approaches in particular: bagging decision tree ensembles (BAG-DT), ADA decision tree ensembles (ADA-DT), RUS decision tree ensembles (RUS-DT), optimizable decision tree (O-DT), optimizable k-nearest neighbor (O-KNN), and optimizable discriminant (O-DSC). the CIC-Darknet-2020 dataset, which consists of current IoT communication traffic from four different classes that combines VPN and Tor traffic in a single dataset covering a wide range of recorded cyberattacks and hidden services offered by the Darknet, is being used to assess the implemented DTDS models.An actual performance investigation shows that bagging ensemble techniques (BAG-DT) outperform other applied supervised learning approaches in terms of accuracy and error rates, attaining a 99.50% classification accuracy with a low inferencing overhead of 9.09 second. In addition, compare our BAG-DT-DTDS with other DTDS models that are already in use and show how our best results outperform the previous state-of-the-art models by a factor of (1.927%) [28]. Rath et al. [29] examines advanced-level security in network and real-time applications using ML in his book Machine Learning and Cognitive Science Applications in Cyber Security. There was an understanding that allowing machines to benefit from themselves was the greatest method to be able to do this task [29].

According to Philipp Morgner, there are three new technical vulnerabilities that could jeopardise the security and privacy of the Internet of Things [30]. The researcher is looking into the privacy dangers associated with using smart heating systems that capture data about the indoor environment. In addition to assuming that the attacker has access to temperature and relative humidity data, the researcher trains ML classifiers to recognise occupants and distinguish between different types of activity. The results show that there are serious privacy repercussions when room temperature data is leaked. The researcher considers a second issue as the proliferation of IoT infrastructure, which is making it possible for new attack methods against hardware security.and focuses on the degree to which malicious things' supply chains may be changed in a way that makes it possible for attackers to commandeer these gadgets after they have been put into use. To do this, The researcher develops and introduces a malicious Internet of Things implant, which is inserted into any electrical product. In order to assess these implants, device-level assaults are used against essential security and safety-related components. They addresses a third risk by looking at the security of the popular network standard for smart homes, ZigBee. The researcher introduces novel attacks that explicitly take advantage of one of the standard's commissioning measures, indicating that it is a dangerous design. Examining these issues reveals that attackers may influence ZigBee devices and networks from more than 100 meters away and eavesdrop on important data. Customers, according to the research, understand suggested labels, have a substantial effect on their purchase decisions, and so have the ability to urge manufacturers to give long-term security support [30]. An overview of the use of ML in

wearable Wireless Body Area Network (WBAN) is provided by Al-Turjman [31]. It emphasizes the key difficulties and unresolved problems with using ML models in such delicate networks. This study intends to report on the many ways that ML is applied to these networks for their benefit, the design elements taken into account when ML algorithms are implemented, and the communication technologies utilized to link wearable WBAN in the IoT age. The research that have been published are supported by comprehensive simulation data and real-world experiment outcomes [31].

It is clear from the review that FL offers a great deal of promise to address the security issues that IoT devices face. However, further study and research are needed to standardize the procedures through suitable frameworks because FL is still in the emergent stage. This will enable the implementation to be improved.

III. Cognitive IoT

The use of cognitive computing technologies to the data produced by the IoT ecosystem's linked devices is known as Cognitive Internet of Things (CIoT). The rise of cognitive aspects in information systems has led to the development of smart systems. CIoT is an example of advanced technology that integrates smart systems like ML and data analytics [32]. However, these systems are highly complex and complicated, due to which security concerns are also increased. In this regard, the use of ML can be effective to solve security issue of CIoT.

The main goal following Joseph Mitola's 1999 coining of the term "cognition" was to fully use the limited radio frequencies that were shared among cooperative users. Cognitive IoT was eventually used when cooperative behavior became even more necessary to boost network performance by making network components more interactive and intelligent. The primary goal of CIoT is to develop self-driven intelligent IoT systems by embedding Artificial Intelligence capabilities that are able to autonomously use sensed knowledge, and take appropriate actions in response to observations. However, because the network's constituent entities are diverse in their nature and focused on different domains, it is not an easy procedure to improvise. To accomplish so, a multi-domain cooperation strategy is applied. As a result, items may be integrated across network domains relatively quickly. The primary motivation behind the development of CIoT is to minimize human involvement by incorporating the cognitive process. The network is improved by combining system design with human level intelligence. With relatively little power consumption and great agility, the sensor nodes are deployed with each of the devices that make up the network. To ensure that all of the resources are evenly divided across all of the devices sharing the network, cognition is required to bring about uniform resource usage in the network. Three hypothetical levels, including the core and foundation of the internet, communication between objects, and ultimately the flow of business process, may be used to understand the general architecture of the CIoT. By identifying the network's constituent items and classifying them into domains, the CIoT network is created. This kind of grouping makes each entity dependent on and loosely associated with the others. The entities in the network are referred to as nodes because they include sensors of many kinds, including RFID, GPS, infrared, and other types, along with transmission and receiving capabilities. Without the internet, the CIoT cannot function as intended in terms of data transfer. The data gathered across the whole network is unprocessed raw data with no cognitive capabilities. A new network typology with the primary role of bridging the gap between hypothetical and real-world notions has emerged as a result of the advent of CIoT in information technology. The basic goal of CIoT is to include an intelligent system with the capacity to learn, think, and comprehend, which prevents the interference of outside organizations.

IV. Machine Learning and Federated Learning

Data is used to train and test models in ML, a significant area of artificial intelligence, which can be deployed to take actions like humans. These trained models automate the processes and enhance the security of the network. Therefore, their use in cognitive IoT also has high significance [33]. The major focus of this study is on FL, which enables more efficient training and performance of models. FL allows mobile devices to work together to build a shared prediction model while still keeping all of the training data on the device since it separates the ability to apply ML from the obligation to store training data in the cloud. (View Fig. 1) Model training on the device goes much further than simply employing local models on mobile devices to produce predictions [2]. As a result, FL approach is cutting edge and ideal for IoT systems (View Fig. 2).

Fig. 1. Centralised Model embedding Data Analytics &Machine Learning

V. Intrusion Detection

When unexpected activity is spotted in network traffic, an intrusion detection system (IDS) watches it for it and sends out alarms. Software is used to identify illegal behaviour or policy violations in a system or network [34]. Typically, a security information and event management (SIEM) system is used to alert administrators to any unauthorised activity or breach or to centrally log it [35]. In order to differentiate between real and fake alarms, a SIEM system aggregates the outputs from multiple sources. Intrusion detection systems are

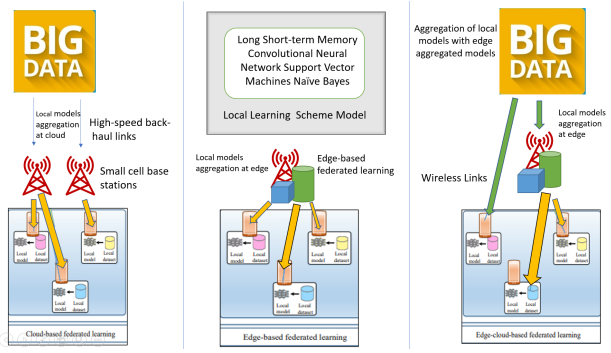

Fig. 2. Federated Learning for IoT Networks

prone to raising erroneous alerts while scanning networks for potentially harmful behaviour [36]. Organizations must therefore modify their IDS systems after the initial installation. Intrusion detection systems need to be configured correctly in order to distinguish between safe network traffic and malicious activity. Intrusion prevention systems check network packets entering the system for malicious activity, and if any is detected, alarms are promptly sent out.The following types of IDS are very commonly used for intrusion detection:

1) System for detecting network intrusions (NIDS)
   Internal system monitoring via HIDS handles connections to log files, user activity, and other things. In contrast, NIDS examines both incoming and outgoing network traffic.
2) Host intrusion detection system (HIDS)
   HIDS's adaptation is reliable despite problems with tampering attempts. In isolation, a host-based IDS is not the best option. It has serious drawbacks like high resource usage that impairs host performance. Such attacks might go undetected unless they manage to get into the target [37].
3) Protocol-based intrusion detection system (PBIDS)
   Anomaly-based detection and signature-based detection are the two primary techniques for intrusion detection. To check for malicious activity, anomaly-based intrusion detection monitors unusual behaviours in the systems [38]. On the other hand, signature-based detection works by examining the sequence (signature) of the activities to determine the intrusions. Recent trends in IDS through anomaly detection have moved towards the use of ML methods. The use of ML models has significantly enhanced the performance of IDS. Due to this, most research and development is being performed in this area to make intrusion detection optimal through the deployment of the most efficient and effective ML models [39]. FL is one of the most recent concepts integrated with intrusion detection to enhance its performance.

VI. Methods

Because of the nature of this study, which attempts to define the terminology of the Cognitive Internet of Things and shed light on security and privacy, and learn about ML and and how to employ them in achieving security and privacy, and based on the study's objectives and questions that it seeks to answer, this study required the use of the descriptive analytical approach, which can be He described it as a method of analysis and interpretation in an organized scientific manner in order to achieve specific goals. By examining its variables, defining its ideas, and addressing what is conceivable from diverse perspectives, the issue under investigation can be better understood. The study population consists of all workers in the field of security and privacy in the CIoT. In addition to selecting a random sample consisting of (***) employees who are responsible for the security and privacy of the CIoT. In order to achieve the theoretical basis for the study and its directives, This study has performed secondary data collection using a literature review to accomplish the research aim. A systematic literature review has been performed to obtain the results and reach the findings. The selection criteria for the literature were based on the years of publication and only the articles published in and after 2019 were used for the review. Through this practice, this research ensured that the data collected was recent. Besides this, data was collected from credible sources like IEEE Access to ensure that the results are accurate.

VII. Contribution

In the CIoT, object and resource management, object identification, access control, network, and communication technologies have received the majority of research and development efforts. In this study, we seek to expand our knowledge of how ML plays a crucial role in defining the Cognitive IoT landscape rather than concentrating on important Cognitive IoT developments. The researchers working on cognitive IoT-based ML will use this study as a foundation. The primary objective of this study is to further knowledge of ML, as well as its applicability and significance to the field of cognitive internet of things (IoT). Significant contributions to this work:

- First, categorising the work being done in terms of research and development for the Cognitive Internet of Things.
- Second, the study offers information on the most recent studies and recent advancements in the Cognitive Internet of Things, with a focus on advancements in ML.
- Thirdly, the study outlines burgeoning Cognitive IoT themes that will leverage machines as its fundamental building blocks to create innovative and long-lasting solutions.
- Finally, the study also aids readers in identifying potential directions for future ML research based on the Cognitive Internet of Things.

The research findings show that the use of federated learning allows IoT networks to become smarter against intrusion detection. Federated learning enables an efficient data training process due to the ease of availability of the data of the IoT systems. The most effective application of federated learning is in the prediction of an activity to be malicious or not. Analysis

of anomalous activities and prediction based on this analysis is the main feature of federated learning that makes it highly effective in intrusion detection in IoT systems. The table below shows the results gained from a systematic review of various studies.

TABLE I
THE SUMMARY OF FINDINGS SYSTEMATIC REVIEW OF VARIOUS STUDIES

Study	Research Method	Findings
[12]	Survey	According to the study's findings, federated learning has a number of benefits for IoT systems since it can be used to ensure the security of local devices and is not constrained by centralised systems. This federated learning capability makes it possible to respond to harmful activity on local devices more quickly and effectively while shielding users from any harm However, the research has also identified challenges to the implementation of federated learning in IoT that include high costs and the need for more research and development.
[13]	Survey	The research finds that federated learning has high potential in catering to the needs of security on IoT systems by ensuring privacy and integrity during the process of data collection. Various attacks can be identified and prevented through the use of federated learning. However, the research also finds that federated learning itself is susceptible to adversarial attacks that can harm the model performance and put local devices at risk.
[14]	Experiment	The authors of the study have performed an experimental study to examine the role of federated learning on malware detection. The results of the study show that the performance of models improves due to the increase in data. The IoT networks usually have a large number of devices connected, which leads to enhanced training of models, which enables them to perform more accurately in malware detection. In order to combat adversarial assaults, a number of robust aggregation functions have been used in this study, with median aggregation providing the most promising but still inadequate results. Therefore, adversarial attacks remain the most significant challenge for federated learning models as a single client can harm the model performance majorly.
[15]	Survey	This research has compared the performance of federated learning models with traditional models. It was noted that the traditional model relies on the use of central systems for intrusion detection, while federated learning is a decentralized approach. Due to this, the performance of federated learning is more effective. However, the lack of standardization in federated learning is a major challenge and there is a need to ensure the development of effective frameworks for improved implementation.
[40]	Experimental	The study used federated learning models and blockchain to fully protect the IoT platforms. According to the report, federated learning ensures that any suspicious activity is quickly identified, and blockchain defends IoT systems from a variety of assaults. Because blockchain also protects federated learning from adversarial attacks, using it with federated learning produces better results in intrusion detection.
[41]	Experimental	The effectiveness of federated learning models and centralised machine learning methodologies is compared in the study. A federated learning-based approach to IoT intrusion detection is more effective and efficient since it protects data privacy through local training and inference of detection models. By broadcasting only their changes to a remote server, which aggregates them and gives participating devices a better detection model, devices in this manner can keep their privacy while also having access to the expertise of their peers.

VIII. Discussion

The Internet of Things (IoT) is a highly inventive technology that has huge potential and is growing exponentially. A network of linked objects known as the Internet of Things (IoT) exchanges data to offer cutting-edge applications. IoT gadgets come in a variety of forms, from low-power devices to smart objects. Processes may be automated with IoT, which saves time and money. Sensors, cameras, and other IoT devices often gather data that is subsequently shared with a server for analysis and monitoring. The integrity of the data stored on the server should be maintained in a way that guards against malicious attempts to change the data. Additionally, other people and systems should always have access to the data [26].

The implementation of IoT networks in several systems has led to a growth in the number of IoT devices. IoT device numbers are predicted to increase from 7.74 in 2019 to 25.44 billion in 2030. The fundamental issue with these gadgets is that security is generally not taken into account. Additionally, the login and password are not modified during deployment. IoT devices like cameras so become the primary target for attackers. They attempt to breach them before using them to launch distributed denial-of-service (DDoS) attacks or utilize them as a botnet to steal data [42].

Internet of Things (IoT) devices are widely employed in several sectors because of advances in network technology and computer capacity. But there are a number of security risks as well. Anomaly detection is a widely used technique, however, conventional approaches have drawbacks including poor accuracy. In order to increase accuracy and make use of the benefits of federated learning to safeguard local data security, a decentralized federated learning approach for anomaly detection has been found to be very effective [9], [10], [12]. The decentralised algorithm avoids the single point of failure connected to traditional federated learning. Due to this, the

performance of federated learning models is more effective. Various studies have been conducted that present different methods and approaches to conducting federated learning on IoT systems. Before federated learning can be widely used, there are still some obstacles that must be removed. The adversarial assaults that can render federated learning models utterly useless are one of the biggest obstacles in this area [13]–[15].

It is possible to employ ML techniques to enhance the functionality of IoT infrastructure and cybersecurity solutions (such as smart sensors and IoT gateways). Based on the most recent understanding of cyberthreats, these algorithms may keep track of network traffic, update threat knowledge databases, and maintain the security of the underlying systems. Extensive, complex datasets are mined for precise information using ML techniques. The collected results can be used to predict and identify the flaws in IoT-based systems. Blockchain (BC) technologies are also becoming more and more well-liked as a solution to the security and privacy issues in modern IoT applications. Both ML algorithms and BC approaches have been the subject of several investigations [27].

A solution for this has been proposed in the shape of blockchain, which can make IoT systems very secure against security attacks. Nonetheless, the costs of implementation and lack of standardization of federated learning remain an issue. In this regard, it is proposed that more research and development should be conducted so that federated learning can be standardized. Standardization of federated learning is essential as it will help enhance the implementation and maintenance of IoT systems [43], [44]. Costs will also go down as a result of this. The table below displays the key characteristics of the federated learning model for intrusion detection in IoT systems. Future research in this field will benefit from the findings of a SWOT analysis for federated learning in order to improve the development of IoT systems.

TABLE II
THE MAIN FEATURE OF THE FEDERATED LEARNING MODEL IN IOT (SID)

Strengths	Weaknesses	Opportunities	Threats
• Decentralized training • Higher performance • Timely detection of malware	Lack of standardization and high costs do not allow large-scale implementation of FL models.	A significant focus of the researchers and experts on federated learning models will lead to their improved development and implementation of IoT networks.	Adversarial attacks are the major threat to the performance of federated learning models.

IX. Conclusion and Future Direction

As a rising paradigm, cognitive IoT introduces several advantages to the IoT ecosystem through employing cognitive learning. Despite the latter, the emergence of cognitive IoT is not security risk-free, and requires guarantees to ensure the security of such a promising technology. In this paper, we have reviewed several research studies which shed the lights on the use of ML to address the security issues of IoT networks. The review focused on the implementation of federated learning approaches to secure IoT networks and make them smarter to detect malware and intrusions. Comparing federated learning to centralized and conventional techniques, it has been discovered that its performance is competitive with a high potential. In addition, since FL models rely on local devices rather than a centralized method, they are more effective in training and detection. Therefore, federated learning implementation needs to be improved to ensure that IoT system security problems are resolved. Hence, the implementation of federated learning should be enhanced to ensure that the security issues of IoT systems are overcome. Nonetheless, there are some threats and weaknesses of federated learning models that need to be overcome in future studies, which include high costs, lack of standardization, and susceptibility to adversarial attacks. It is advised that more study be done in this area so that FL models can perform even better in the future. Before FL models can be widely used for intrusion detection, there are a number of important holes that must be filled. Future research is necessary to strengthen FL model security so that they can function effectively in any situation without being vulnerable to any kind of attacks.

References

[1] M. Liu, H. Zhang, Z. Liu, and N. Zhao, "Attacking spectrum sensing with adversarial deep learning in cognitive radio-enabled internet of things," *IEEE Transactions on Reliability*, 2022.

[2] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for internet of things: Recent advances, taxonomy, and open challenges," *IEEE Communications Surveys & Tutorials*, 2021.

[3] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for iot security attacks," *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 2545–2554, 2021.

[4] A. Kaur and K. Kumar, "A comprehensive survey on machine learning approaches for dynamic spectrum access in cognitive radio networks," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 34, no. 1, pp. 1–40, 2022.

[5] L. Cui, Y. Qu, G. Xie, D. Zeng, R. Li, S. Shen, and S. Yu, "Security and privacy-enhanced federated learning for anomaly detection in iot infrastructures," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3492–3500, 2021.

[6] A. Yazdinejad, R. M. Parizi, A. Dehghantanha, and H. Karimipour, "Federated learning for drone authentication," *Ad Hoc Networks*, vol. 120, p. 102574, 2021.

[7] A. R. Javed, M. A. Hassan, F. Shahzad, W. Ahmed, S. Singh, T. Baker, and T. R. Gadekallu, "Integration of blockchain technology and federated learning in vehicular (iot) networks: A comprehensive survey," *Sensors*, vol. 22, no. 12, p. 4394, 2022.

[8] Z. Chen, W. Liao, K. Hua, C. Lu, and W. Yu, "Towards asynchronous federated learning for heterogeneous edge-powered internet of things," *Digital Communications and Networks*, vol. 7, no. 3, pp. 317–326, 2021.

[9] T. Alam and R. Gupta, "Federated learning and its role in the privacy preservation of iot devices," *Future Internet*, vol. 14, no. 9, p. 246, 2022.

[10] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6g communications: Challenges, methods, and future directions," *China Communications*, vol. 17, no. 9, pp. 105–118, 2020.

[11] H. Yang, J. Yuan, C. Li, G. Zhao, Z. Sun, Q. Yao, B. Bao, A. V. Vasilakos, and J. Zhang, "Brainiot: Brain-like productive services provisioning with federated learning in industrial iot," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 2014–2024, 2021.

[12] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1622–1658, 2021.

[13] J. C. Jiang, B. Kantarci, S. Oktug, and T. Soyata, "Federated learning in smart city sensing: Challenges and opportunities," *Sensors*, vol. 20, no. 21, p. 6230, 2020.

[14] V. Rey, P. M. S. Sánchez, A. H. Celdrán, and G. Bovet, "Federated learning for malware detection in iot devices," *Computer Networks*, vol. 204, p. 108693, 2022.

[15] I. Kholod, E. Yanaki, D. Fomichev, E. Shalugin, E. Novikova, E. Filippov, and M. Nordlund, "Open-source federated learning frameworks for iot: A comparative review and analysis," *Sensors*, vol. 21, no. 1, p. 167, 2020.

[16] K. A. da Costa, J. P. Papa, C. O. Lisboa, R. Munoz, and V. H. C. de Albuquerque, "Internet of things: A survey on machine learning-based intrusion detection approaches," *Computer Networks*, vol. 151, pp. 147–157, 2019.

[17] L. U. Khan, M. Alsenwi, I. Yaqoob, M. Imran, Z. Han, and C. S. Hong, "Resource optimized federated learning-enabled cognitive internet of things for smart industries," *IEEE Access*, vol. 8, pp. 168 854–168 864, 2020.

[18] F. Zahra, N. Jhanjhi, S. N. Brohi, N. A. Khan, M. Masud, and M. A. AlZain, "Rank and wormhole attack detection model for rpl-based internet of things using machine learning," *Sensors*, vol. 22, no. 18, p. 6765, 2022.

[19] S. F. A. Shah, M. Iqbal, Z. Aziz, T. A. Rana, A. Khalid, Y.-N. Cheah, and M. Arif, "The role of machine learning and the internet of things in smart buildings for energy efficiency," *Applied Sciences*, vol. 12, no. 15, p. 7882, 2022.

[20] N. Pathik, R. K. Gupta, Y. Sahu, A. Sharma, M. Masud, and M. Baz, "Ai enabled accident detection and alert system using iot and deep learning for smart cities," *Sustainability*, vol. 14, no. 13, p. 7701, 2022.

[21] S. Wadhwa, S. Rani, G. Kaur, D. Koundal, A. Zaguia, and W. Enbeyle, "Heterofl blockchain approach-based security for cognitive internet of things," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.

[22] S. Mehta, B. Bhushan, and R. Kumar, "Machine learning approaches for smart city applications: Emergence, challenges and opportunities," *Recent Advances in Internet of Things and Machine Learning*, pp. 147–163, 2022.

[23] J. Bzai, F. Alam, A. Dhafer, M. Bojović, S. M. Altowaijri, I. K. Niazi, and R. Mehmood, "Machine learning-enabled internet of things (iot): Data, applications, and industry perspective," *Electronics*, vol. 11, no. 17, p. 2676, 2022.

[24] S. Sengan, O. I. Khalaf, G. R. K. Rao, D. K. Sharma, K. Amarendra, and A. A. Hamad, "Security-aware routing on wireless communication for e-health records monitoring using machine learning," *International Journal of Reliable and Quality E-Healthcare (IJRQEH)*, vol. 11, no. 3, pp. 1–10, 2022.

[25] N. Ahmed, R. Amin, H. Aldabbas, D. Koundal, B. Alouffi, and T. Shah, "Machine learning techniques for spam detection in email and iot platforms: analysis and research challenges," *Security and Communication Networks*, vol. 2022, 2022.

[26] N. Waheed, X. He, M. Ikram, M. Usman, S. S. Hashmi, and M. Usman, "Security and privacy in iot using machine learning and blockchain: Threats and countermeasures," *ACM Computing Surveys (CSUR)*, vol. 53, no. 6, pp. 1–37, 2020.

[27] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, "Machine learning in iot security: Current solutions and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1686–1721, 2020.

[28] Q. Abu Al-Haija, M. Krichen, and W. Abu Elhaija, "Machine-learning-based darknet traffic detection system for iot applications," *Electronics*, vol. 11, no. 4, p. 556, 2022.

[29] M. Rath and S. Mishra, "Advanced-level security in network and real-time applications using machine learning approaches," in *Research Anthology on Machine Learning Techniques, Methods, and Applications*. IGI Global, 2022, pp. 664–680.

[30] P. Morgner, "Security and privacy in the internet of things: Technical and economic perspectives," Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2019.

[31] F. Al-Turjman and I. Baali, "Machine learning for wearable iot-based applications: A survey," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 8, p. e3635, 2022.

[32] B. Ghimire and D. B. Rawat, "Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for internet of things," *IEEE Internet of Things Journal*, 2022.

[33] J. Pang, Y. Huang, Z. Xie, Q. Han, and Z. Cai, "Realizing the heterogeneity: A self-organized federated learning framework for iot," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3088–3098, 2020.

[34] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 16–24, 2013.

[35] Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4150, 2021.

[36] T. Saranya, S. Sridevi, C. Deisy, T. D. Chung, and M. A. Khan, "Performance analysis of machine learning algorithms in intrusion detection system: a review," *Procedia Computer Science*, vol. 171, pp. 1251–1260, 2020.

[37] M. Liu, Z. Xue, X. Xu, C. Zhong, and J. Chen, "Host-based intrusion detection system with system calls: Review and future trends," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–36, 2018.

[38] M. Aloqaily, S. Otoum, I. Al Ridhawi, and Y. Jararweh, "An intrusion detection system for connected vehicles in smart cities," *Ad Hoc Networks*, vol. 90, p. 101842, 2019.

[39] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer networks*, vol. 174, p. 107247, 2020.

[40] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial iot," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4177–4186, 2019.

[41] S. A. Rahman, H. Tout, C. Talhi, and A. Mourad, "Internet of things intrusion detection: Centralized, on-device, or federated learning?" *IEEE Network*, vol. 34, no. 6, pp. 310–317, 2020.

[42] X. Li, W. Chen, Q. Zhang, and L. Wu, "Building auto-encoder intrusion detection system based on random forest feature selection," *Computers & Security*, vol. 95, p. 101851, 2020.

[43] E. Anthi, L. Williams, M. Słowińska, G. Theodorakopoulos, and P. Burnap, "A supervised intrusion detection system for smart home iot devices," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9042–9053, 2019.

[44] I. Butun, P. Österberg, and H. Song, "Security of the internet of things: Vulnerabilities, attacks, and countermeasures," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 616–644, 2019.