International Journal of Engineering Technology and Applied Science

(ISSN: 2395 3853), Vol. 10 Issue 2 February 2024

A Review on Data Privacy Using Federated Learning and Deep Learning

Vipin Verma[1]

[1] Ph.D Scholar, School of Information Technology, UTD, RGPV, Bhopal 462033, MP, India
vipinvermacs@gmail.com

Abstract: A robust security framework that effectively safeguards IoT systems against potential adversarial attacks is in high demand. However, creating an appropriate security model for IoT is complex due to its dynamic and distributed nature. This complexity has spurred researchers to delve into the role of machine learning (ML) in shaping security models. An examination of various ML algorithms for IoT security is presented, accompanied by an exploration of their strengths and limitations. Existing research highlights concerns such as significant computational overhead and privacy risks associated with ML algorithms. This review concentrates on the application of federated learning (FL) and deep learning (DL) algorithms to enhance IoT security. Unlike traditional ML methods, FL models can preserve data privacy while sharing information with other systems. The study proposes that FL has the potential to address the shortcomings of conventional ML approaches by maintaining data privacy during information sharing. The review delves into diverse models, provides an overview, conducts comparisons, and offers a summarization of FL and DL-based techniques for bolstering IoT security.

Keywords: Internet of Things, Machine learning, Federated learning, Deep learning, Privacy preservation.

1. Introduction

IoT constitutes a network comprising diverse interconnected devices, including sensors and actuators, functioning at an accelerated data collection pace. By employing a network of modest sensors and interconnected entities, comprehensive data acquisition concerning our environment and surroundings can be achieved at an elevated magnitude. The ubiquity of IoT is steadily on the rise, with projections indicating a proliferation of approximately 27 billion IoT-based applications by the year 2024 [1]. The escalating prominence of IoT can be attributed to its remarkable attributes encompassing automation, dependability, scalability, and resilience. These distinctive traits possess the potential to revolutionize forthcoming IoT applications and elevate the caliber of service (QoS) provided by such applications, spanning domains like smart urban centers, intelligent healthcare [2], industrial automation [3], and advanced transportation systems [4].

Nonetheless, the incorporation of the IoT system with a multitude of devices introduces security apprehensions within IoT applications. IoT establishes connections with other interconnected devices via a centralized server, a factor that amplifies concerns regarding privacy and security. The diverse and ever-changing nature of IoT devices renders them susceptible to various forms of threats and security breaches [5]. Furthermore, the heightened vulnerability of IoT devices increases the risk of device impersonation, thus presenting significant challenges to the data security and privacy of IoT systems. Navigating these challenges necessitates the establishment of a robust security framework capable of thwarting adversarial attacks on IoT. However, the process of devising an effective security strategy for IoT is immensely intricate and demanding, primarily due to the inherent limitations imposed by resource constraints within the IoT system [6]. A substantial number of IoT devices possess restricted resources such as computational capabilities, bandwidth, and memory, which do not align with the demands of intricate security solutions.

Traditional security methodologies, including techniques like malware identification, access control, device authentication, and cryptography-based methods, have been previously suggested to enhance IoT security [7,8,9]. Yet, the task of identifying various categories of cyber attacks and security vulnerabilities through these approaches is exceptionally intricate. Moreover, unresolved and critical operational issues introduce security vulnerabilities that erode the reliability of the IoT framework. The need of the hour is a transformation of existing security strategies to effectively detect emerging cyber attacks. This necessitates the deployment of an intelligent and astute model capable of recognizing diverse forms of attacks within the IoT realm, encompassing denial of service (DoS), distributed DoS (DDoS), flooding attacks, jamming attacks, and botnet attacks [10]. Leveraging artificial intelligence (AI) based techniques in devising and crafting an intelligent attack detection model holds promise for fortifying IoT systems. Harnessing the depth of AI knowledge, particularly machine learning (ML), empowers researchers to identify anomalies and unwarranted malicious activities within the IoT ecosystem. Consequently, this offers a dynamic security solution that remains in a state of continuous enhancement and relevance [11]. Specifically, machine and deep learning (DL) models encompass an array of principles, methods, and intricate transformation functions designed to extract valuable insights and noteworthy data patterns from security-related data. This lays the foundation for training machines to prognosticate threats or risks at their nascent stages, thereby enabling proactive defense mechanisms.

Nonetheless, traditional machine learning (ML) algorithms demand a substantial volume of training data to execute specific tasks. The acquisition of extensive datasets for ML models escalates concerns related to privacy and security. Furthermore, ML models encounter challenges linked to privacy breaches arising from the necessity to transmit device data to a centralized third-party server. The feasibility of implementing centralized ML models within the IoT context is hampered by the considerable data size, and the computational expense entailed in training such extensive models [12]. More recently, federated learning (FL) has emerged as a viable alternative to surmount the limitations inherent in conventional ML algorithms [13,14]. In contrast to conventional ML models, the migration of data into a central server is not obligatory in the framework of federated learning. This reduction in reliance on centralized servers significantly curtails the potential for privacy breaches, rendering it a favored approach compared to traditional ML algorithms. Another promising technology extensively employed in fortifying the security of IoT systems is deep learning (DL). The effectiveness of DL models in enhancing the security of IoT systems has been well demonstrated [15].

While both Federated Learning (FL) and Deep Learning (DL) fall under the umbrella of machine learning (ML), this review delves into these two models to offer an all-encompassing analysis, encompassing comparisons and their respective importance within IoT security. The objectives of this survey are as follows:

- Recognize and categorize privacy risks within IoT systems.
- Highlight the obstacles that hinder the effective utilization of federated learning and deep learning for safeguarding privacy.
- Conduct an exhaustive assessment of research endeavors employing federated learning and deep learning techniques to ensure privacy within the IoT domain.
- Pinpoint the most significant and prospective avenues for future exploration.

2. Related Work

Over recent years, the realms of privacy, machine learning (ML), and the Internet of Things (IoT) have intersected. Given the pivotal role that ML and cyber security play within IoT landscapes, numerous researchers have undertaken surveys and tutorials to provide pragmatic guidance in addressing cyber security threats and charting pathways for forthcoming solutions. However, the majority of existing surveys primarily tackle cyber security concerns in IoT settings while overlooking the aspect of privacy. These surveys commonly assign greater prominence to other cyber security threats such as network attacks or software-based breaches. In contrast, the limited numbers of surveys that center on privacy provide a constrained perspective by focusing solely on specific privacy-preserving solutions.

2.1. Cyber security based on machine learning

Multiple surveys have examined machine learning (ML) based cybersecurity solutions, yet few have delved into the realm of privacy. For instance, [16] conducted an analysis of ML-based Internet of Things (IoT) solutions, focusing primarily on security. They scrutinized vulnerabilities and attack surfaces, evaluating the pros and cons of various ML models within different IoT layers. In the context of privacy, they briefly explored studies highlighting the potential data leakage from ML algorithms [17]. This susceptibility affects privacy-preserving

ML and deep learning (DL) algorithms, leaving them exposed to predominant attacks, such as distributed DL methods that cannot maintain the confidentiality of the training set. The survey ultimately underscores that privacy preservation in this domain is in its early stages and warrants further investigation. Author [18] conducted a review of ML-based solutions for IoT cyberattacks across network, malware, and privacy dimensions. In the privacy sphere, their examination centered on works employing Federated Learning (FL), an Artificial Intelligence (AI) model development framework distributed across edge devices. FL ensures secure models while safeguarding user privacy and enhancing performance. Their focus extended to healthcare, exploring adaptive access control mechanisms, like Ciphertext Policy Attribute-based Encryption (CP-ABE) and dual encryption, as well as Merkle Hash Trees (MHTs).

In a similar vein, [19] provided an expansive overview of ML-based cybersecurity solutions, particularly diverging in their exploration of Blockchain (BC) technologies to bolster user privacy. They introduced taxonomy for IoT security and privacy solutions involving ML algorithms and BC techniques, predicting their future integration. Privacy breaches were categorized into Man-in-the-middle (MiTM) attacks and Data Privacy. ML-based approaches like Stochastic Gradient Descent (SGD), Logistic Regression (LR), Oblivious Evaluation of Multivariate Polynomial (OMPE), Support Vector Machines (SVM), and Naive Bayes (NB) were scrutinized.

Author [20] contributed by evaluating ML-based cybersecurity solutions within 5G networks. Their focus centered on recent studies addressing privacy concerns arising from service providers' usage of private user data. However, their survey mainly offered guidelines on the responsible use of private data to mitigate privacy issues. Turning to DL models, [21] analyzed existing cybersecurity threats and DL-based solutions. Within the privacy domain, they tackled adversarial attacks that could compromise user privacy, including scenarios where Deep Neural Networks (DNNs) could be manipulated to yield false inputs and misclassifications. They highlighted existing remedies based on defensive distillation and targeted gradient sign methods [22]. Author [23] expanded their survey to encompass DL-based solutions for various cyberattacks within mobile networks. Privacy concerns prompted a review of DL-based privacy preservation strategies, classified into collaborative DL, differential privacy, and training on encrypted data. They observed that most of these works adopted supervised learning. Collaborative DL solutions only shared partial data, and the utilization of differential privacy averted dataset exposure. Furthermore, solutions employing encrypted datasets for DNN training achieved reasonable accuracy but exhibited notably slower performance compared to non-encrypted data. Lastly, [24] homed in on privacy and security issues tied to Federated Learning (FL). They dissected unintentional data leakage, model reconstruction attacks, Generative Adversarial Network (GAN)-based inference attacks, and inference attacks. Their evaluation encompassed privacy-preserving techniques, including gradient noise addition and compression, Secure Multi-Party Computation (SMC), Differential Privacy, and Blockchain technology. Their findings underscored inference-based attacks as particularly menacing to FL privacy. Overall, these surveys collectively highlight a consistent observation: privacy protection within IoT environments remains nascent.

2.2. Surveys on Privacy in IoT

Numerous studies have examined privacy concerns within IoT environments, yet their focus has largely been on privacy threats and attacks. In contrast, [25] introduced a comprehensive approach by proposing a threat model and taxonomy that classifies diverse attack types based on involved actors and assets. They identified key assets susceptible to privacy breaches in machine learning (ML) contexts, encompassing the training dataset, the model itself, its parameters, hyper-parameters, and architecture. The relevant actors include data owners, model owners, and model consumers. The duo extensively analyzed privacy-related attacks, targeting knowledge acquisition such as training data and model information. These attacks were grouped into four primary categories: membership inference, reconstruction, property inference, and model extraction. After a thorough assessment of various ML models, the authors concluded that Decision Trees (DT), Linear Regression (LR), and Support Vector Machine (SVM) exhibit heightened vulnerability to such attacks.

Lastly, in the realm of IoT environments, the surveys predominantly center on specific privacy-preserving solutions. However, most of these inquiries tend to concentrate on differential privacy without delving into the variations in learning architectures: centralized, distributed, and federated. In [26] conduct a review that encompasses significant privacy threats within IoT settings, categorizing them into four groups: authentication and authorization, edge computing mediators, data anonymization, and data summarization. They highlight primary shortcomings in existing solutions, notably the lack of thorough performance evaluation, real-life scenario assessments, and precision in data access granularity. [27] Undertakes an exploration of centralized privacy-preserving solutions that harness ML within IoT contexts. Initially, they classify data generated across distinct IoT layers and subsequently scrutinize ML-based privacy solutions. The studies examined reveal that ML techniques employed for safeguarding privacy are tailored to specific data categories. The authors also emphasize the absence of standardization and interoperability as impediments in crafting more comprehensive global privacy-preserving solutions. Similar in approach, [28] analyze privacy preservation solutions to identify fundamental traits that an IoT privacy framework should embody to ensure user privacy while aligning with GDPR requisites. Their scrutiny leads them to propose a fusion of the following components: (1) ML techniques for enhancing user privacy protection, (2) policy languages to delineate user privacy preferences and express intricate policies, and (3) deployment of negotiation techniques to enhance user services while upholding their privacy.

In [29] delve into the domain of privacy-preserving ML, distinguishing between solutions geared toward safeguarding privacy during prediction and those during training. In the prediction sphere, these solutions adopt strategies like differential privacy, secure multi-party computation, and homomorphic encryption. Training-focused privacy-preserving solutions are categorized into collaborative and aggregation scenarios, both of which embrace differential privacy, homomorphic encryption, and secret-sharing mechanisms. Lastly, [30] delve into the gamut of techniques aimed at shielding user privacy within Federated Learning (FL). Their classification groups these techniques into three primary categories: firstly, methods such as k-anonymity, l-diversity, and t-closeness that secure user privacy before data availability; secondly, techniques that ensure privacy during the training phase; and thirdly, techniques grounded in Differential Privacy, providing user protection at all stages of data model training. Their conclusion underscores that while Differential Privacy effectively guarantees privacy preservation in FL, it is not without its drawbacks, particularly in terms of privacy degradation with increasing composition time.

3. Privacy Challenges and Threats in IoT

In this segment, we undertake an examination of privacy-related challenges and threats within the realm of IoT. To facilitate this, we initiate by providing a comprehensive overview of IoT systems in order to discern potential privacy risks. The IoT architecture comprises multiple hierarchical layers, conventionally classified into four primary strata [31] (depicted in Figure 1): the Perception Layer, Network Layer, Middleware Layer, and Application Layer. The Perception Layer encompasses various devices such as sensors and actuators, which play a pivotal role in gathering information from the surrounding environment to oversee and manage physical conditions. Sensors are responsible for detecting factors like temperature, humidity, air quality, motion, and acceleration, while actuators regulate the operations of physical apparatuses, such as controlling a vehicle's acceleration. Devices situated within the Perception Layer generate copious amounts of data, which subsequently flow to the Network Layer for further processing and secure routing. Operating within the IoT infrastructure, the Network Layer undertakes data processing and transmission tasks across the system. In certain instances, a Middleware Layer serves as a bridge between the Network Layer and the Application Layer, assisting in the management of vendor-specific services and catered application requirements. Occupying the zenith of the IoT framework, the Application Layer leverages data that has undergone processing from diverse IoT devices to execute functions tailored to specific applications.

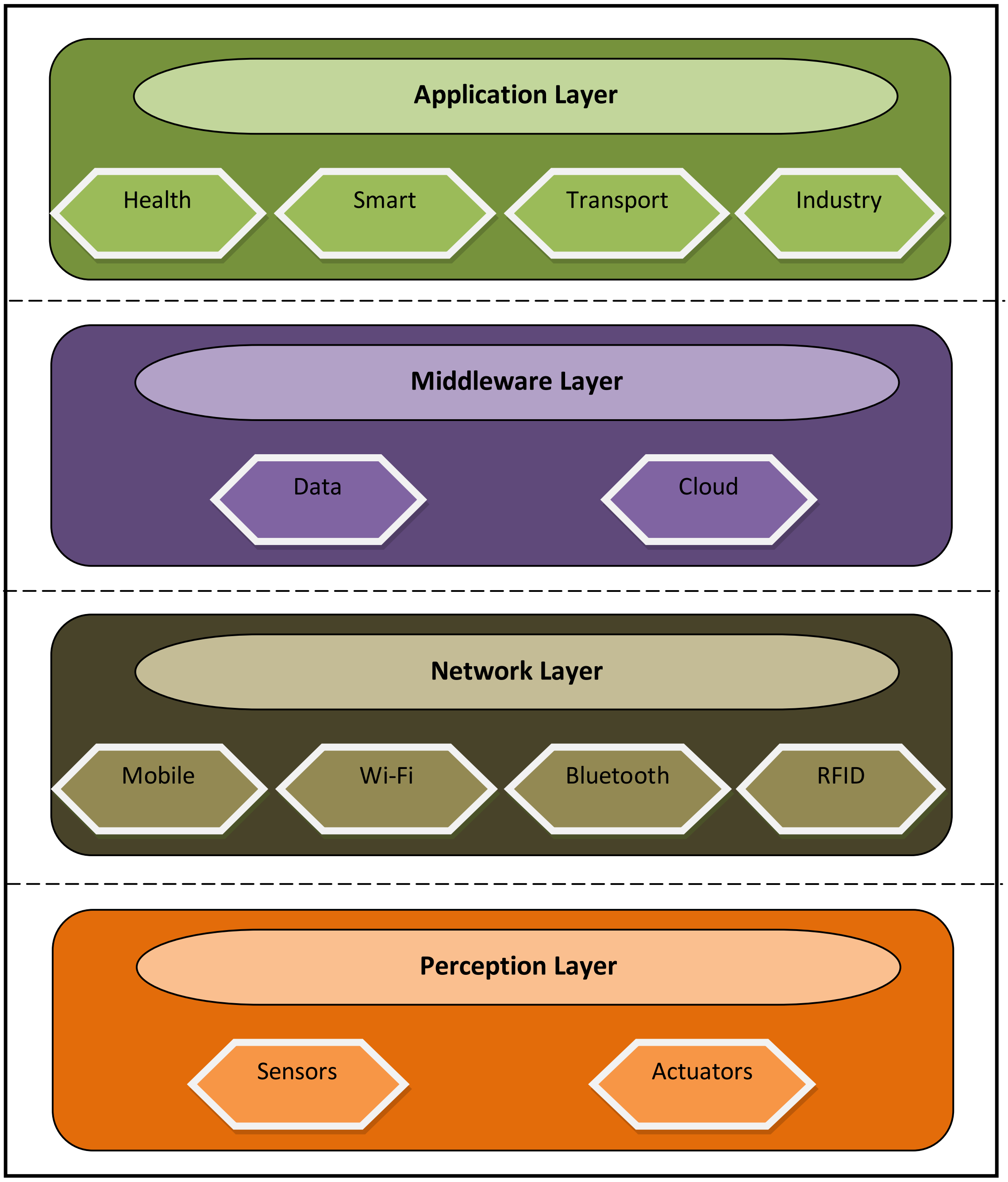

Fig.1 The IoT architecture

The substantial volume of sensitive data exchanged within the IoT infrastructure, coupled with the resource-constrained nature of IoT devices and the data processing operations executed by IoT applications, raises significant privacy apprehensions. These concerns can be categorized as follows: (1) direct threats, wherein attackers manage to access sensitive information directly, and (2) indirect threats, where attackers deduce or speculate about information without obtaining actual data access. Instances of direct information exposure can manifest across various strata of the IoT architecture. Data leakage may stem from deliberate actions (malicious users, malware, viruses, etc.) or unintentional occurrences (inadvertent disclosure of sensitive data). Notably, the Intel Security report [32] underscores that 45% of data leakage is attributed to employees, with 51% of these instances being accidental. Ill-intentioned parties exploit system vulnerabilities to bypass authentication mechanisms and gain unauthorized access to sensitive data.

Membership inference attacks concentrate on the attributes of individual data instances. In such attacks, an adversary assesses whether a given individual data record was employed in training a machine learning (ML) model. Privacy infringement occurs when the inclusion of an individual's data in the dataset is itself sensitive. For instance, if it's revealed that a personal medical record contributed to a health-related ML model, this disclosure directly jeopardizes the person's identification and health information. The pioneering membership inference attack on ML, introduced by [33], involves an attacker with black-box access to the model and utilizes confidence scores of inputs to ascertain data participation in training. More recently, Author in [34] presented more versatile attacks with reduced cost and undiminished efficacy. Similarly, [35] investigated scenarios where the attacker possesses access to both the model and average training loss. The model inversion attack [36] empowers adversaries to extract underlying participant training data by analyzing ML model predictions. Author demonstrated notable instances of these attacks. In the first case [37], they retrieved genomic data of patients by inverting a model used for medicine dosage predictions. The second instance [38] involved extracting training data instances from observed model predictions. Through confidence scores, they reconstructed images of faces resembling given examples. [39] extended this approach by proposing a model inversion attack in healthcare analytics, considering differences in the model before and after gradient updates. [40] devised an attack that recovers model parameters based on observed predictions, leveraging input/output pairs for equation solving. Model stealing attacks are geared toward recovering internal training parameters or sensitive details of ML models. As trained models are considered proprietary, particularly in the realm of Deep Neural Networks (DNNs), revealing model parameters can inadvertently lead to data disclosure. Property inference attacks [41] deduce distinct patterns of information from a target ML model, identifying sensitive patterns within the model's training data. These attacks have been applied to various models, including Hidden Markov Models (HMM) [42], Support Vector Machines (SVM) [43], and Neural Networks (NN) [44]. Re-identification attacks [45] combine data from multiple sources to re-identify a record from outsourced, public, or open data repositories. An illustrative re-identification attack exploits a voter registration list to re-identify a government health record from data released by a health insurance company. Impersonation attacks [46] involve monitoring a device's communication pattern and attempting to replicate it. The adversary gains access to the device modifies privacy preferences and ultimately inserts counterfeit data into the system.

3.1. Federated Learning for IoT security

Federated learning (FL) is centered on the notion of constructing global models through decentralized data sources. This inherent approach substantially enhances privacy levels within distributed systems. FL champions privacy by extending the training of machine learning (ML) models to Internet of Things (IoT) devices, effectively enabling these client devices to engage in model learning. Within this paradigm, dispersed nodes employ their individual data to locally train ML models, subsequently transmitting model parameters to other nodes. This orchestrated process enables clients to generalize their localized models based on the collective model parameters overseen by a central server. FL's design involves participants exclusively sharing gradients for aggregation purposes, thereby heightening the safeguarding of localized data privacy. In its application to IoT networks, initial efforts have predominantly revolved around extensive data analytics pertinent to sensor devices operating within IoT environments.

The growing significance and adaptability of IoT applications have heightened the vulnerability of IoT devices to adversarial attacks and security threats, subsequently impacting ML, FL, and DL models. These threats manipulate data inputs and tweak network parameters to yield erroneous outcomes [47]. Numerous research endeavors have explored the integration of FL as a means to devise potential solutions for enhancing IoT security. Approaches such as ensemble or adversarial training [48] have been suggested to fortify IoT systems. Nonetheless, these techniques exhibit limited scalability when applied to distributed networks and are tailored to specific attack types. FL, with its compatibility with the distributed nature of IoT networks, emerges as a potent tool for detecting a wide spectrum of security threats and attacks, thereby contributing to the development of robust defense mechanisms. Leveraging FL's privacy enhancement attributes, security frameworks are designed to empower individual IoT devices to concurrently execute neural network models. This parallel execution bolsters the security model against diverse adversaries. Integrating FL with IoT not only expedites the learning process but also accelerates attack detection while minimizing potential risks. In the intricate landscape of heterogeneous IoT learning, the integration of an attack detection module within the FL environment becomes imperative. A notable example of such a framework is presented in [49], which deploys a dynamic model for assessing aggregated parameters to mitigate attacks within the IoT-FL interaction. This study underscores the need for innovative approaches to detect and prevent various attack types through FL, as discussed in [50]. To address the challenge of detecting anomalies, an efficient anomaly detection process has been established and deployed at the global server [51]. This process identifies rare and distinct updates, eliminating malicious data instances and retaining crucial data features. The integration of FL enables the continuous monitoring and updating of models, facilitating the identification and prevention of malicious activities and unauthorized users [52]. Additionally, FL models can be trained to identify unauthorized users in dynamic IoT networks [53]. These models further expedite the identification of compromised IoT devices within FL networks.

Nonetheless, the FL-based IoT environment is confronted with communication bottlenecks that contribute to increased communication delays. Efforts to alleviate this bottleneck are explored in a comprehensive survey presented in [54]. Communication congestion may arise due to factors such as the proliferation of participating devices, restricted network bandwidth, limited computation at edge nodes, and network heterogeneity. A solution to this limitation involves updating the model, selecting clients to manage the number of participants and communication load, employing decentralized training, and embracing Peer-to-Peer learning. Addressing the communication bottleneck concern, an approach known as decentralized FL (DFL) is implemented to mitigate congestion around the central server. DFL operates in both synchronous (Sync-DFL) and asynchronous (Async-DFL) modes, with Async-DFL pioneering a generic FL framework that is asynchronous and eliminates waiting periods. This innovation facilitates efficient training of distributed models within the heterogeneous IoT landscape. In conjunction with communication bottlenecks, intermittent connectivity of IoT devices presents another challenge for FL-based IoT. This issue is tackled in [55], where a novel framework is proposed utilizing the Message Queue Telemetry Transport (MQTT) protocol and the Open Mobile Alliance (OMA) Lightweight Machine-to-Machine (LwM2M) semantics. This framework enhances the FL model's robustness while managing IoT devices and optimizing connectivity and communication efficiency. The feasibility of the proposed protocol is examined; demonstrating its capacity to improve connectivity compared to existing Proof-of-Concept (PoC) setups. The stability of the communication process is also impacted by intermittent connectivity. The challenge posed by this aspect in FL-based IoT is discussed in [55], where a novel framework harnessing the Message Queue Telemetry Transport (MQTT) protocol and the Open Mobile Alliance (OMA) Lightweight Machine-to-Machine (LwM2M) semantics is proposed. This framework enhances the robustness of FL models while managing IoT devices and optimizing connectivity and communication efficiency. The feasibility of the protocol is assessed and compared against existing Proof-of-Concept (PoC) implementations to validate its scalability. In conclusion, a concise evaluation of the FL process for IoT is outlined in Table 1.

Table 1 Evaluation of Federated Learning in IoT security

Aspect	Evaluation for Federated Learning in IoT Security
Advantages	- **Privacy Preservation:** Data remains decentralized, enhancing privacy. - **Data Efficiency:** Reduces data transmission, conserving bandwidth. - **Local Learning:** Devices learn locally, minimizing data exposure. - **Security:** Centralized data storage vulnerability is reduced. - **Scalability:** Supports large-scale IoT deployments. - **Heterogeneity:** Accommodates diverse IoT device capabilities.
Challenges	- **Communication Overhead:** Communication between devices can be taxing. - **Data Heterogeneity:** Differing data quality and formats affect learning. - **Model Aggregation:** Ensuring accurate model synthesis is challenging. - **Device Constraints:** Resource-limited devices may struggle with training. - **Security Concerns:** Ensuring device authenticity and model privacy. - **Network Latency:** Slow or intermittent connections can hinder learning.

Use Cases	- **Anomaly Detection:** Collaborative model learns to identify deviations. - **Predictive Maintenance:** Devices forecast maintenance needs collectively. - **Healthcare Monitoring:** Devices learn health trends without sharing data. - **Smart Grid Management:** Optimizing energy consumption without data sharing. - **Industrial IoT Security:** Detecting breaches and vulnerabilities collaboratively.
Implementation Steps	- **Device Selection:** Choose suitable devices based on capabilities. - **Model Design:** Design a robust model accommodating data variety. - **Initial Model Distribution:** Deploy initial model to devices. - **Local Training:** Devices train on local data with iterative updates. - **Model Aggregation:** Federated server aggregates local models into global. - **Privacy Measures:** Employ techniques like differential privacy for protection. - **Secure Communication:** Ensure encrypted communication between devices.
Benefits for IoT Security	- Enhanced security due to data decentralization and reduced exposure. - Privacy preservation through localized learning and limited data sharing. - Scalability for large IoT deployments without centralized resource strain. - Improved efficiency with reduced data transmission and processing. - Adaptability to heterogeneous IoT device landscape.

3.2. Deep Learning for IoT security

Given the rapid evolution and increased complexity of security breaches, the task of detecting such breaches using existing models, which primarily rely on detecting deviations from the normal behavior of networks and IoT devices with minimal false alarms and swift detection times, has become considerably intricate. In addition to delving into Federated Learning (FL), this section also explores the contribution of deep learning (DL) models to fortifying the security of IoT systems.

The Internet of Things is rapidly evolving into a ubiquitous computing service necessitating substantial data storage and processing capacities [56]. Regrettably, due to limitations in resources, self-organization, and the unique attributes of close-range communication within the IoT realm, there exist challenges. Employing Deep Learning (DL) methodologies presents opportunities to address a range of security concerns. Presented here is a deep learning model tailored to bolster IoT security. In summary, the investigation underscores that leveraging Deep Learning for the implementation of security measures – encompassing authentication, Hardware Device Integrity Testing, and confidentiality – in IoT devices and their inherent vulnerabilities proves advantageous.

Deep learning (DL) stands as the forefront methodology for delving into data exploration within the IoT context, discerning the patterns of 'normal' and 'abnormal' behavior among IoT components based on their interactions within the IoT ecosystem [57]. DL models possess the remarkable ability to anticipate emerging attacks by drawing insights from previous attack instances, and they exhibit an impressive aptitude for predicting new future attacks by assimilating knowledge from past occurrences. It is essential to recognize that, in light of the escalating potency and resources of attackers, traditional machine learning techniques employed for attack detection prove inadequate in detecting intricate cyber assaults. These methods falter in recognizing variations in the characteristics of threats and attacks, and they struggle to extract pertinent attributes that differentiate novel or modified attacks from innocuous events. The utilization of deep learning-powered neural networks offers a remedy to this challenge, given their proficiency in managing intricate classification and attack detection tasks [58]. A depiction of how DL models can offer potential for enhancing IoT security is presented in Figure 2.

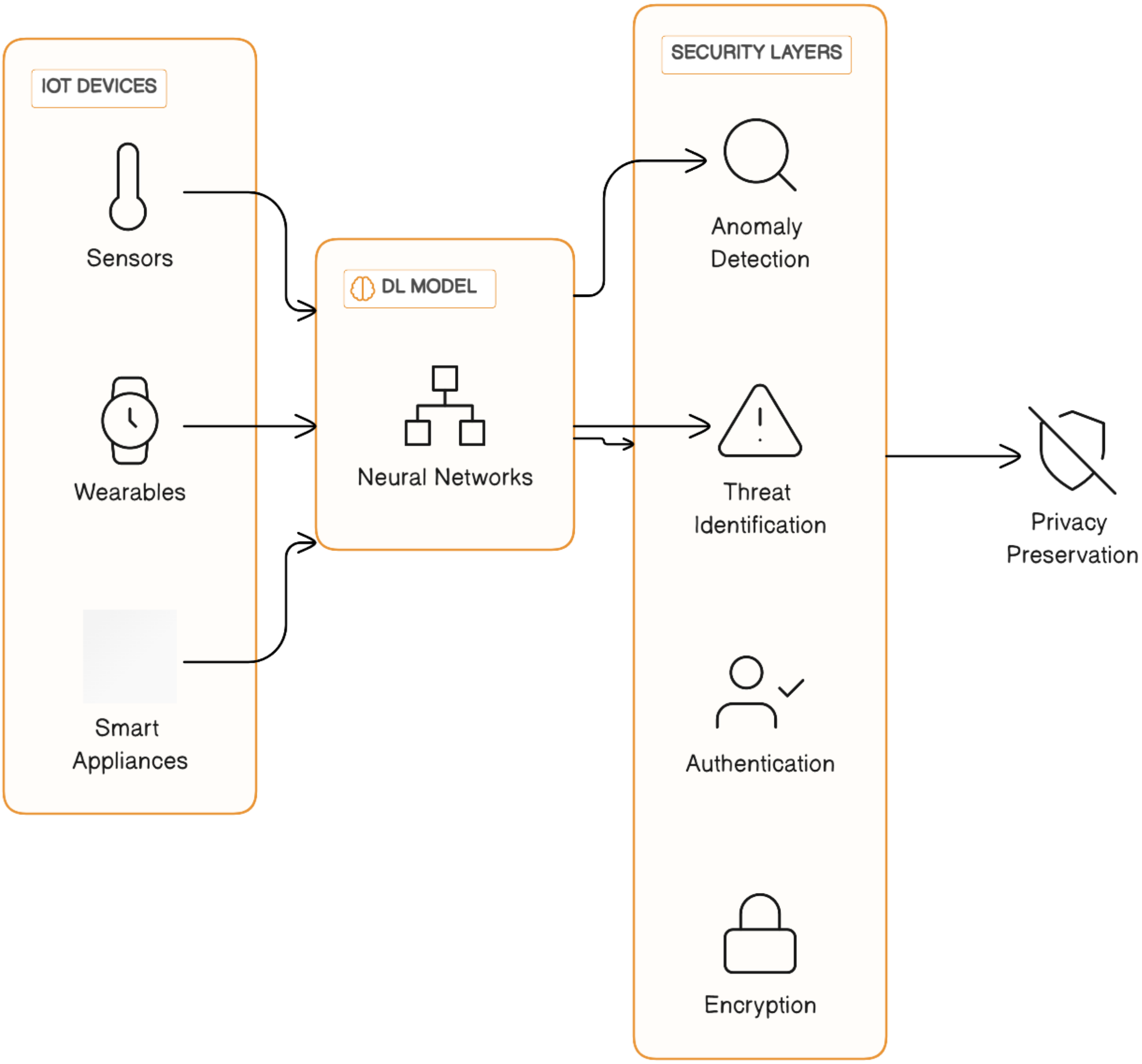

Fig.2 Deep Learning model IoT security

The workflow of the DL process in the context of attack detection is outlined in the following key points:

- Preprocessing of data: As an initial step within the attack detection process, data preprocessing assumes a crucial role. This preprocessing phase serves to mitigate uncertainties inherent in the data, encompassing external noise, missing values, null entries, redundant information, and more. The raw input data undergoes refinement to render it amenable to classification. The presence of uncertainties in the input data can significantly impact classification accuracy, underscoring the necessity to eliminate them to enhance the effectiveness of detecting and classifying attacks within IoT environments.
- Extraction of features: This step holds significant prominence within the attack detection procedure. Broadly, feature extraction entails isolating pertinent and crucial attributes from the input data. The extraction of solely significant features serves to diminish dimensionality, as numerous features may not substantially contribute to the overall efficacy of the attack detection endeavor. By curtailing unnecessary and redundant features, not only does this action streamline computational demands, but it also bolsters the overall performance of the attack detection process.
- Detection of attacks and classification: During this procedure, the extracted attributes serve as inputs for the DL model, facilitating the identification of security breaches. Subsequent to pinpointing attacks

within the input data, the classifier proceeds to categorize various attack types, such as Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) attacks, leveraging the extracted attack-associated features. The DL models undergo training to perpetually surveil the network, enabling the timely detection of anomalous shifts in network behavior. Upon recognizing such alterations, the model undertakes the classification of data into either normal or malicious categories.

- Evaluation of performance: The evaluation of DL model performance involves the assessment of various performance metrics, encompassing measures like accuracy, precision, recall, and support. All the above steps depicted in figure 3.

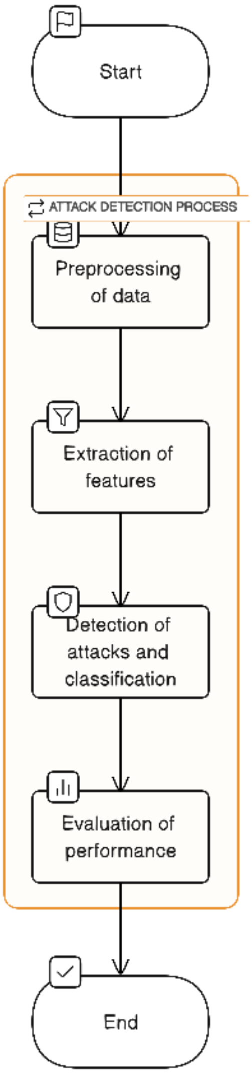

Fig.3 Workflow of DL process in the detection of attacks

4. Difficulties linked with Federated Learning (FL) and Deep Learning (DL) models.

4.1. Difficulties linked with Federated Learning (FL):

- Communication Overhead: FL involves constant communication between edge devices and a central server, which can lead to significant communication overhead, especially in large-scale and resource-constrained environments [59].
- Heterogeneity: Devices in FL may differ in terms of hardware capabilities, data distribution, and network connectivity, posing challenges in aggregating and reconciling diverse data sources.

- Privacy Concerns: While FL is designed to protect data privacy, there can still be risks of data leakage during aggregation or inference if not properly managed [60].
- Non-IID Data: FL assumes independently and identically distributed (IID) data, but real-world data is often non-IID, leading to challenges in achieving accurate model convergence.
- Model Drift: Devices in FL may have different data distributions over time, leading to model drift and degradation of global model performance.
- Secure Aggregation: Ensuring secure aggregation of model updates from multiple devices while preserving privacy is a complex challenge.

4.2. Difficulties linked with Deep Learning (DL):

- Data Quality and Quantity: DL models require large volumes of high-quality labeled data, which can be scarce or expensive to obtain, especially in niche domains.
- Overfitting: DL models are prone to overfitting, where they perform well on training data but generalize poorly to new, unseen data [61].
- Computation and Resource Intensity: Training DL models can be computationally intensive, requiring powerful hardware and significant energy consumption.
- Interpretability: DL models are often considered black-box models, making it challenging to understand their decision-making processes, especially in critical applications.
- Hyperparameter Tuning: Selecting optimal hyperparameters for DL models is a time-consuming and complex task that significantly impacts model performance.
- Transfer Learning: Adapting pre-trained DL models to new tasks can be challenging due to domain differences and the need for fine-tuning.
- Bias and Fairness: DL models can inherit biases present in training data, leading to biased predictions and potential fairness issues [62].
- Scalability: As DL models become larger and more complex, deploying and managing them at scale becomes a challenge, especially in edge and distributed computing environments.

These challenges underline the need for ongoing research and innovation to address the limitations and enhance the effectiveness of both Federated Learning and Deep Learning models in various applications.

5. Conclusion

This comprehensive paper delves into the utilization of Machine Learning (ML) algorithms, particularly emphasizing Federated Learning (FL) and Deep Learning (DL), to bolster the security of Internet of Things (IoT) environments. The examination encompasses a diverse array of FL and DL techniques, all geared toward identifying various security threats and potential attacks targeting IoT. The synergistic deployment of FL and DL holds the potential to furnish robust security measures against a multitude of malicious incursions, a feat attributed to their adeptness in accommodating the resource-constrained and heterogeneous nature intrinsic to IoT devices.

Moreover, the review intricately outlines contemporary techniques introduced within existing literature and undertakes a thorough analysis of layer-wise attacks in IoT. The discernment of these attacks proves pivotal for fortifying systems against adversarial maneuvers. Alongside, the study embarks on an exploration of diverse ML algorithms poised to counteract security breaches. Considering the attributes and functionalities intrinsic to IoT devices, the formulation of an adept security model necessitates the judicious selection of a suitable DL or FL model. Subsequent training empowers this security framework to make intelligent real-time decisions by assimilating knowledge from available instances.

In addition to its exploration of ML-based security paradigms, the paper engages with the complexities and challenges inherent in implementing such approaches within IoT systems. The challenges brought to light in this research not only underscore existing hurdles but also beckon promising avenues for future research in the realm of IoT security.

References:

[1] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, B. Sikdar (2019), A survey on IoT security: Application areas, security threats, and solution architectures, IEEE Access 7, 82721–82743.

[2] Shah, R.; Chircu, A. IoT and AI in healthcare: A systematic literature review. *Issues Inf. Syst.* 2018, *19*.

[3] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson (2018) "The industrial internet of things (iiot): An analysis framework," Computers in Industry, vol. 101, pp. 1–12.

[4] J. Zhang, Y. Wang, S. Li, S. Shi (2020), An architecture for IoT-enabled smart transportation security system: A geospatial approach, IEEE Internet Things J. 8 (8), 6205–6213.

[5] Djamel Eddine Kouicem, Abdelmadjid Bouabdallah, Hicham Lakhlef 2018, Internet of things security: A top-down survey, Computer Networks, Volume 141, Pages 199-221, ISSN 1389-1286, https://doi.org/10.1016/j.comnet.2018.03.012.

[6] J. Granjal, E. Monteiro, and J. S. Silva (2015), "Security for the internet of things: a survey of existing protocols and open research issues," IEEE Communications Surveys & Tutorials, vol. 17, no. 3, pp. 1294–1312.

[7] Q.-D. Ngo, H.-T. Nguyen, V.-H. Le, D.-H. Nguyen (2020), A survey of IoT malware and detection methods based on static features, ICT Express 6 (4), 280–286.

[8] Jaikla, Tinthid, Chalee Vorakulpipat, Ekkachan Rattanalerdnusorn and Dang-Hai Hoang (2019), "A Secure Network Architecture for Heterogeneous IoT Devices using Role-based Access Control." *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)* : 1-5.

[9] Kumari, Swati, Maninder Singh, Raman Singh and Hitesh Tewari (2021):. "To Secure the Communication in Powerful Internet of Things Using Innovative Post-Quantum Cryptographic Method." *Arabian Journal for Science and Engineering* 47, 2419 - 2434.

[10] E. Bertino and N. Islam 2017, "Botnets and internet of things security," Computer, vol. 50, no. 2, pp. 76–79.

[11] G, Sripriyanka and Anish Mahendran (2022), "Issues and Solution Techniques for IoT Security and Privacy - A Survey." *International Journal of Computing and Digital Systems* : n. pag.

[12] T. T. Huong (2022), "Federated Learning-Based Explainable Anomaly Detection for Industrial Control Systems," in *IEEE Access*, vol. 10, pp. 53854-53872, doi: 10.1109/ACCESS.2022.3173288.

[13] Cunha Neto, Helio & Hribar, Jernej & Dusparic, Ivana & Menezes, Diogo & Fernandes, Natalia. (2023). Securing Federated Learning: A Security Analysis on Applications, Attacks, Challenges, and Trends. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3269980.

[14] Viraaji Mothukuri, Reza M. Parizi, Seyedamin Pouriyeh, Yan Huang, Ali Dehghantanha, Gautam Srivastava (2021), A survey on security and privacy of federated learning, Future Generation Computer Systems, Volume 115, Pages 619-640, ISSN 0167-739X, https://doi.org/10.1016/j.future.2020.10.007.

[15] Mujaheed Abdullahi, Hitham Alhussian, Norshakirah Aziz, Said Jadid Abdulkadir, Yahia Baashar (2022), "Deep Learning Model for Cybersecurity Attack Detection in Cyber-Physical Systems", *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA*, pp.1-5.

[16] Y. Xin (2018), "Machine Learning and Deep Learning Methods for Cybersecurity," IEEE Access.

[17] Hitaj, B.; Ateniese, G.; Perez-Cruz, F. Deep models under the GAN: Information leakage from collaborative deep learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, New York, NY, USA, 3 November 2017; pp. 603–618.

[18] F. Restuccia, S. DOro, and T. Melodia, "Securing the internet of things in the age of machine learning and software-defined networking," IEEE Internet of Things Journal, vol. 5, pp. 4829–4842, Dec 2018.

[19] Ul Hassan, Muneeb & Rehmani, Mubashir Husain & Chen, Jinjun. (2019). Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions. Future Generation Computer Systems. 97. 10.1016/j.future.2019.02.060.

[20] G. Choudhary and V. Sharma, “A Survey on the Security and the Evolution of Osmotic and Catalytic Computing for 5G Networks,” in 5G Enabled Secure Wireless Networks. Springer, 2019, pp. 69–102.

[21] Torre, Damiano & Mesadieu, Frantzy & Chennamaneni, Anitha. (2023). Deep learning techniques to detect cybersecurity attacks: a systematic mapping study. Empirical Software Engineering. 28. 10.1007/s10664-023-10302-1.

[22] Akhtar, Naveed & Mian, Ajmal. (2018). Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. IEEE Access. 6. 10.1109/ACCESS.2018.2807385.

[23] Macas, Mayra & Wu, Chunming & Fuertes, Walter. (2022). A survey on deep learning for cybersecurity: Progress, challenges, and opportunities. Computer Networks. 212. 109032. 10.1016/j.comnet.2022.109032.

[24] Truong, N.; Sun, K.; Wang, S.; Guitton, F.; Guo, Y. Privacy preservation in federated learning: An insightful survey from the GDPR perspective. *Comput. Secur.* 2021, *110*, 102402.

[25] Karafiát, Bc. Vít. “Machine learning privacy : analysis and implementation of model extraction attacks.” (2021).

[26] Torre, Damiano & Chennamaneni, Anitha & Rodriguez, Alex. (2023). Privacy-Preservation Techniques for IoT Devices: A Systematic Mapping Study. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3245524.

[27] Amiri-Zarandi, M.; Dara, R.A.; Fraser, E. A survey of machine learning-based solutions to protect privacy in the Internet of Things. Comput. Secur. J. 2020, 96, 21–45.

[28] Eugenia Politou and others, Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions, Journal of Cybersecurity, Volume 4, Issue 1, 2018, tyy001, https://doi.org/10.1093/cybsec/tyy001.

[29] Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth. 2017. Practical Secure Aggregation for Privacy-Preserving Machine Learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17). Association for Computing Machinery, New York, NY, USA, 1175–1191. https://doi.org/10.1145/3133956.3133982.

[30] El Ouadrhiri, A.; Abdelhadi, A. Differential privacy for deep and federated learning: A survey. IEEE Access 2022, 10, 22359–22380.

[31] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali and M. Guizani, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," in IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 1646-1685, thirdquarter 2020, doi: 10.1109/COMST.2020.2988293.

[32] Intel Security Report. Available online:https://www.intel.com/content/dam/www/public/us/en/security-advisory/documents/intel-2020-product-security-report.pdf (accessed on 14 November 2022).

[33] L. Liu, Y. Wang, G. Liu, K. Peng and C. Wang, "Membership Inference Attacks Against Machine Learning Models via Prediction Sensitivity," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 3, pp. 2341-2347, 1 May-June 2023, doi: 10.1109/TDSC.2022.3180828.

[34] S. Truex, L. Liu, M. E. Gursoy, L. Yu and W. Wei, "Demystifying Membership Inference Attacks in Machine Learning as a Service," in IEEE Transactions on Services Computing, vol. 14, no. 6, pp. 2073-2089, 1 Nov.-Dec. 2021, doi: 10.1109/TSC.2019.2897554.

[35] M. S. Islam, B. Omidi, I. Alouani and K. N. Khasawneh, "VPP: Privacy Preserving Machine Learning via Undervolting," 2023 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), San Jose, CA, USA, 2023, pp. 315-325, doi: 10.1109/HOST55118.2023.10133266.

[36] A. Krall, D. Finke and H. Yang, "Gradient Mechanism to Preserve Differential Privacy and Deter Against Model Inversion Attacks in Healthcare Analytics," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 5714-5717, doi: 10.1109/EMBC44109.2020.9176834.

[37] Ganta, S.R., Kasiviswanathan, S.P., Smith, A.: Composition attacks and auxiliary information in data privacy. In: KDD 2008, pp. 265–273 (2008)

[38] Kasiviswanathan, Shiva & Rudelson, Mark & Smith, Adam, " The Power of Linear Reconstruction Attacks". Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (2012). 10.1137/1.9781611973105.102.

[39] A. Krall, D. Finke and H. Yang, "Gradient Mechanism to Preserve Differential Privacy and Deter Against Model Inversion Attacks in Healthcare Analytics," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 5714-5717, doi: 10.1109/EMBC44109.2020.9176834.

[40] M. Juuti, S. Szyller, S. Marchal and N. Asokan, "PRADA: Protecting Against DNN Model Stealing Attacks," 2019 IEEE European Symposium on Security and Privacy (EuroS&P), Stockholm, Sweden, 2019, pp. 512-527, doi: 10.1109/EuroSP.2019.00044.

[41] Xiaoyu Zhang, Chao Chen, Yi Xie, Xiaofeng Chen, Jun Zhang, Yang Xiang, A survey on privacy inference attacks and defenses in cloud-based Deep Neural Network, Computer Standards & Interfaces, Volume 83, 2023, 103672, ISSN 0920-5489.

[42] Tun Li, Yutian Liu, Yanbing Liu, Yunpeng Xiao, Nang (2020), An Nguyen, Attack plan recognition using hidden Markov and probabilistic inference, Computers & Security, Volume 97, 101974, ISSN 0167-4048.

[43] B. Dakhale, K. Vipinkumar, K. Narotham, S. Pungati, A. A. Bhurane and A. G. Kothari (2023), "Analysis of Adversarial Attacks on Support Vector Machine," 2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS), Nagpur, India, pp. 1-5, doi: 10.1109/PCEMS58491.2023.10136124.

[44] K. Ganju, Q. Wang, W. Yang, C. A. Gunter, and N. Borisov (2018), “Property inference attacks on fully connected neural networks using permutation invariant representations,” in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 619–633, ACM.

[45] Narayanan, A.; Huey, J.; Felten, E.W (2016), A Precautionary Approach to Big Data Privacy. In Data Protection on the Move; Gutwirth, S., Leenes, R., De Hert, P., Eds.; Springer: Dordrecht, The Netherlands, pp. 357–385.

[46] Yan X, Yan K, Rehman M.U, Ullah S (2022) Impersonation Attack Detection in Mobile Edge Computing by Levering SARSA Technique in Physical Layer Security. Appl. Sci, 12, 10225. https://doi.org/10.3390/app122010225.

[47] G. Abad (2023), "Sniper Backdoor: Single Client Targeted Backdoor Attack in Federated Learning," in 2023 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML), Raleigh, NC, USA, pp. 377-391. doi: 10.1109/SaTML54575.2023.00033.

[48] Fatimah Aloraini, Amir Javed, Omer Rana, Pete Burnap, Adversarial machine learning in IoT from an insider point of view, Journal of Information Security and Applications, Volume 70, 2022, 103341, ISSN 2214-2126, https://doi.org/10.1016/j.jisa.2022.103341.

[49] Y. Fraboni, R. Vidal, M. Lorenzi (2021), Free-rider attacks on model aggregation in federated learning, in: International Conference on Artificial Intelligence and Statistics, PMLR, pp. 1846–1854.

[50] Linhao Meng, Yating Wei, Rusheng Pan, Shuyue Zhou, Jianwei Zhang, and Wei Chen (December 2021), VADAF: Visualization for Abnormal Client Detection and Analysis in Federated Learning. ACM Trans. Interact. Intell. Syst. 11, 3–4, Article 26, 23 pages. https://doi.org/10.1145/3426866

[51] A. Gerard, R. Latif, S. Latif, W. Iqbal, T. Saba and N. Gerard (2020), "MAD-Malicious Activity Detection Framework in Federated Cloud Computing," 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, pp. 273-278, doi: 10.1109/DeSE51703.2020.9450728.

[52] E. Rizk, S. Vlaski and A. H. Sayed (2020), "Dynamic Federated Learning," 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Atlanta, GA, USA, pp. 1-5, doi: 10.1109/SPAWC48557.2020.9154327.

[53] Pouriyeh S, Shahid O, Parizi R.M, Sheng Q.Z, Srivastava G, Zhao L, Nasajpour M (2022), Secure Smart Communication Efficiency in Federated Learning: Achievements and Challenges. Appl. Sci, 12, 8980. https://doi.org/10.3390/app12188980.

[54] P. Pinyoanuntapong, W. H. Huff, M. Lee, C. Chen and P. Wang (2022), "Toward Scalable and Robust AIoT via Decentralized Federated Learning," in IEEE Internet of Things Magazine, vol. 5, no. 1, pp. 30-35, doi: 10.1109/IOTM.006.2100216.

[55] Genovese G., Singh G., Campolo C., Molinaro A (2022), Enabling edge-based federated learning through MQTT and OMA lightweight-M2M, in: IEEE VTC Spring.

[56] N. V. Rajeesh Kumar and M (2022). Arun, "Deep Learning Model to Improve Security in IOT Systems," 2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), Villupuram, India, pp. 1-5, doi: 10.1109/ICSTSN53084.2022.9761347.

[57] M. Hassaan Khalid (2023), "A Brief Overview of Deep Learning Approaches for IoT Security," 2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, pp. 1-5, doi: 10.1109/iCoMET57998.2023.10099306.

[58] Mirdula S, Roopa M (2023), MUD enabled deep learning framework for anomaly detection in IoT integrated smart building, e-Prime - Advances in Electrical Engineering, Electronics and Energy, Volume 5, 100186, ISSN 2772-6711, https://doi.org/10.1016/j.prime.2023.100186.

[59] Divya Jatain, Vikram Singh, Naveen Dahiya (2022), A contemplative perspective on federated machine learning: Taxonomy, threats & vulnerability assessment and challenges, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 9, Pages 6681-6698, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2021.05.016.

[60] Nuria Rodríguez-Barroso, Daniel Jiménez-López, M. Victoria Luzón, Francisco Herrera, Eugenio Martínez-Cámara (2023), Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges, Information Fusion, Volume 90, Pages 148-173, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2022.09.011.

[61] V. K. BP, K. SM and P. LV (2023), "Deep machine learning based Usage Pattern and Application classifier in Network Traffic for Anomaly Detection," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, pp. 50-54, doi: 10.1109/ICAECIS58353.2023.10169914.

[62] K.C. Ravikumar, Pandi Chiranjeevi, N. Manikanda Devarajan, Chamandeep Kaur, Ahmed I. Taloba (2022), Challenges in internet of things towards the security using deep learning techniques, Measurement: Sensors, Volume 24, 100473, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100473.