Review

A Survey on Privacy Preservation Techniques in IoT Systems

Rupinder Kaur *, Tiago Rodrigues, Nourin Kadir, Rasha Kashef

Posted Date: 13 March 2025

doi: 10.20944/preprints202503.0979.v1

Keywords: IoT; privacy; privacy preservation; security threats; sensors; privacy threats; edge nodes; IoMT; industrial IoT

 doi:10.20944/preprints202503.0979.v1

Review

A Survey on Privacy Preservation Techniques in IoT Systems

Rupinder Kaur *, Tiago Rodrigues, Nourin Kadir and Rasha Kashef

Electrical, Computer and Biomedical Engineering Department, Toronto Metropolitan University, 350 Victoria St, Toronto, M5B2K3, Ontario, Canada

* Correspondence: rupinder.kaur.ece@torontomu.ca

Abstract: We are living in the era of IoT systems where this technology supports our daily life in various aspects, whether it is our daily used Apple watches, our smart home systems, or our laptops. The goal is to make our everyday activities more convenient and more accessible by means of connectivity anytime, anywhere. The growth of IoT-based systems booms in the post-2020s with advanced network technologies like 5G and Edge Computing. The IoT-based systems are comparatively cost-effective and convenient to apply, especially in remote operations where manual intervention is impossible. Like other technologies, IoT applications also come with flaws and concerns. One of the main severe concerns nowadays is privacy and security issues related to IoT systems. Knowingly or unknowingly, we are being monitored by smart sensors or edge devices almost every moment, and our personal and professional sensitive information is being exposed to untrusted third parties like Google or Amazon. This is becoming a serious concern with the rapid expansion of IoT-based systems. In this survey paper, we have tried to categorically list and present the state-of-the-art techniques for privacy preservation in IoT-based systems in various application fields. Our work is a summary based on the 39 papers and two online reports that we chose to analyze to understand the current situation, privacy attacks, how to handle the sensitive data of the clients without breaching privacy and future directives in this fast-growing IoT-based systems.

Keywords: IoT; privacy; privacy preservation; security threats; sensors; privacy threats; edge nodes; IoMT; industrial IoT

1. Introduction

Who could have imagined that after 1990, the world would experience a technology boom in IoT-based systems within only three decades of the Internet explosion [1–10]? The rise of the Internet has not only connected individuals but also machines, sensors, and devices, creating a highly interconnected ecosystem. Today, more than 15 billion active IoT devices are operating worldwide, whereas the global population currently stands at 7.8 billion. This means roughly every individual has two IoT devices of some form. This number is expected to double by 2030, a staggering prediction that highlights the rapid expansion of IoT systems [11–18]. Figure 1 shows the exponential growth of active IoT devices across the globe.

 doi:10.20944/preprints202503.0979.v1

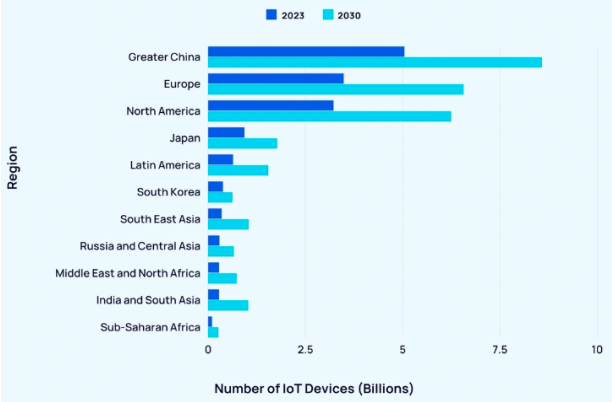

Figure 1. Number of active devices by region [1].

The increasing penetration of IoT technologies into nearly every aspect of our lives signals an ever-growing dependence on IoT-based systems. These systems range from home automation and healthcare applications to industrial control and environmental monitoring, each connected via the internet. As the reach of IoT continues to grow, so does the volume and sensitivity of the data generated by IoT devices. It is estimated that the number of connected devices will continue to rise, amplifying the challenges of managing and securing the massive amounts of data they generate. The coverage area of IoT-based systems is vast, as nearly any physical object can be transformed into an IoT-enabled device. This ecosystem is driving technological advancements that connect not only urban areas but also remote regions that were previously underserved.

However, with the proliferation of IoT devices and the increasing integration of IoT systems into critical infrastructures, a pressing concern has emerged—data security and privacy. Specifically, the vast amount of data transmitted from IoT edge nodes or sensors to cloud storage, often involving untrusted third parties, presents significant security risks. These data transfers, whether in the transmission phase or while being stored, are susceptible to breaches of privacy. As IoT systems continue to evolve, the need for robust security measures becomes more urgent to prevent unauthorized access to sensitive information [19–28].

Various academic efforts have been dedicated to addressing these security challenges, and significant progress has been made in securing IoT networks, enhancing data protection mechanisms, and developing privacy-preserving technologies [29–40]. Furthermore, emerging trends like edge computing, blockchain, and artificial intelligence are being explored as potential solutions to safeguard IoT systems and enhance data privacy and security [41–53]. Nevertheless, despite the ongoing advancements in IoT security, there remains a continuous need for innovative approaches that can mitigate the risks associated with data privacy breaches, especially considering the increasing complexity of modern IoT networks.

This paper explores the evolving landscape of IoT security, focusing on privacy-preserving mechanisms that aim to protect sensitive data as it is transmitted across networks and stored in cloud-based systems. By reviewing current research and evaluating state-of-the-art technologies, we seek to provide insights into the future direction of IoT security in light of its growing importance across various domains, including healthcare, industrial automation, and smart cities.In this survey paper, we have investigated recent research papers discussing privacy preservation in IoT systems in various application fields. In IoT-based systems, edge devices and sensors collect or gather information from the ambience related to our personal and professional behaviours and preferences because of the pervasive nature of the systems. When the collected data is transferred to cloud storage for further processing, sensitive information like the name, address, location, image or even personal security PIN or password can be exposed to untrusted third parties. In May 2023, Meta was fined US$ 1.3 billion by the European Union for leaking user's personal information across the Atlantic.

 doi:10.20944/preprints202503.0979.v1

Figure 2. presents the security threats and cyber-attacks specifically targeting IoT devices. The trend is up and growing. So, privacy preservation is a must for IoT-based systems.

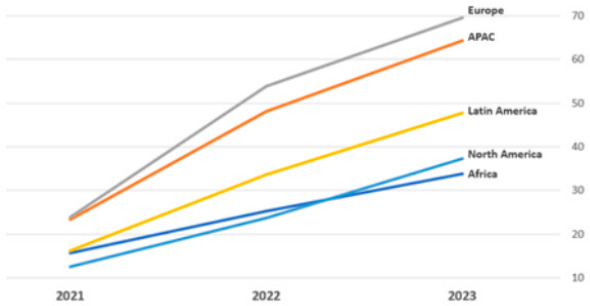

Figure 2. Cyber attacks trend targeting IoT devices across all regions [2].

We have investigated the various privacy preservation techniques in IoT-based systems and organized the rest of this paper as follows: section II. Literature review and background, section III. Data sources and types, section IV. Analytical models used V. discussion and section VI. conclusion and future directions.

2. Literature Review and Background

2.1. Research Questions

Much academic research has been published on privacy preservation techniques in IoT-based systems. We have read some of those works and categorically listed the taxonomy for privacy preservation. We have set the following research questions for our survey paper, and the answers are found at the end of our study based on the selected papers [14]:

- What are the privacy preservation techniques used in the published works?
- What type of IoT edge devices are used and affected by the applied techniques?
- What are the privacy attacks and threats that have been studied and mitigated?

All the research papers that have been used in our survey were published between 2016 to 2023. The papers have been published on various platforms like IEEE, Springer Link, MDPI, IET, ELSEVIER, and ACM.

2.2. Related Work

We have selected the research papers for our study so that different emerging fields depending on state-of-the-art IoT systems can be covered for our analysis. Xue, Wanli, et al. [3] privacy-preserving face recognition technique using Bloom filter space for storage and analytics tasks. Evaluation of BFace is done by using the practical dataset (YaleB). The proposed technique shows the trade-off between utility and privacy for practical scenarios. BFace is a lightweight and less complex technique where false positives are possible, but false negatives are not. The baseline accuracy for SVM on the Yale B dataset is 80%, while BFace has a better classification accuracy of 92%. Privacy-preserving stochastic gradient descent-based algorithms and Private Aggregation of Teacher Ensembles (PATE) have been used as privacy-preserving-based learning techniques.

Peterson Yuhala [4] focuses on enhancing the security and privacy of IoT-based systems, while hardware peripheral devices are isolated in his work. The paper presents a Trusted Execution Environment (TEE) and leveraged machine learning classification methods for filtering out the sensitive data, e.g., speech, images, etc., from associated peripheral devices to prevent the unwilling or unaware data leakage to the untrusted parties in the cloud like Amazon or Google. However, there is a trade-off between the performance and privacy of the IoT-based systems for implementing the proposed model. OPP-115 Corpus dataset is used to evaluate the performance of different machine

 doi:10.20944/preprints202503.0979.v1

learning algorithms to classify the privacy policies of websites in the published work of Carson, DiSalvo and Ray [5]. If the ML algorithm is trained with annotated privacy policies of websites, what is its performance for privacy policies' classification of IoT devices. The work is based on "Towards Automatic Classification of Privacy Policy Text," which indicates that segment-based classification is effective for automated classification in privacy policies. The proposed model is 70% accurate on the OPP-115 Corpus dataset using the bag of words method.

Elkahlout, Mohammed, et al. [6] present an elaborate study on the existing IoT-based healthcare and monitoring systems for elderly people. Also, it shows how vulnerable the system is when it comes to security and privacy while handling the patient's sensitive information, which involves the personal identity and medical condition of the patient. Two of the strongest threats in IoT-based healthcare systems are selective forwarding (SF) and wormhole (WH). The problem is severe because most elderly people are unaware of the current technological advancements, and hence, they have minimum experience with related security issues. One of the optimal solutions for the privacy and security threat is to use blockchain technology where the decentralization principle is applied, and even if any one of the node's securities is breached, the original dataset is available in all other nodes in the network.

Meenakshiamma, Bharathi and Krishnakumar [7] propose a novel technique to detect breast cancer using an IoT system while ensuring the privacy of the patient's personal and medical data. When the collected data is transferred to untrusted third parties like cloud storage, the model encrypts the data using the AES algorithm. Then, the triple DES encryption follows a double DES encryption to ensure no private data is leaked over the transmission. At the receiver end, the data is decrypted using a secret key and then re-encrypted by a different method. After the procedure, the data is submitted for prediction analysis for early detection of breast cancer. The authors use the Wisconsin Diagnostic Breast Cancer (WDBC) dataset to check the performance of their proposed model, which has 98.5% accuracy in CNN classification and 99.2% accuracy in ANN classification. However, CNN classifiers create a considerable volume of synthetic data during the training phase. This requires ample storage for the data. The proposed IoMT architecture needs extensive hardware resources, though it may be mitigated by using cloud computing services. Schiliro, Moustafa and Beheshti [8] presents an LSTM-based cognitive privacy technique that ensures data privacy of EEG signals along with user classifications and their events. The data privacy is achieved through a normalized correlation analysis model to prevent unauthorized access. The proposed model performs highly on the PhysioNet BCI dataset for user classification with data privacy. The model is validated against four other techniques: Mahalanobis Distance-Based Classification model (SEDP), Magnitude Squared Coherence with a K-Nearest Neighbor algorithm (MSCKNN), Mahalanobis Distance with Spectral Coherence Features (SCCDP) and Similar Distance-Based Classification technique along with Alpha-Delta Bands Power features (ADPDP). The authors conclude that AI-enabled cybersecurity technology can protect individuals' cognitive information and prevent violations of cognitive privacy.

Fazeldehkordi, Owe and Noll [9] present a framework built on a security classification method in the work. The authors use a case study featuring a pacemaker functioning as a medical device communicating with its environment. The paper investigates various security and privacy challenges associated with two distinct scenarios. The analysis results reveal that after implementing the proposed framework, there is an upgrade in the security classification of both the sensor and the pacemaker, advancing from class F to C and class D to A, respectively. Overall security improves significantly. The main objective of the paper is to offer the functionalities of IoT privacy and security.

Gochoo, Munkhjargal, et al. [10] present a novel IoT-based human posture recognition system which is device-free and privacy-preserving. The model consists of a low-resolution infrared sensor based on three WSN nodes and a server for classifying posture images by Deep Convolutional Neural Network (DNCC). The data set has 93200 posture images of 26 postures. The classifier is evaluated with the tenfold cross-validation technique. The average F1 score for using three WSN nodes is 0.9989, while the average F1 score for using a single WSN node is 0.9854. The proposed model is a

 doi:10.20944/preprints202503.0979.v1

novel approach to device-free privacy-preserving sensing technology. Jui, Tania Tahmina, et al. [11] investigate different combinations of data preprocessing, feature selections and classification algorithms to detect intrusion in IoT-based networks while improving accuracy and reducing time. They use MQTT-IoT_IDS-2020, which is the most recent IoT-specific network intrusion detection dataset and NSL-KDD, which is one of the most popular benchmark datasets for traditional network traffic for the experiment. For the MQTT-IoT_IDS-2020 dataset, the achieved accuracy is 99.86% in 2.81 seconds while using normalization for data preprocessing, AdaBoost and Best First Search for feature selection and J48 for classification algorithm. For the NSL-KDD dataset, the achieved accuracy is 84.35% in 0.32 seconds using oversampling for data preprocessing, Bagging and Generic Search for feature selection and Bagging for classification algorithms. The authors propose a novel model to detect intrusion in the IoT network by combining data preprocessing, feature selection and classification algorithms that have not been practiced before.

Fazeldehkordi, Owe and Noll [12] give a comprehensive presentation of the state-of-the-art regarding security and privacy functionalities and requirements in IoT networks. Fagbohungbe, Omobayode, et al. [13] present a novel framework for implementing efficient privacy-preserving edge intelligent computing where training the autoencoder at each edge node is flexible. At the same time, compressed and encrypted raw data is transmitted to the edge server, maintaining privacy and security. The classifier on the server is trained with the less controlled features provided by the autoencoder. Thus, by decoupling the training of the autoencoder and the edge server, the classifier lessens the frequent communication between them and yields improved privacy and security. In future, the authors want to expand the experiments for classifier performance with communication cost and model complexity for image classification by federated learning and SplitNN.

Alotaibi et al. [15] propose a stacked deep learning architecture from five pretrained residual networks for cyberattack detection in the IoT field. This approach can distinguish between suspicious and normal traffic activities. Each of the five pretrained residual networks has ten residual blocks with the same settings. The authors test the architecture in two heterogeneous environments, such as the smart home and smart grid. The dataset of the smart home is publicly available by the ICS Cyberattack Dataset collection consisting of 15 subdatasets. In comparison, the smart grid dataset is composed of three subdatasets from the IoT Botnet Attack dataset. This method can differentiate between regular and malicious traffic activities effectively, thus detecting possible IoT attacks. This high predictive performance results in better findings than related approaches by instantaneously providing a higher detection rate than existing classification models. In future, the authors want to further improve the current accuracy and detection time by altering the number of pretrained residual networks and using alternative state-of-the-art pretrained models.

Moving users' health records to the cloud benefits healthcare practitioners as it allows for on-demand self-service data availability. Such infrastructure requires both authentication and access control to keep data confidential. These health records might include allergy information, blood tests, dental information, and health information gathered from smartwatches. Thus, kahani et al. [16] propose a zero-knowledge protocol combined with a two-stage keyes access control to protect patient's and medical centers' data in the e-health systems. The zero-knowledge protocol verifies and maintains the anonymity of the user's identity. At the same time, the access control is based on scalable fine-grained access responsible for decisions such that a dental surgeon will not have access to a user's blood records but will have access to a patient's dental x-rays. Other related works have not considered the need for managing data access policies, have intense computational overhead, have compromised data files or require users to be online, while the proposed model is a one-stop for providing privacy in e-health services.

Meisami et al. [17] propose a blockchain-based protocol for e-health approaches that does not use a trusted third party and incorporates an efficient privacy-preserving access control method. Wearable devices that automatically calculate a patient's vital information, like heart rate, require a central storage or cloud infrastructure to save such data for further analysis by machine learning or medical staff for treatment. Meisami et al. [17] thus propose a model in which the patient's data is

 doi:10.20944/preprints202503.0979.v1

protected by privacy and confidentiality. One related work applies anonymity to the patient's data, but recent research has shown that the data can still be decrypted with little information. Other related work involves a differential privacy method that adds noise during the computational process, but it is not sufficient and efficient for healthcare applications. Another related work uses an encryption algorithm named fully homomorphic encryption but is found not suitable for practical application.

Another privacy problem is caused by traffic inference attacks on IoT devices, which employ packet sizes and information regarding timing. Many questions arise for the effective solution. Firstly, it is essential to identify the type of privacy required. Secondly, the design of new protocols and standards should be such that the effect of different types of IoT traffic is deeply considered. These questions are addressed in [18], which introduces an event-level differential privacy (DP) model applied over infinite packet streams. A novel shaping mechanism is designed under the recognized traffic and privacy models. Also, a constrained problem is formulated to reduce the expected delay for a packet and identify the ideal shapers for specified stages of privacy efficiency. The experimental results in [18] showcase privacy-overhead trade-offs, i.e., better privacy is achieved for a larger shaping overhead.

Ref. [19] focuses on the need for security and privacy in healthcare systems. The authors present a secure and protected health monitoring system that can operate remotely. The distributed architecture of blockchains is used instead of a centralized communication system. The architectural design of the system is shown in Figure 3. The design considers security, scalability, and processing time as some of the essential characteristics. The security is ensured by employing the re-encryption proxy with Blockchain for data encryption. To guarantee scalability, an InterPlanetary file system (IPFS) off-chain database is used for data storage. Ethereum Blockchain, based on proof of authority (PoA), works to speed up the process.

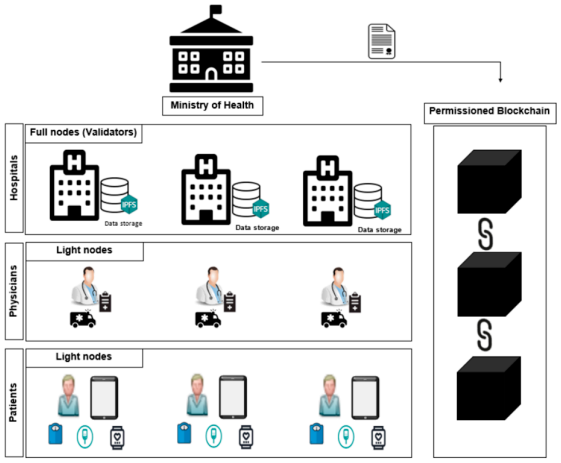

Figure 3. Health Monitoring System using Blockchains presented in [19].

Though many privacy-preserving data aggregation techniques have been proposed in past studies, most are restricted to homogenous data or definite functions. [20] emphasizes the necessity of a publicly available tool to ensure the privacy of heterogeneous data. A new technique to provide user and data privacy is proposed, which can be adopted for data aggregation in fog and cloud settings. This technique benefits from a Trusted Execution Environment (TEE) to confirm privacy and process heterogeneous data.

Smart cars are a classic example of IoT applications, but they come with many risks involved with security and privacy. Remote vehicle hijacks and road accidents are some of these issues. Blockchain is a promising candidate to cater to security problems in vehicular IoT. [21] presented a methodology based on blockchain for secure vehicular IoT. This technique puts the vehicle data on the server, but the associated hash values are on the blockchain. Therefore, security is maintained by

 doi:10.20944/preprints202503.0979.v1

the decentralization of data. Blockchain utilizes peer-to-peer computing to deliver a safer platform for information sharing. Figure 4. reflects the methodology in a flow chart.

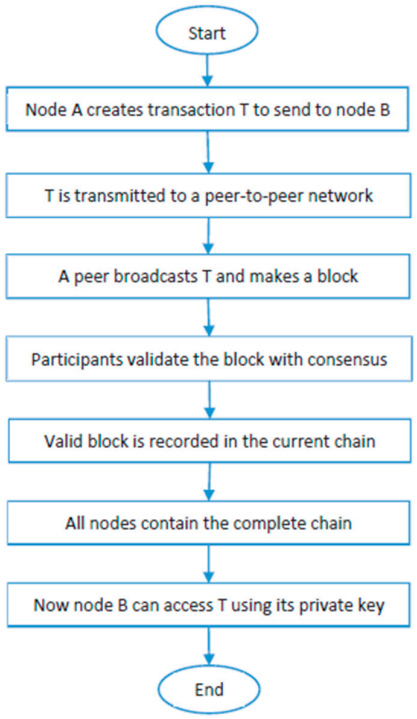

Figure 4. Secure Vehicular IoT using blockchains proposed in [21].

Another paper [22] also outlines blockchain technology's ability to ensure security and privacy. The authors also provide an impression of cache memory, which is based on IoT. Also, an architecture of a single-bit six-transistor static random-access memory cell sense amplifier is presented. Each chip's memory is recorded as a block. It is encrypted and used as a blockchain further. Power optimization techniques are applied to different parts of architecture. The results illustrate a trade-off between power and area. Another article [23] utilizes blockchain to safeguard the privacy of federated learning (FL). The paper presents a novel framework with two models: one is private peer-to-peer computing, and the other is private federated learning. Ethereum and interplanar file systems (IPFS) are used to build these models. The article also outlines the realization of automatic on-chain identification and autonomous FL when the technique is united with radio frequency identification (RFID) technology.

Ref. [24] carries out a detailed review to investigate security issues in IoT. The authors provide an observable report of the possible security threats, their solutions based on blockchain, security features and challenges associated with Blockchain in IoT. However, this review concludes that Blockchain technology is one of the hopeful technologies that offer many benefits to enhance the security and privacy of IoT data. It also identifies the issues to reflect that the use of this technology in IoT still demands enough research to overcome the related challenges and complexities.

Ref. [25] aims to design a personal assistant which can cater to the queries about Privacy by Design (PbD) methods in the designing stage of IoT systems. Semantic web technologies are utilized to demonstrate PbD measurements, their relationships with privacy patterns and the requirements of IoT systems. The study also delivers a modelled agenda of the framed questions and their actions that bring added concerns around the associated IoT systems.

Ref. [26] proposes an effective privacy-ensured healthcare finance system that is based on blockchain technology and is appropriate for lightweight computer devices. The proposed strategy brings into light noninteractive zero-knowledge proof. This technique effectively reduces the cost of communication. The authentication of the transactions with this method takes milliseconds. The

 doi:10.20944/preprints202503.0979.v1

entire system and its use in a healthcare financial system is clarified deeply at a micro-level. The benefits of this technique can be applied to other general financial systems. To lessen privacy concerns in securing sensitive financial data, a noninteractive version of ZKPs using the Fiat–Shamir heuristic is employed [26]. It mainly reduces the communication cost. Numerousl ZKPs, namely range-proof, proof of asset, proof of balance, range-proof, proof of consistency, etc., are used to safeguard the correctness of transactions.

Another paper [27] proposes a distributed platform to maintain privacy in the healthcare system. It establishes a medium between the blockchain system and systems like healthcare centers. Along with it, numerous smart contracts are employed for the registration of users, generation of certificates and guaranteeing access to the users. There is no definite need for the maintenance of a certificate. Any IoT device can be utilized to maintain certificates by the user. In this process, users first get a unique ID after registration. To obtain any healthcare certificate, the officials verify the user's data. Finally, the required certificate is issued using blockchain technology.

The Industrial Internet of Things (IIoT) faces numerous data privacy and security challenges in the processing of data collection in industrial applications. Old-style single-factor authentication does not offer flexibility with an increasing number of users and various other categories. [28] provides an artificial intelligence-based model to preserve privacy in IIOT. The design follows two stages, sanitization, and restoration of the data. Data sanitization aims to secure the information with the Grasshopper-Black Hole Optimization(G-BHO) algorithm. The efficiency of the sanitization stage is supported by optimal key generation through the algorithm proposed in [28].

There can be a scenario when data privacy is affected by a nasty node sending private data accessing requests to other nodes from an edge-computing-based IoT cloud storage system. [29] manages to solve this issue by developing a signaling game to strengthen privacy preservation. A payoff matrix represents the relationship between the edge nodes and IoT devices. The two parties in this signaling game adjust their strategies where IoT devices tend to carry out nasty privacy requests, and the edge nodes can reject them. A signalling Q-learning algorithm is utilized in the paper to optimize evolutionary learning, and the simulation results confirm its perfection.

Federated learning (FL) is considered a pioneering artificial intelligence methodology that can be a problem-solver in machine learning training by training larger data. [30] inspects federated learning approach based on methodology, application and system stages, real-life applications, and process backgrounds. FL results in highly secure privacy strategies as it generates vigorous classifiers that do not require information revelation. FL is the first choice of researchers to work with a secure environment at the forefront. It offers a secure model to share prototypes to different roles if the users have inadequate data labels. FL can provide high shared and federated security for a broad spectrum of applications, thus supporting the expansion of artificial intelligence.

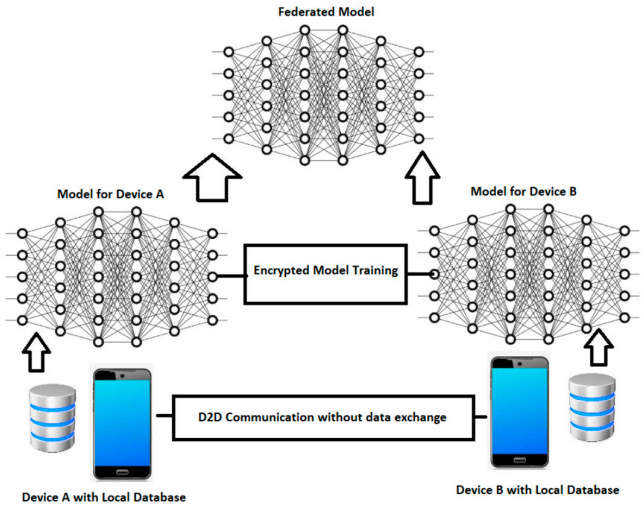

Figure 5. D2D Communication without data exchange in FL [30].

 doi:10.20944/preprints202503.0979.v1

As the things on The Internet of Things (IoT) are self-configuring and autonomous, there is a broad scope of privacy issues. [31] classifies the privacy and security factors in IoT systems, focusing on (a) privacy-preserving models, (b) available solutions, and (c) reference of privacy models for various IoT applications. Figure 6. depicts a balanced measure of privacy in the different IoT-based layers. The review summarizes those newer methodologies such as Blockchain, Machine Learning, Data Minimization, and Data Encryption have significant effects on privacy and security concerns. Also, the study supports that the lesser the data, the better the protection. So, this paper also presents a machine learning-based data minimization methodology, which proves to be helpful for privacy-preserving.

Ref. [32] presents an effective method to improve privacy preservation in IoT-based healthcare applications by utilizing homomorphic encryption techniques and blockchain technology. Homomorphic encryption supports efficient calculations in data encryptions without decrypting the data. The data can be further used after encryption without sharing the actual data. In addition to it, this paper includes smart contracts in the blockchain network and incorporates access control.

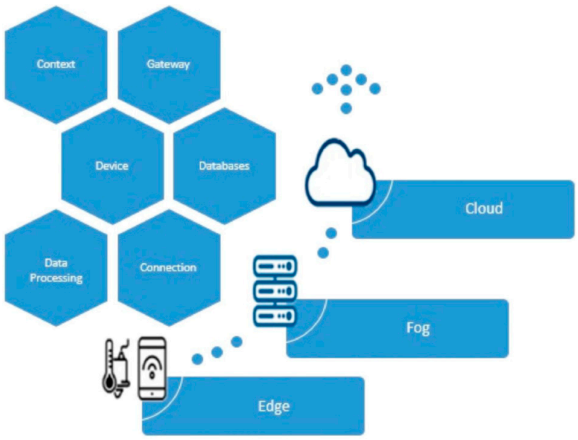

Figure 6. Three-tier layers of IoT architecture presented in [31].

Other related work mentions propose blockchains, hashing algorithms, key encryptions, and machine learning for privacy in IoT. [33] propose PriModChain, trained on the MNIST dataset to enforce privacy and trustworthiness on IIoT systems, [34] propose OPECA that aids in transmitting information over antennas of IoT devices without eavesdropper obtaining information. [35] introduce a PPCDR algorithm for IoT devices in the industrial field capable of avoiding privacy leakage from directly sharing information. [36] research into various directions of encryption protocols, software protection and sensor data loss. [37] proposed FedL, a lightweight computational algorithm for privacy-preserving in IoT devices. [38] research into the current state of security and privacy of the IoT in the healthcare system. [39] generate privacy fingerprints extracted from sensitive information in mobile networks to be quantified for analysis. [40] introduces CASE, a context-aware attribute learning scheme to automatically learn, exploit and generate attributes to reduce the post-encryption data volume. [41] propose HASHA, a hashed-based protocol for location privacy, particularly in IoT wireless sensor networks, through address anonymity.

3. Data Sources and Types

In this section, we are going to discuss in detail the data sources and data types used in some of the selected research papers that have been studied for our work—and a tabular summary of the data sources and types used in those papers for the survey. We elaborate on the used datasets and their applications for the developed models. Jui, Tania Tahmina, et al. [11] have used two major datasets for their proposed intrusion detection model in IoT-based networks. The used datasets are MQTT-IoT-IDS-2020 and NSL-KDD, which are benchmark datasets for network traffic. First, the authors have applied some preprocessing techniques to simplify the dataset, and after preprocessing, they

 doi:10.20944/preprints202503.0979.v1

have extracted important features through feature selection techniques. In the next step, with the reduced features, different classification algorithms have been applied. Accuracy and time efficiency are the performance measuring parameters for the applied classification algorithms for finding out the best result for combining preprocessing techniques with feature selection techniques and classification algorithms. The main goal of the work is to detect the intrusion in the network that can be seen as privacy preservation of the IoT-based systems. Figure 7. presents the flow chart for the proposed model.

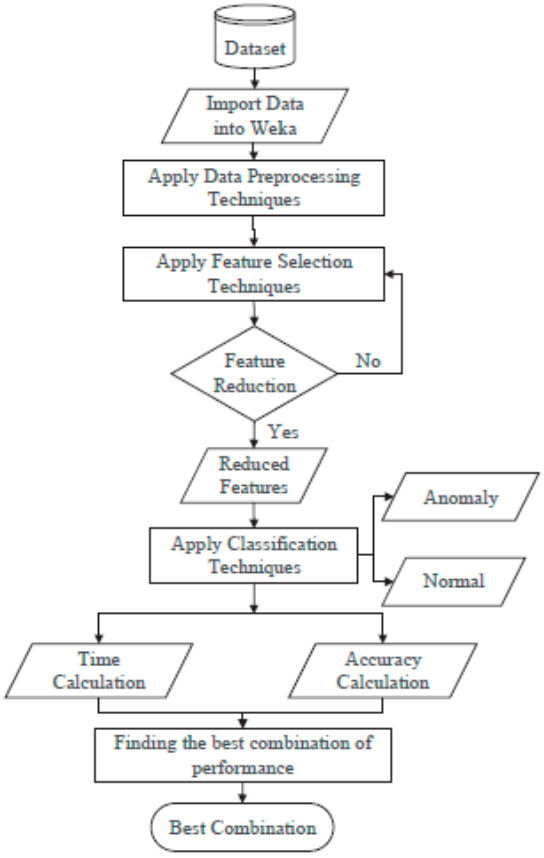

Figure 7. Flow chart of the proposed model for feature reduction through data preprocessing for intrusion detection in IoT networks [11].

The MQTT-IoT-IDS-2020 dataset is used to detect intrusion in the IoT networks, where the dataset is built based on the Message Queueing Telemetry Transport (MQTT) protocol. This dataset has three feature levels, namely Packet-based features, unidirectional-flow-level-based features, and Bidirectional-flow-based features. It has been shown that to distinguish MQTT-based attacks from benign traffic, it is essential to use the Packet-based features. The MQTT-IoT-IDS-2020 comprises five different kinds of normal records and four different kinds of attack scenarios. The four attack types are Aggressive scan (Scan-A), User Datagram Protocol (UDP) scan (Scan-sU), Sparta SSH brute-force (Sparta), MQTT brute-force attack (MQTTBF). There are two different types of features in the MQTT-IoT-IDS-2020 dataset [11].

NSL-KDD is a superset of the "KDD Cup'99" dataset, which is taken from the University of New Brunswick. It is a benchmark dataset for contemporary internet traffic. It combines four datasets, namely KDDTrain+, KDDTrain+_20Percent, KDDTest+, and KDDTest-21Percent, where KDDTest-21Percent is a subset of KDDTrain+. Like the MQTT-IoT-IDS-2020 dataset, the NSL-KDD dataset also has four different kinds of attack scenarios: Denial of Service (DoS), Probe, User to Root(U2R), and Remote to Local (R2L). The NSL-KDD dataset has a total of 43 features. Among those, 41 present the

 doi:10.20944/preprints202503.0979.v1

traffic input directly, and the other two are labels and scores. The label indicates if the attack is normal or not, and the score indicates the severity of the traffic input. TABLE I and TABLE II present the data distributions in MQTT-IoT-IDS-2020 and NSL-KDD [11].

Table 1. Distributions of data in MQTT-IoT-IDS-2020 [11].

Dataset	Total	Normal	Scan-A	Scan-sU	Sparta	MQTT-BF
Packet	259379	86008	25693	39664	91318	16696
Uniflow	49525	171836	51358	56845	182407	33079
Bi flow	259379	86008	25693	39664	91318	16696

Table 2. Distributions of data in NSL-KDD [11].

Dataset	Total	Normal	DOS	Probe	U2R	R2L
KDDTrain+	125973	67343	45927	11656	52	995
KDDTest+	22544	9711	7458	2421	200	2654

Alotaibi et al. [15] use two datasets corresponding to two heterogeneous environments. The first dataset is publicly available and is part of the ICS Cyberattack Dataset collection that represents smart home IoT devices, such as security sensors, alarms, cameras, thermostats, and solar panels. Alotaibi et al. [15] focus on the binary-class subdataset composed of 15 subdatasets with 78,377 samples, of which 22,714 represent normal traffic activities, while 55,663 represent malicious traffic activities. These datasets have a feature size of 128. Each instance of the binary dataset is classified as a regular or malicious event. The second dataset is also publicly available and is part of the IoT Botnet Attack dataset (N-BaIoT) collection that represents nine IoT devices, such as baby monitors, security cameras, doorbells, and thermostats. Each of the nine datasets is classified as either an attack or benign class. These attacks are generated using two botnets. Only three datasets are used, including the Ecobee thermostat, Ennio doorbell and Samsung SNH 1011 N webcam, with 1,566,598 samples, of which 104,363 represent the benign class, while 1,462,235 describe the attack class. For all subdatasets, we split the data into 70% for training and 30% for testing.

Kahani et al. [16] did not use any dataset. Instead, data is generated when testing the implementation to represent the patients' accounts and health records. Meisami et al. [17] did not use any dataset as well—instead, a discussion of the theoretical aspects of the model, including its security and privacy features.

Table 3 illustrates the various datasets that have been utilized in the research papers that are surveyed as part of this review.

Table 3. Summary table of datasets used in the research papers for this survey.

Paper	Datasets Used
[3]	YaleB dataset.
[4]	Speech, Images collected from the peripheral devices. OPP-115 Corpus dataset.
[6]	NA (the paper is a literature survey).
[7]	Wisconsin Diagnostic Breast Cancer (WDBC) dataset.
[8]	PhysioNet BCI dataset (EEG).
[9]	Data was collected from the pacemaker used in the patient.
[10]	Dataset has 93200 posture images of 26 yoga postures.
[11]	MQTT-IoT_IDS-2020 dataset and NSL-KDD dataset.
[12]	Data was collected from the pacemaker used in the patient.
[13]	Images collected from the edge devices.
[14]	NA (the paper is a literature survey).
[15]	ICS Cyberattack Dataset and IoT Botnet Attack Dataset.
[16]	Custom data generated by authors.
[17]	No data was used.
[18]	Synthetic data from sleep monitor, nest camera, and WeMo switch.

 doi:10.20944/preprints202503.0979.v1

[19]	Blood Glucose, weight and health assessment from smart phone and raspberry pi.
[20]	No dataset was used.
[21]	No dataset was used.
[22]	No dataset was used.
[23]	No dataset was used.
[24]	NA (the paper is a review).
[25]	Workshop data for six realistic IoT cases with location, personal information, and photos.
[26]	No dataset was used.
[27]	No dataset was used.
[28]	Three datasets were collected: a) Student activities time in the lab at the University of Genoa, b) Electrical values from household, and c) recorded gas r, humidity, and temperature sensor data.
[29]	No dataset was used.
[30]	No dataset was used.
[31]	NA (the paper is a review).
[32]	No dataset was used.
[33]	MNIST dataset.
[34]	No dataset was used.
[35]	Amazon dataset (2018).
[36]	NA (the paper is a literature survey).
[37]	MNIST dataset.
[38]	NA (the paper is a literature survey).
[39]	Private dataset of a three-day IoT traffic communication from a mobile network operator in China.
[40]	UCI activity data.
[41]	A private dataset was used.

4. Analytical Models Used

In this section, we will discuss the analytical models used in some of the selected research papers studied for our work—and a tabular summary of the analytical models used in those papers for the survey.

Alotaibi et al. [15] propose a stacked deep learning architecture from five pretrained residual networks (ResNet) for cyberattack detection against IoT devices. Each pretraines ResNet model is made with ten ResNet blocks with two convolutional layers with the same settings assigned. After the ResNet blocks, the data is transferred to the new meta-algorithm, consisting of two dense layers (40 and 20 neurons), followed by softmax to compute the class score. Each of the two convolutional layers comprises 16 convolutional filters that output 16 feature maps, where the activation function ReLu is used. Additionally, average pooling of size two with striding two is used after the last ResNet block to reduce the feature size as it had superior results to max pooling. The authors use cross entropy as the loss function when training and the Adam optimizer because of the advantages of both RMSProp and AdaGrad optimizers.

Kahani et al. [16] propose an authentication and access control manager (AAM) server responsible for both the authentication and access control. Authentication is achieved with the Schnorr zero-knowledge identification protocol, which is used as a challenge-response protocol to authenticate the identity of the healthcare practitioner anonymously. To establish secure communication between multiple actors, a combination of public and private keys generated by the Derive Unique Key per Transaction (DUKPT) scheme is used, which alters the session key each new session to strengthen the security of the communication. Access control is defined by an intention tree, where different tree nodes represent a hierarchical relationship-like structure that represents various healthcare professions. The result of the tree is the minimum rights that a healthcare provider needs to satisfy the data request or will be denied if outside his rights. Figure 8. summarizes the data access control procedure that a user, the healthcare provider, needs to follow to retrieve a patient's data. Firstly, the user registers with the service provider where the ID is generated. Secondly, the user selects a Base Derivation Key (BDKAM) to create a session key and sends it encrypted alongside some key parameters to the AAM, responsible for the authentication and access control, where the AAM

 doi:10.20944/preprints202503.0979.v1

responds with an encrypted random number and timestamp. Thirdly, the user combines his ID, the random number and previous key parameters to calculate y, as seen in Equation 1 [16], which is transmitted to AAM.

$$y = ID^{*}e + r^{*}modq \quad (1)$$

After being authenticated and granted access, the AAM responds with a validation token. Finally, the user uses the token to retrieve the patient's encrypted data from the cloud server, where the user can decrypt it with the shared key. This protocol is fundamental to the proposed solution of Kahani et al. [16], which manages secured data sharing for e-health services.

Meisami et al. [17] propose a blockchain-based protocol for e-health approaches that does not use a trusted third party and incorporates an efficient privacy-preserving access control method. The proposed model architecture, Figure 9., shows four main modules. Wearable IoT devices and patients' phones are responsible for gathering and temporarily storing patient data. IoT devices typically have low storing space and computational power, while phones have larger storing space and computational speeds and can transmit data through wireless communications. Medical staff are the physicians and nurses who want to download the patient's data for analysis and treatment. The blockchain module stores the pointers to the data and not the actual data. It also stores access policies and removes the need for a trusted third party. The off-chain storage module is responsible for storing the patient's encrypted data. It uses an Interplanetary File System for the peer-to-peer distributed file system.

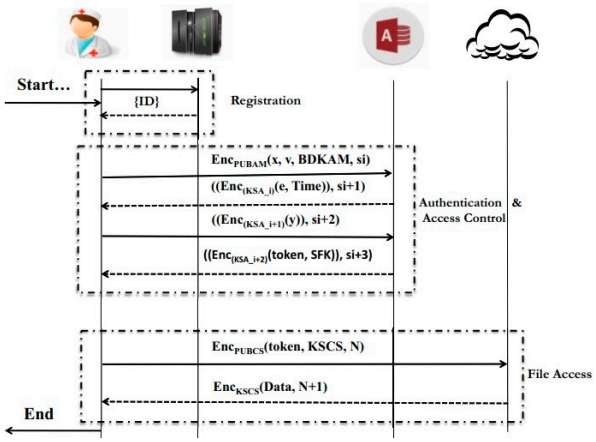

Figure 8. Data access protocol [16].

Throughout the protocol, Meisami et al. [17] use three cryptographic functions: the SHA-256 as the hash function, AES for the symmetric key encryption, and ECDSA with a secp256k1 curve for the digital signature algorithm. Initially, a staff m and patient p generate private and public keys to sign and send transactions to the blockchain server and a secret key for the AES encryption, as seen in Figure 10. Secondly, the patient registers the data access permissions by assigning policies that indicate what permission message a medical staff has over the patient, as seen in Figure 11. These permission policies can later be changed or revoked. After the patient's data is saved on the off-chain storage, it can be accessed by the medical practitioner by first checking if the staff has permission access. This can be seen in Figure 12. If the staff is granted access permissions, the patient's encrypted data can be downloaded with the following protocol, as seen in Figure 13.

 doi:10.20944/preprints202503.0979.v1

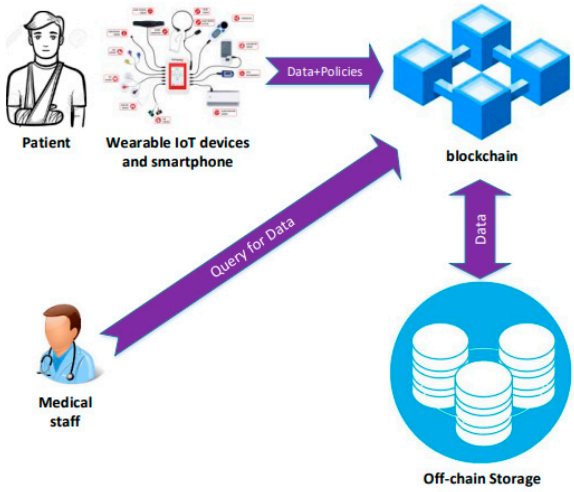

Figure 9. Proposed model architecture in [17].

Protocol 1 Joining the Blockchain

procedure GENERATING(p,m)
  p executes:
    $(pk^{p}_{sig}, sk^{p}_{sig}) \leftarrow G_{sig}()$
    $sk^{p,m}_{enc} \leftarrow G_{enc}()$
  p shares pk^{p}_{sig} (as address) on the network
  m executes:
    $(pk^{m}_{sig}, sk^{m}_{sig}) \leftarrow G_{sig}()$
  m shares pk^{m}_{sig} (as address) on the network
  p shares $sk^{p,m}_{enc}$ with m from secure channel
  // Both p and m have $sk^{p,m}_{enc}$
  return $pk^{p}_{sig}, pk^{m}_{sig}, sk^{p,m}_{enc}$
end procedure

Figure 10. Protocol used to generate secret keys in [17].

Protocol 2 Access Control Transaction

procedure ACCESSTX(pk^{k}_{sig}, $message$)
  $s \leftarrow 0$
  $(pk^{p}_{sig} \,||\, pk^{m}_{sig} \,||\, POLICY_{p,m}) = message$
  if $pk^{k}_{sig} = pk^{p}_{sig}$ **then**
    $L[H(pk^{k}_{sig})] \leftarrow L[H(pk^{k}_{sig})] \cup message$
    // L is Blockchain memory
    $s \leftarrow 1$
  end if
  return s
end procedure

Figure 11. Protocol used to set access policies [17].

Protocol 3 Blockchain Permissions Checking

procedure POLICYCHECK(pk^{k}_{sig}, T) //T=type of data
  $s \leftarrow 0$
  if $L[H(pk^{k}_{sig})] \neq \emptyset$ **then**
    $(pk^{p}_{sig} \,||\, pk^{m}_{sig} \,||\, POLICY_{p,m}) \leftarrow L[H(pk^{k}_{sig})]$
    if $pk^{k}_{sig}=pk^{p}_{sig}$ **or**
      $(pk^{k}_{sig}=pk^{m}_{sig}$ and $T \in POLICY_{p,m})$ **then**
        $s \leftarrow 1$
    end if
  end if
  return s
end procedure

 doi:10.20944/preprints202503.0979.v1

Figure 12. Protocol used to check access permissions [17].

Protocol 4 Data Transaction

procedure DATATX(pk^k_{sig} ,*message*)
  (C || T || RW) = message
  // C=encrypted data , T=type of data
  //RW=read data(=1) or write data(=0)
  if POLICYCHECK(pk^k_{sig}, T)=True **then**
    $(pk^p_{sig} || pk^m_{sig} || POLICY_{p,m}) \leftarrow L[H(pk^k_{sig})]$
      if $RW = 0$ **then**
      $L[pk^k_{sig} || T] \leftarrow L[pk^k_{sig} || T] \cup H(C)$
      (IPFS) ds[H(C)] $\leftarrow C$
        return *H(C)*
      else if $C \in L[pk^k_{sig} || T]$ **then**
        (IPFS) **return** *ds[H(C)]*
    end if
  end if
  return Ø
end procedure

Figure 13. Protocol used for data download [17].

Table 4 displays the analytical models used in the research papers that are surveyed as part of this review.

Table 4. Summary table of analytical models used in the research papers for this survey.

Paper	Analytical Models Used
[3]	SVM, BFace, Privacy-preserving stochastic gradient descent-based algorithms, Private Aggregation of Teacher Ensembles (PATE).
[4]	Trusted Execution Environment (TEE) and leveraged machine learning classification methods for filtering.
[5]	Segment-based classification, Bag of words method.
[6]	NA (the paper is a literature survey).
[7]	CNN classification and ANN classification.
[8]	LSTM-EEG, Mahalanobis distance-based classification model (SEDP), magnitude squared coherence with the k-nearest neighbour algorithm (MSCKNN), Mahalanobis distance with spectral coherence features (SCCDP) and similar distance-based classification technique along with alpha-delta bands power features (ADPDP).
[9]	The paper presents a new taxonomy framework organizing all aspects of security and privacy baselines, guidelines, and recommendations.
[10]	Deep Convolutional Neural Network (DNCC).
[11]	Normalization, SMOTE, Under Sampling for data preprocessing, Best First Search and Genetic Search for feature reduction, Naïve Bayes, AdaBoost, Bagging, J48 and Random Forest for classification.
[12]	The paper presents a new taxonomy framework organizing all aspects of security and privacy baselines, guidelines, and recommendations.
[13]	Autoencoder model for image dataset at each edge device and CNN Model for latent variables at the edge server.
[14]	NA (the paper is a literature survey).
[15]	Stacked deep learning architecture of five pretrained ResNet and meta-algorithm.
[16]	Authentication and access control manager (AAM) protocol.
[17]	Blockchain-based protocol for an efficient privacy-preserving access control method.
[18]	Event-level Adversary model, Traffic shaping Mechanisms.
[19]	Blockchain technology, Ethereum platform, Proof of Authority (PoA), InterPlanetary File System (IPFS).
[20]	Data Aggregation, Trusted Execution Environment (TEE) namely Intel SGX.
[21]	Blockchain technology, Ethereum platform, Smart Contracts.
[22]	Blockchain, IoT-based single-bit cache memory.
[23]	Federated Learning, Ethereum and Interplanar File Systems (IPFS).
[24]	NA (the paper is a literature survey).

 doi:10.20944/preprints202503.0979.v1

[25]	PARROT ontology, Privacy by Design (PbD).
[26]	Blockchain, Raft Consensus algorithm.
[27]	Distributed application (DA) based on blockchain.
[28]	Hybrid met-heuristic algorithm, Grasshopper-Black Hole Optimization.
[29]	Edge-based computing, Signaling Q- learning game technique.
[30]	Edge-based IoT system, Evolutionary game-based model.
[31]	NA (the paper is a literature survey).
[32]	The fused Heuristic approach is termed the Elliptic Curve Cryptography Blockchain model.
[33]	Privacy-preserving trustworthy machine learning model training and sharing framework based on blockchain (PriModChain).
[34]	OPCECA.
[35]	Privacy-Preserving Cross-Domain Recommendation (PPCDR) algorithm.
[36]	The literature survey research directions for encryption protocols, software protection and sensor data loss.
[37]	Lightweight privacy-preserving federated learning (FedL).
[38]	The literature survey researched the current state of security and privacy of the IoT in the healthcare system.
[39]	The model generates privacy fingerprints from the IoT traffic flow, traffic block generation and sensitive marker selection.
[40]	Context-aware attribute learning scheme (CASE).
[41]	The model HASHA is proposed, which uses hash functions.

5. Discussions

This section comprehends the results and findings of the selected research articles for our survey. We have already discussed the data sources and types in section III. The analytical model used by Jui, Tania Tahmina, et al. [11] is shown in Figure 14. The experiment for the proposed model has been performed by following several steps.

5.1. Data Preprocessing

Data preprocessing is the first step in the experiment to make the used datasets simplified and consistent. For preprocessing, the authors have applied different techniques, which are given below.

- Normalization technique transforms the dataset's features into the same scale. The process changes the numeric columns into a common range without value distortion. But normalization is not applied to the class feature.
- Synthetic Minority Oversampling Technique (SMOTE) is used for the synthetic data generation from the minority class. The oversampling technique creates new instances which continuously produce anomalous data for the dataset randomization.
- Sampling technique is used to delete the instances from the minority class. Overall, the data preprocessing technique deletes examples from the minority class for dataset balancing.

 doi:10.20944/preprints202503.0979.v1

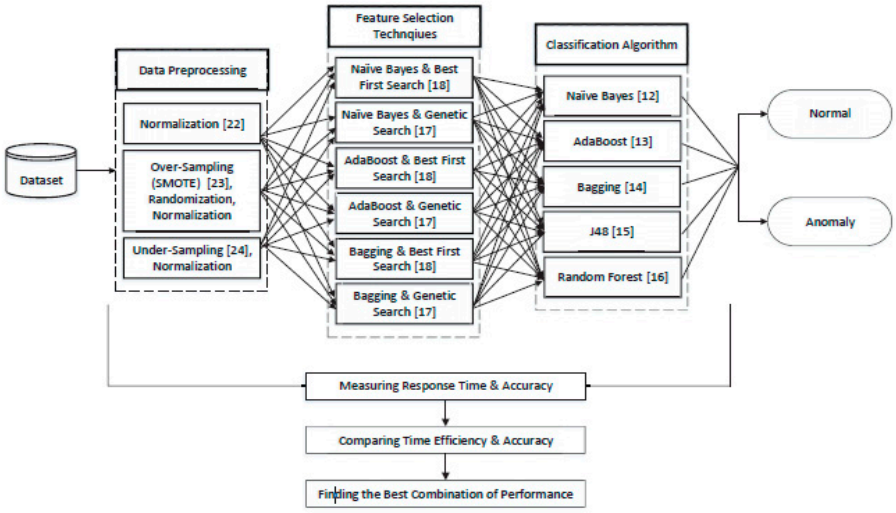

Figure 14. Analytical model for combining data preprocessing techniques, feature reduction techniques, and classification algorithms [11].

5.2. Feature Reduction

The feature reduction process ensures core feature selection and redundancy reduction. Random and irrelevant features can be misleading for the developed model. So, feature reduction is significant before model evaluation. Jui, Tania Tahmina, et al. [11] have used three feature selection algorithms, along with two searching techniques. Naive Bayes, Bagging, and AdaBoost are the feature selection algorithms used in the experiment for performance evaluation. Best First Search and Genetic Search are the search techniques the authors use here. Those features have been kept that contain all the characteristics of the dataset and the rest of the features have been removed.

5.3. Classification Models

The classification models perform training and testing on the datasets. The considered problem [11] is a binary classification problem. Five different classification models have been used in the experiment, which are described below.

- Naïve Bayes algorithm is used, which is a supervised machine learning technique.
- AdaBoost algorithm is used, a boosting technique in machine learning applied as an ensemble method.
- Bagging algorithm is used, which is an ensemble learning technique. Bagging is used to reduce variance where noise is present in the dataset. In MQTT-IoT-IDS, some noise level is present, which is why the Bagging algorithm is used in the experiment.
- J48 algorithm is used, which is a tree-based machine learning algorithm. It uses entropy for building decision trees. A decision tree-based algorithm can detect missing values.
- Random Forst algorithm is used, a supervised machine learning algorithm for accurate and stable results. The main advantage of this algorithm is that it can be applied to classification and regression problems.

Several experiments have been conducted to determine the best combinations of preprocessing techniques, feature selection techniques and classification algorithms. Based on accuracy and execution timing for model building, classification algorithms have been selected. Figure 15. Shows the data types of features in both MQTT-IoT-IDS-2020 and NSL-KDD datasets.

 doi:10.20944/preprints202503.0979.v1

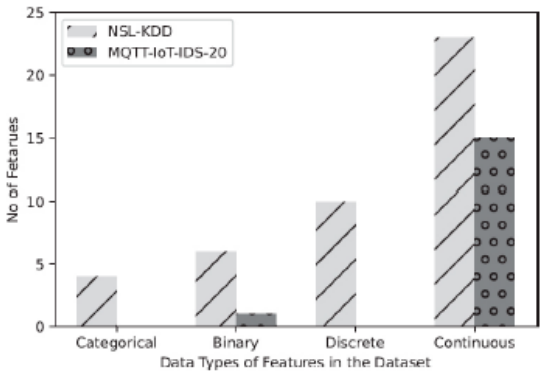

Figure 15. Data types of feature types in both MQTT-IoT-IDS-2020 and NSL-KDD datasets [11].

Table 5 and Table 6 present the result summary of feature selection techniques on MQTT-IoT-IDS-2020 and NSL-KDD datasets, respectively. Table 7 and Table 8 present the result summary of MQTT-IoT-IDS-2020 and NSL-KDD datasets, respectively, for the best combination of preprocessing techniques, feature selection techniques and classification algorithms regarding accuracy and execution time for model building. For all cases, normalized (Set A), Over-Sampled (Set B) and Under-Sampled (Set C) datasets are used.

Table 5. Summary of results of feature selection techniques on MQTT-IoT-IDS-2020 for various preprocessing techniques [11].

Feature Selection	Normalized Data (Set A)	Over-Sampled Data (Set B)	Under-Sampled Data (Set C)
Naive Bayes and Best search	18.50% (3 out of 16)	37.50% (6 out of 16)	31.25% (5 out of 16)
Naive Bayes and Genetic Search	37.50% (6 out of 16)	43.75% (7 out of 16)	37.50% (6 out of 16)
AdaBoost and Best First Search	25.00% (4 out of 16)	25.00% (4 out of 16)	25.00% (4 out of 16)
AdaBoost and Genetic Search	25.00% (4 out of 16)	Memory Exceed	31.25% (5 out of 16)
Bagging and best first Search	56.25% (9 out of 16)	Memory Exceed	43.75% (7 out of 16)
Bagging and Genetic Search	Memory Exceed	Memory Exceed	50.00% (8 out of 16)

 doi:10.20944/preprints202503.0979.v1

Table 6. Summary of results of feature selection techniques on NSL-KDD for various preprocessing techniques [11].

Feature Selection	Normalized Data (Set A)	Over-Sampled Data (Set B)	Under-Sampled Data (Set C)
Naive Bayes and Best search	06.98% (3 out of 43)	06.98% (3 out of 43)	16.287% (7 out of 43)
Naive Bayes and Genetic Search	39.53% (17 out of 43)	34.88% (15 out of 43)	37.21% (16 out of 43)
AdaBoost and Best First Search	06.98% (3 out of 43)	06.00% (3 out of 43)	09.327% (4 out of 43)
AdaBoost and Genetic Search	44.20%	23.67% (10 out of 43)	32.567% (14 out of 43)
Bagging and best first Search	Memory Exceed	Memory Exceed	Memory Exceed
Bagging and Genetic Search	Memory Exceed	60.47% (26 out of 43)	Memory Exceed

Table 7. Summary of results of MQTT-IoT-IDS-2020 [11].

Preprocessing	Feature Selection	Classifier	Accuracy	Time (sec)
Normalization	AdaBoost & Best First Search	J48	99.86	2.81
Over-Sampling & Normalization	AdaBoost & Best First Search	J48	99.85	5.40
Under-Sampling & Normalization	Naïve Bayes & Genetic Search	J48	99.82	4.00

Table 8. Summary of results of NSL-KDD [11].

Preprocessing	Feature Selection	Classifier	Accuracy	Time (sec)
Normalization	AdaBoost & Genetic Search	Bagging	84.21	0.79
Over-Sampling & Normalization	Bagging & Best First Search	Bagging	84.35	0.32
Under-Sampling & Normalization	AdaBoost & Genetic Search	Bagging	83.80	0.36

Alotaibi et al. [15] use a stacked deep learning architecture made with five pretrained ResNet models, where each model has 10 ResNet blocks followed by two convolutional layers. Following the ResNet models, average pooling is used before the two dense layers (40 and 20 neurons) before the

 doi:10.20944/preprints202503.0979.v1

Softmax classification function. This architecture is trained on two datasets, 70% for training and 30% for testing, the ICS Cyberattack Dataset and the IoT Botnet Attack dataset, where all subdatasets had a binary-class representation of benign or malicious traffic activities. The authors' architecture outperforms the other current methods when testing on the ICS Cyberattack Dataset containing 15 binary class subdatasets with an average accuracy of 97.5%, and the second most accurate method is the random forest with an accuracy of 94%. The proposed model also outperforms the IoT Botnet Attack dataset containing three subdatasets with an average accuracy of 100%, followed by the random forest model with 99.9991% accuracy. Aside from the high accuracy, the stacked deep learning model achieves a low test time of just 1.49 ms per packet. A strength of the proposed model is the better predictive performance achieved than related approaches with high accuracy. The first weakness is the lack of mentioned performance metrics thus, a balanced or overfitted performance is unknown, which brings it to the second weakness of the top four models, achieving an accuracy of over 99.98% using the N-BaIoT dataset. Such high accuracy could suggest that the dataset may be relatively easy to learn or have well-defined patterns, or the proposed model is overfitted and thus achieves the perfect accuracy.

Kahani et al. [16] propose an authentication and access control manager (AAM) server responsible for the authentication and access control of patient's healthcare data in the e-health systems. Kahani et al. [16] developed the prototype in Python and JSON format to transmit messages between the entities. The entities include the user (healthcare practitioners), cloud server (holds encrypted healthcare records), service provider (administrates access policies), and AAM server (responsible for the authentication and access control), executed on three Amazon EC2 instances. Kahani et al. [16] solution kept users' data confidential against hackers in an encrypted, semi-trusted cloud server. Operating the instances on a cloud server allows for decreasing computational and communication overhead on the data owners, making it friendly for anyone to use and implement. Experimental results show that 100 concurrent authentication requests can be made with a response time of 2.16 seconds. Secondly, the AAM server can manage 13,984 transactions over 5 minutes while responding to 4,742 requests simultaneously. The response time is, on average, composed of 54% for the validation operation, 40% for the request operation, and 6% for the network latency. While these results are reasonable, higher results can be achieved by operating the instances on better virtual machines. Overall, the approach successfully resists common network security attacks, making it ideal for the IoT cloud infrastructure where customer's data must be kept encrypted and confidential to resist network security attacks. It can withstand replay attacks through the use of sequence numbers and timestamps, resist man-in-the-middle attacks by employing public key and session key encryption, withstand brute force attacks thanks to the session key changes with each new session, and withstand denial of service attacks with the use of timestamps, requiring the user to respond within a specified time frame.

Meisami et al. [17] propose a model that protects sensitive data collected from patients' IoT devices or phones, which are not limited to tracking the heart rate, blood pressure and body temperature, using a blockchain architecture in conjunction with an access control without utilizing a trusted third party. This model consists of wearable IoT devices and smartphones, medical staff, blockchain and off-chain storage. The authors did not test the model with any data types but discussed the theoretical aspects. This model achieves confidentiality, integrity and authentication. Confidentiality is achieved by using symmetric key encryption to transmit messages. Integrity is achieved by using a hashing algorithm in the blockchain. By using a digital signature, the model achieved authentication. One limitation of the proposed model is the unavailable result outcomes. Thus, the computational overhead is unknown. A strength is the high resilience to various attacks in the IoT environment, such as denial of service attacks, modification attacks, public blockchain modifications, storage attacks, appending attacks and distributed denial of service attacks.

Table 9 summarizes the research papers in terms of their strengths and limitations.

 doi:10.20944/preprints202503.0979.v1

Table 9. Summary table of strengths and limitations in the research papers for the survey.

Paper	Strengths	Limitations
[3]	The Practical dataset YaleB is used.	The tradeoff between utility and privacy.
[4]	Prevent unwilling or unaware data leakage to untrusted third parties in clouds like Amazon or Google.	The tradeoff between performance, security, and privacy of the IoT-based systems.
[5]	Segment-based classification is effective for automated classification in privacy policies.	Accuracy is only 70%.
[6]	Shows that blockchain technology is effective for decentralization techniques.	Some semiautomatic systems require manual data log entry by patients, which limits flexible operations.
[7]	The triple DES encryption follows the AES algorithm and then a double DES encryption to prevent data leakage.	CNN classifiers create a big volume of synthetic data during training, requiring large storage for the data. The proposed IoMT architecture needs large hardware resources.
[8]	AI-enabled cybersecurity technology can protect individuals' cognitive information and prevent violations of cognitive privacy.	The proposed LSTM-EEG model has 99.5% accuracy, which can lead to overfitting.
[9]	The proposed framework improves security significantly.	The only case study considered is a pacemaker functioning as a medical device communicating with its environment.
[10]	The proposed model is a novel approach to device-free privacy-preserving sensing technology. Low-resolution infrared sensor based WSN nodes are used, which might affect the image quality.	--
[11]	Use of MQTT-IoT_IDS-2020, which is the most recent IoT-specific network intrusion detection dataset and NSL-KDD, which is one of the most popular benchmark datasets for traditional network traffic.	For the MQTT-IoT_IDS-2020 dataset, the achieved accuracy is 99.86%, which might lead to overfitting.
[12]	The proposed framework improves security significantly.	The only case study considered is a pacemaker functioning as a medical device communicating with its environment.
[13]	Decoupling the training of the autoencoder and the edge server classifier lessens the frequent communication between them and yields improved privacy and security.	The tradeoff between classifier accuracy, data dimensionality compression ratio, and various choices of classifiers.
[14]	A systematic mapping study for privacy-preserving techniques in IoT systems.	Industrial IoT systems are not studied elaborately.
[15]	Better predictive performance than related approaches.	Lack of mentioned performance metrics and potential easy-to-learn dataset.
[16]	Resists common network security attacks, ideal for the IoT cloud infrastructure.	High computational time as responding to multiple requests takes a long time.
[17]	High resilience to various common attacks in the IoT environment.	Unavailable result as the model was not implemented.
[18]	Better handling of bursty traffic, better privacy-overhead tradeoffs.	Inefficient in hiding larger bursts of event packets.
[19]	Data Encryption, Integrity, Privacy, Access control, Confidentiality, and Resistance to attacks.	The current system designed is the web type and can be upgraded to real-time examination.
[20]	Better privacy in handling heterogeneous data.	The scheme is yet to be implemented for real-time IOT scenarios
[21]	Blockchain provides secure vehicular communication between vehicles and users by handling fake requests.	Latency and fault tolerance are not discussed.
[22]	Security issues in the cache are addressed, and power optimization techniques such as dual	Results are not compared with previous studies.

 doi:10.20944/preprints202503.0979.v1

	sleep, sleep transistor and forced stack are utilized.	
[23]	Improved security and privacy with satisfactory costs.	Real-time scenarios are not considered.
[24]	A complete examination of the blockchain technique is provided and classifies projecting issues that must be addressed.	Though the issues are classified in detail, slight emphasis is given to possible solutions.
[25]	Around 56% of questions were answered by the proposed technology.	The analysis is mainly carried out on only six use cases. Also, it does not report all the potential issues in these cases.
[26]	Transaction authentication is done in milliseconds.	The practical results of the technique are not included.
[27]	Better Latency, processing time, and response time than other existing schemes.	The work is limited to medical certificates and can be extended to prescriptions.
[28]	Higher security (more than 89% superiority) for industrial IoT applications than other schemes.	The model should be implemented for real-life cases.
[29]	Reduced probability of nasty requests from IoT nodes.	The real-life scenario is missing.
[30]	The proposed algorithm increases the projected revenue and yields the finest evolutionary approach.	The research work finds a rough equilibrium argument.
[31]	Makes an accurate assessment of privacy representations. Offers a method of privacy shield by data minimization.	The proposed method has not been implemented yet.
[32]	Better grades of accuracy in comparison to other models.	Real-world application of the algorithm to health care networks is missing.
[33]	Combines differential privacy, federated ML, smart contracts, and blockchain.	High federation interval, 5 minutes for 5,000 tuples and 25 minutes for 60,000 tuples.
[34]	Capable of receiving artificial noise at both the transmitting and receiving nodes.	The model setup is antennas communicating with each other and not over the network.
[35]	Substantial increase in precision and recall compared to other proposed models.	The computational heavy model is composed of embeddings, CIN, DNN and AutoRec Training.
[36]	The literature survey focuses on creating complete security and privacy solutions for IoT systems, including various techniques.	The literature survey did not include any deep analysis of potential techniques and methods outlined.
[37]	Strong computation efficiency in online and offline phases, with linear time increase as the number of users increases.	Homomorphic encryption was used instead of other cryptographic encryption but requires high computational complexity.
[38]	The literature survey highlights novel solutions such as ML, DL, blockchain, SDN and more.	While the literature survey explored novel solutions, it did not include the results of such implementations.
[39]	Quantifiable IoT privacy leakage in mobile networks.	The dataset utilized potentially had easily selectable privacy-sensitive markers, which might not be accurate in the real world.
[40]	Reduces post-encryption data size compared to similar approaches.	The CASE model is scalable but has a linear increase in average network delay.
[41]	Provides location privacy against attacks through address anonymity.	It uses a lot of hashing algorithms; thus, it introduces additional computational overhead, especially for low-performance wireless sensor networks.

 doi:10.20944/preprints202503.0979.v1

6. Conclusions and Future Directions

This survey presents an accurate look at the security and privacy preservation techniques and models seen in recent IoT systems. The paper makes a detailed review of effective methods to overcome the challenges of security and privacy in IoT. A thorough assessment of the existing research work establishes that some achievements have been accomplished, which include the security of delicate and public data, protection against malicious requests, prevention of unauthorized use of data and reduction in operative costs. The review also makes some future recommendations and identifies some challenges that still need to be addressed. Firstly, the mass-produced data can be optimized and reduced to essential data. Secondly, a series of high-standard solutions are critical due to the growing content scale. In addition to this, the proposed schemes have great potential to be implemented in real-life scenarios for the evaluation of performance metrics. Table 10 shows a detailed summary of this review, identifies the research gaps for each work, and suggests their respective future recommendations.

Table 10. Summary table of research gaps and recommendations in the research papers for the survey.

Paper	Research Gaps	Recommendations
[3]	The bloom filter encoding technique is non-reversible.	A parallel key-based data encryption system can compensate for the 'reversibility' of data backup.
[4]	The developed TEE model has small memory resources, whereas running complex ML models needs larger memory and generates additional overhead.	This gap can be minimized by minimizing TCB with lesser driver functions and smaller ML models.
[5]	The privacy policy is suitable for websites, but if it needs to be extended other than websites, this framework is not sufficient.	More data regarding privacy policies for IoT devices must be included, covering different fields.
[6]	Some semi-automatic systems require manual record logging.	In the future, semi-automatic systems can be upgraded into automatic by using sensors and wearable IoT devices.
[7]	The deep learning-based model needs powerful hardware and bigger RAM. Due to the large amount of synthetic data produced by CNN, large data storage is required, which is not adequately addressed.	The storage problem can be mitigated through cloud computing services.
[8]	The proposed model was developed based on the EEG data records from only 109 subjects.	In the future, a large volume of data should be used to validate the model's performance.
[9]	Only a single case study is not enough to propose the taxonomy.	Elaborate studies covering different case studies are recommended.
[10]	The proposed novel model is developed based on 26 yoga postures only.	More data for yoga postures should be included in future studies for the robustness of the systems.
[11]	Nothing is mentioned in the study about handling data with significantly different values.	Outliers' detection techniques can be included in the developed model.
[12]	Only a single case study is not enough to propose the taxonomy.	Elaborate studies covering different case studies are recommended.
[13]	The current comparison technique cannot quantify the advantages and disadvantages of the proposed model.	Comparison with federated learning and SplitNN can be done in future for classifier performance vs. communication cost and model complexity for image classification.

 doi:10.20944/preprints202503.0979.v1

[14]	Industrial IoT systems are not studied elaborately.	In the future, privacy preservation techniques in industrial IoT systems should be included in the study.
[15]	Additional testing to reduce the computational overhead of the large model.	Addition of performance metrics to check for overfitting.
[16]	No testing of security threats other than common ones.	Further testing must be done to test other network security threats.
[17]	A comprehensive analysis of the resilience of various network attacks is needed.	Implementation of the model to determine computational overhead.
[18]	Difficulty in noting real-time cycle length and onset times during the design phase.	Shaper design efficiency can be improved by modelling the correlation of traffic input.
[19]	The current system designates a web category and can be upgraded to real-time patient examination.	The data can be filtered before being sent to the hospital, and Real-time investigation can be added.
[20]	Proof of concept for study for the methodology is yet to be implemented.	Security analysis of the proposed scheme can be carried out to bring strengths and limitations.
[21]	A comparison table regarding security and efficiency results to previous works is missing.	Simulation results for Proximity, Fault tolerance and latency can also be analyzed.
[22]	The power consumption and area overhead witness a trade-off.	This work can be utilized further in the array form.
[23]	There is a need to support the flexibility of federations. Pre-shared information must not be disclosed.	The technique can be expanded with RFID to comprehend automated on-chain identification.
[24]	Challenges and complications of blockchain technology need to be addressed.	Hybrid consensus protocols can be established by including speed, computational requirements, and other parameters.
[25]	The proposed solution is required to be tested for a broader set of areas.	A chatbot interface can be included further to enhance the interaction.
[26]	The prototype can be practically developed on a real-time dataset to generate strong results.	The proposed technique can be expanded to other finance systems.
[27]	Research can be done to achieve access control for mutual verification.	Government rules and standards can also be studied to produce certificates with stamps.
[28]	The model can be applied to broader applications in the industry.	The model can be utilized for other real-time applications.
[29]	The research work found an approximate value of the theoretic equilibrium.	The results obtained can be applied to real-world datasets.
[30]	Other game models can also be utilized to grip privacy protection.	Privacy protection for different IoT devices under data aggregation can be the next step with the proposed algorithm.
[31]	Experimental evaluation of the proposed method can be more specific.	The presentation of the proposed method can be examined and compared with previous methods.
[32]	Associations with healthcare workers can aid in reflecting legal and ethical contemplations.	The study of user involvement and usableness is a vital part of upcoming work for this technique.
[33]	The authors did not compare with other models for a baseline comparison.	Further testing different approaches to reduce latency and improve efficiency
[34]	The authors did not compare with other proposed models for comparison.	Further analysis with other proposed models for a better comparison.
[35]	The authors did not include the training time of the model.	Further testing to include more metric analysis for comparison with other proposed models and computational overhead.
[36]	The literature survey was too short and did not include a deep analysis of each research direction.	A deeper analysis of each research direction, comparing multiple models.

 doi:10.20944/preprints202503.0979.v1

[37]	Further analysis into other methods that can balance privacy preservation and model performance.	Further analysis of other models is necessary to achieve lower time performance.
[38]	The literature survey had an in-depth knowledge of various state-of-the-art solutions, but it did not include further analysis of them.	The literature survey should include an example of how each state-of-the-art solution is applied in practice.
[39]	The results are unknown to be accurate as metrics are not included, or results are not compared with other models.	Extend the model to other datasets for further analysis.
[40]	The authors compared three different models: the NB, SVM and DT. All three models achieved mostly different accuracy, clock cycles and time response thus, further models need to be tested to achieve more uniform results.	Further analysis of other models that can achieve higher accuracy, but lower average clock cycles and average time is necessary for better performance.
[41]	The authors only compared the power consumption of their model with another routing algorithm but did not compare other analyses.	Further analysis with the compared routing algorithm for a better analytic comparison.

References

1. Duarte, F. EXPLODING TOPICS, 22 February 2023. Available online: https://explodingtopics.com/blog/number-of-iot-devices (accessed on 27 October 2023).
2. CHECK POINT. Check Point Research, 11 April 2013. Available online: https://blog.checkpoint.com/security/the-tipping-point-exploring-the-surge-in-iot-cyberattacks-plaguing-the-education-sector/ (accessed on 27 October 2023).
3. Xue, W.H.W.G.P.S.A.; J.S. An efficient privacy-preserving IoT system for face recognition. In Proceedings of the 2020 Workshop on Emerging Technologies for Security in IoT (ETSecIoT), IEEE, April 2020; pp. 7-11.
4. Yuhala, P. Enhancing IoT Security and Privacy with Trusted Execution Environments and Machine Learning. arXiv preprint arXiv 2023, 2305.02584.
5. L., J.D.; Carson, R.L. Automatic Classification of Web and IoT Privacy Policies. In Proceedings of the IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS), October 2022; pp. 732-735.
6. Elkahlout, M.A.-S.M.A.A.I.A.; D. M. 2. A. .. I. 2. I. C. o. A. a. R. T. IoT-Based Healthcare and Monitoring Systems for the Elderly: A Literature Survey Study. In Proceedings of the International Conference on Assistive and Rehabilitation Technologies (iCareTech), IEEE, 28 August 2020; pp. 92-96.
7. R., B.; Meenakshiammal, K. Preserving Patient Privacy in IoT Based Breast Cancer Monitoring System. In Proceedings of the 2nd International Conference on Edge Computing and Applications (ICECAA), IEEE, 19 July 2023; pp. 1370-1374.
8. N., F.M.; Schiliro, B.A. Cognitive privacy: AI-enabled privacy using EEG signals in the internet of things. In Proceedings of the IEEE 6th International Conference on Dependability in Sensor, Cloud and Big Data Systems and Application (DependSys), 14 December 2020; pp. 73-79.
9. O., E.; Fazeldehkordi, N.J. Security and privacy in IoT systems: a case study of healthcare products. In Proceedings of the 13th International Symposium on Medical Information and Communication Technology (ISMICT), IEEE, 8 May 2019; pp. 1-8.
10. Gochoo, M.T.T.H.S.B.T.H.J.A.F.C.Y. Novel IoT-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning. IEEE Internet Things J. 2019, 6, 7192-7200.
11. Jui, T.H.M.M.S.H.M. Feature Reduction through Data Preprocessing for Intrusion Detection in IoT Networks. In Proceedings of the Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), IEEE, 13 December 2021; pp. 41-50.
12. O., E.; Fazeldehkordi, N.J. Security and Privacy Functionalities in IoT. In Proceedings of the 17th International Conference on Privacy, Security and Trust (PST), IEEE, 26 August 2019; pp. 1-12.
13. Agbohungbe, O.R.S.D.X.Q.L. Efficient privacy-preserving edge intelligent computing framework for image classification in IoT. IEEE Trans. Emerg. Top. Comput. Intell. 2021, 6, 941-956.

 doi:10.20944/preprints202503.0979.v1

14. Fagbohungbe, R.S.D.X.Q.L. Efficient privacy-preserving edge intelligent computing framework for image classification in IoT. IEEE Trans. Emerg. Top. Comput. Intell. 2021, 6, 941-956.
15. Alotaibi, B.; Alotaibi, M. A stacked deep learning approach for IoT cyberattack detection. J. Sens. 2020, 2020, 8828591.
16. Kahani, N.; Elgazzar, K.; Cordy, J.R. Authentication and access control in e-health systems in the cloud. In Proceedings of the 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), 13-23 July 2016.
17. Meisami, S.; Atashgah, M.B.; Aref, M.R. Using blockchain to achieve decentralized privacy in IoT Healthcare. Int. J. Cybern. Inform. 2023, 12, 97-108.
18. Xiong, S.; Sarwate, A.D.; Mandayam, N.B. Network Traffic Shaping for Enhancing Privacy in IoT Systems. IEEE/ACM Trans. Netw. 2022, 30, 1162-1177.
19. Azbeg, K.; Ouchetto, O.; Andaloussi, S.J. Access Control and Privacy-Preserving Blockchain-Based System for Diseases Management. IEEE Trans. Comput. Soc. Syst. 2023, 10, 1515-1527.
20. Will, N.C. A Privacy-Preserving Data Aggregation Scheme for Fog/Cloud-Enhanced IoT Applications Using a Trusted Execution Environment. In Proceedings of the 2022 IEEE International Systems Conference (SysCon), Montreal, QC, Canada, 2022; pp. 1-5.
21. Agrawal, R.; Faujdar, N.; Kumar, P.; Kumar, A. Security and Privacy of Blockchain-Based Single-Bit Cache Memory Architecture for IoT Systems. IEEE Access 2022, 10, 35273-35286.
22. Ouyang, L.; Wang, F.Y.; Tian, Y.; Jia, X.; Qi, H.; Wang, G. Artificial Identification: A Novel Privacy Framework for Federated Learning Based on Blockchain. IEEE Transactions on Computational Social Systems 2022.
23. Gugueoth, V.; Safavat, S.; Shetty, S.; Rawat, D. A Review of IoT Security and Privacy Using Decentralized Blockchain Techniques. Computer Science Review 2023, 50, 100585.
24. Alkhariji, L.; De, S.; Rana, O.; Perera, C. Semantics-Based Privacy by Design for Internet of Things Applications. Future Generation Computer Systems 2023, 138, 280-295.
25. Singh, R.; Dwivedi, A.D.; Srivastava, G.; Chatterjee, P.; Lin, J.C.W. A Privacy-Preserving Internet of Things Smart Healthcare Financial System. IEEE Internet of Things Journal 2023, 10, 18452-18460.
26. Sharma, P.; Namasudra, S.; Chilamkurti, N.; Kim, B.G.; Crespo, R.G. Blockchain-Based Privacy Preservation for IoT-Enabled Healthcare System. ACM Transactions on Sensor Networks 2023, 19, 56.
27. Kumar, M.; Mukherjee, P.; Verma, S.; et al. A Smart Privacy Preserving Framework for Industrial IoT Using Hybrid Meta-Heuristic Algorithm. Scientific Reports 2023, 13, 5372.
28. Tayeb, H.; Bramas, B.; Faverge, M.; Guermouche, A. Dynamic Tasks Scheduling with Multiple Priorities on Heterogeneous Computing Systems. 2024.
29. Shen, S.; Wu, X.; Sun, P.; Zhou, H.; Wu, Z.; Yu, S. Optimal Privacy Preservation Strategies with Signaling Q-Learning for Edge-Computing-Based IoT. Expert Systems with Applications 2023, 225, 120192. https://doi.org/10.1016/j.eswa.2023.120192.
30. Alam, T.; Gupta, R. Federated Learning and Its Role in the Privacy Preservation of IoT Devices. Future Internet 2022, 14, 246. https://doi.org/10.3390/fi14090246.
31. Yaraziz, M.S.; Jalili, A.; Gheisari, M.; Liu, Y. Recent Trends Towards Privacy-Preservation in Internet of Things, Its Challenges and Future Directions. IET Circuits, Devices & Systems 2023, 17. https://doi.org/10.1049/cds2.12138.
32. Ali, A.; Al-rimy, B.A.S.; Alsubaei, F.S.; Almazroi, A.A.; Almazroi, A.A. HealthLock: Blockchain-Based Privacy Preservation Using Homomorphic Encryption in Internet of Things Healthcare Applications. Sensors 2023, 23, 6762. https://doi.org/10.3390/s23156762.
33. Arachchige, P.C.; et al. A Trustworthy Privacy-Preserving Framework for Machine Learning in Industrial IoT Systems. IEEE Transactions on Industrial Informatics 2020, 16, 6092–6102. https://doi.org/10.1109/TII.2020.2974555.
34. Tayeb, H.; Bramas, B.; Faverge, M.; Guermouche, A. Dynamic Tasks Scheduling with Multiple Priorities on Heterogeneous Computing Systems. 2024.

 doi:10.20944/preprints202503.0979.v1

35. Anajemba, J.H.; Iwendi, C.; Razzak, I.; Ansere, J.A.; Okpalaoguchi, I.M. A Counter-Eavesdropping Technique for Optimized Privacy of Wireless Industrial IoT Communications. IEEE Transactions on Industrial Informatics 2022, 18, 6445–6454. https://doi.org/10.1109/TII.2021.3140109.
36. Ogunseyi, T.B.; Bo, T.; Yang, C. A Privacy-Preserving Framework for Cross-Domain Recommender Systems. Computers & Electrical Engineering 2021, 93, 107213. https://doi.org/10.1016/j.compeleceng.2021.107213.
37. Bertino, E. Data Security and Privacy in the IoT. Purdue University 2016, 1–3. https://doi.org/10.5441/002/edbt.2016.02.
38. Wei, Z.; et al. Lightweight Federated Learning for Large-Scale IoT Devices with Privacy Guarantee. IEEE Internet of Things Journal 2023, 10, 3179–3191. https://doi.org/10.1109/JIOT.2021.3127886.
39. Karunarathne, S.M.; Saxena, N.; Khan, M.K. Security and Privacy in IoT Smart Healthcare. IEEE Internet Computing 2021, 25, 37–48. https://doi.org/10.1109/MIC.2021.3051675.
40. Hui, S.; et al. Systematically Quantifying IoT Privacy Leakage in Mobile Networks. IEEE Internet of Things Journal 2021, 8, 7115–7125. https://doi.org/10.1109/JIOT.2020.3038639.
41. Ghosh, T.; Roy, A.; Misra, S.; Raghuwanshi, N.S. CASE: A Context-Aware Security Scheme for Preserving Data Privacy in IoT-Enabled Society 5.0. IEEE Internet of Things Journal 2022, 9, 2497–2504. https://doi.org/10.1109/JIOT.2021.3101115.
42. Kadir, N.; Kaur, R.; Rodrigues, T.; Kashef, R. Post COVID-19 Vaccination: Infection Rate Analysis Using Time Series Modeling. Proceedings of the 2024 International Conference on Machine Intelligence and Smart Innovation (ICMISI), Alexandria, Egypt, 2024; pp. 266–271. https://doi.org/10.1109/ICMISI61517.2024.10580825.
43. Kaur, R.; Mohammadi, F. Comparative Analysis of Power Efficiency in Heterogeneous CPU-GPU Processors. In Proceedings of the 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), 24 July **2023**; pp. 756-758. IEEE.
44. Deebak, B. D.; Hwang, S. O. Privacy-Preserving Learning Model Using Lightweight Encryption for Visual Sensing Industrial IoT Devices. IEEE Trans. Emerg. Top. Comput. Intell. 2025.
45. Asad, A.; Kaur, R.; Mohammadi, F. Noise Suppression Using Gated Recurrent Units and Nearest Neighbor Filtering. In Proceedings of the 2022 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, **2022**, pp. 368-372. doi: 10.1109/CSCI58124.2022.00072.
46. Kaur, R.; Asad, A.; Mohammadi, F. A Comprehensive Review of Processing-in-Memory Architectures for Deep Neural Networks. Computers 2024, 13, 174. https://doi.org/10.3390/computers13070174.
47. Kaur, R.; Asad, A.; Al Abdul Wahid, S.; Mohammadi, F. A Survey of Advancements in Scheduling Techniques for Efficient Deep Learning Computations on GPUs. Electronics 2025, 14, 1048. https://doi.org/10.3390/electronics14051048.
48. Wang, R.; Lai, J.; Li, X.; He, D.; Khan, M. K. RPIFL: Reliable and Privacy-Preserving Federated Learning for the Internet of Things. J. Netw. Comput. Appl. 2024, 221, 103768.
49. Asad, A.; Kaur, R.; Mohammadi, F. A Survey on Memory Subsystems for Deep Neural Network Accelerators. Future Internet 2022, 14, 146. https://doi.org/10.3390/fi14050146.
50. Abdel-Basset, Mohamed, Hossam Hawash, Nour Moustafa, Imran Razzak, and Mohamed Abd Elfattah. "Privacy-preserved learning from non-iid data in fog-assisted IoT: A federated learning approach." Digital Communications and Networks 10, no. 2 (2024): 404-415.
51. Kaur, R.; Bansal, M. BDD Ordering and Minimization Using Various Crossover Operators in Genetic Algorithm. Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng. 2014, 2, 1247–1253. Available online: www.ijireeice.com.
52. Joshi, P.; Hasanuzzaman, M.; Thapa, C.; Afli, H.; Scully, T. Enabling All In-Edge Deep Learning: A Literature Review. IEEE Access 2023, 11, 3431–3460.
53. Abdul Wahid, S. A.; Asad, A.; Kaur, R.; Mohammadi, F. Quantum Computing Circuit Design: A Tutorial. Proceedings of the 2024 International Conference on Advanced Scientific Computing (ICASC), Cluj-Napoca, Romania, 2024; pp. 1–6. https://doi.org/10.1109/ICASC63229.2024.10785069.

 doi:10.20944/preprints202503.0979.v1