




**REVIEW**

# Recent trends towards privacy-preservation in Internet of Things, its challenges and future directions

Mahdi Safaei Yaraziz<sup>1</sup>  | Ahmad Jalili<sup>2</sup> | Mehdi Gheisari<sup>3</sup>  | Yang Liu<sup>4</sup> 

<sup>1</sup>Department of Computer Engineering, Islamic Azad University, Zanjan Branch, Zanjan, Zanjan Province, Iran

<sup>2</sup>Department of Computer Engineering, Gonbad Kavous University, Gonbad-e Kavus, Iran

<sup>3</sup>Faculty of Automation, Guangdong University of Technology, Guangzhou, China

<sup>4</sup>Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China

**Correspondence**

Mahdi Safaei Yaraziz, Department of Computer Engineering, Islamic Azad University, Zanjan Branch, Zanjan, Zanjan Province, Iran.  
Email: safaei.mahdi1988@gmail.com

**Abstract**

The Internet of Things (IoT) is a self-configuring, intelligent system in which autonomous things connect to the Internet and communicate with each other. As 'things' are autonomous, it may raise privacy concerns. In this study, the authors describe the background of IoT systems and privacy and security measures, including (a) approaches to preserving privacy in IoT-based systems, (b) existing privacy solutions, and (c) recommending privacy models for different layers of IoT applications. Based on the results of our study, it is clear that new methods such as Blockchain, Machine Learning, Data Minimisation, and Data Encryption can greatly impact privacy issues to ensure security and privacy. Moreover, it makes sense that users can protect their personal information easier if there is fewer data to collect, store, and share by smart devices. Thus, this study proposes a machine learning-based data minimisation method that, in these networks, can be very beneficial for privacy-preserving.

**KEYWORDS**

data flow analysis, data handling, data privacy, Internet of Things, security of data

## 1 | INTRODUCTION

Technology is evolving, and the development of new devices and equipment has led to the emergence of extensive equipment networks for collecting and storing data. In the meantime, the Internet of Things has introduced and developed an efficient system and network for managing and controlling mass data. IoT aims to connect all devices worldwide via the Internet. Undoubtedly, many smart devices and electronic gadgets enter the world of information every year. It means that more and more data is produced every year, and the need for data processing centres to extract useful information from these data. The IoT system has been used in various applications, including smart cities, healthcare, transportation, agriculture etc., because of its high efficiency. The Internet of Things combines several technologies and cost-effective small-scale equipment, including GPS, Sensors, smartphones, wearable gadgets etc., to provide a low-cost communication network in a large space with Internet access. IoT networks have the advantage of being scalable, heterogeneous, and open

so that equipment can connect without needing to be reconfigured as it gets upgraded.

The Internet of Things covers a wide variety of capabilities and services. However, like other networks, these networks need information security and privacy protection. For instance, in the aspect of smart healthcare, medical information faces more security problems than other systems, such as smart agriculture. This system requires special attention to the issue of information and device security, as well as privacy when deployed in a large space. Although the Internet of Things offers many benefits, it can still pose security, privacy, and vulnerability concerns. These include Denial of Service attacks, water holes, eavesdropping, ransomware etc., and may compromise infrastructure as well as data. However, the need to apply and implement comprehensive approaches in identifying, authenticating users and managing equipment to provide security and privacy in this type of distributed and heterogeneous networks should be anticipated and presented. Sharing information rather than exchanging duplicate data significantly reduces data transmission and local storage stress. Effective

This is an open access article under the terms of the Creative Commons Attribution-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited and no modifications or adaptations are made.

© 2022 The Authors. *IET Circuits, Devices & Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

learning models may therefore be shared and exchanged as knowledge assets. Due to the sensitive nature of this data, it is crucial to secure it from inauthentic entities that may pose numerous threats to users' privacy [1].

IoT Analytics predicts that approximately 27 billion endpoints will be active in 2025 [2]. However, some difficulties, such as Covid-19 disease may affect the growth of this type of network. The number of devices may be additive soon because most people are becoming addicted more to new technologies, such as mobile solutions, after the pandemic lockdown. This means producing data will accelerate and increase. Thus, the possibility of information leakage or data leakage will increase, and the privacy protection method can be used on each device to prevent these things from happening [3]. There is a discussion about security and privacy-based applications, and optimised solutions have become a major concern of IoT applications. Taking a close look at some technical surveys and review articles in the field of privacy in IoT does not allow you to find useful key information about IoT applications. However, this paper aims to explain different IoT applications with analysis to understand their approaches, which have been published recently.

These approaches require security and privacy because most of these networks exchange big data on a large scale. However, the authors discussed privacy-related approaches to IoT networks in this article. A comprehensive study has been conducted to cover the solutions and problems and route active research in the field. As shown in Figure 1, categorised privacy applications are into three layers: Edge, Fog, and Cloud. Due to Figure 1, privacy can be implemented in any of the layers. With the development and implementation of new IoT systems, privacy and data protection have become important and critical issue that needs to be addressed [4]. The Privacy Mechanism can be managed on devices, gateways, data centres, communications, and data processing, although there

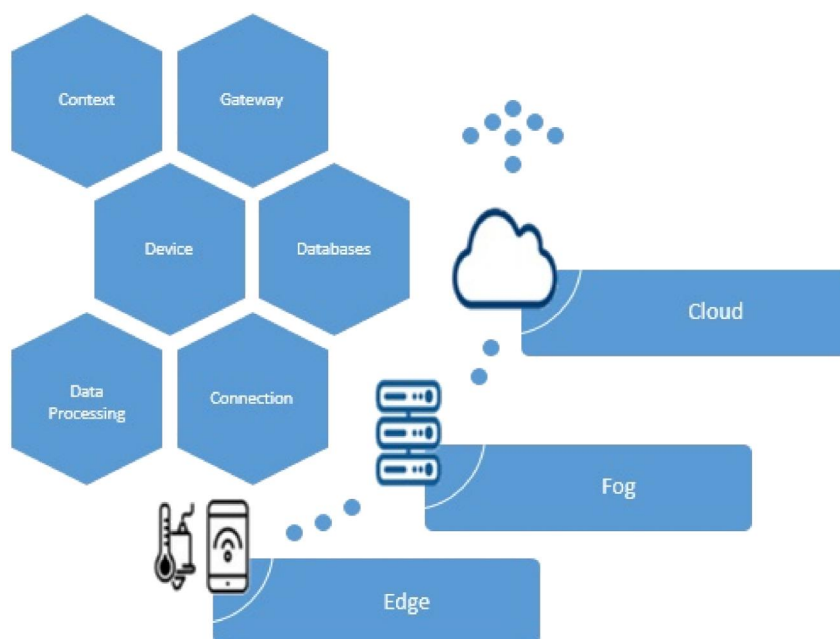
are challenges in each that make it difficult to come up with an optimal approach. However, security breaches and unauthorised access to sensitive data should be avoided. Controlling access to data is the first line of defence, limiting its use to authorised people.

A review was conducted to evaluate privacy-based applications and solutions. Therefore, the points chosen for describing this paper are titled as follows:

- Illustrate the sort of privacy-based application by a technical taxonomy
- Illustrating and discussing the fundamental challenges for privacy-based IoT applications
- Indicate a clear path and research challenges in this area that may need to be addressed in the future

Figure 1 provides a proportioned figure of privacy in the diversity of parts of IoT-based layers to demonstrate domains. It means privacy-preserving is necessary for the operation of systems. For instance, state-of-the-art data security and privacy solutions in edge-related paradigms, such as devices, context etc., are crucial to prevent privacy leakage from data transition among multiple edge nodes. Moreover, another challenge is related to the fog's design purposes of reducing the latency and improving the bandwidth (in connection and gateway), where the privacy-preserving methods work to overcome the issues. Privacy preservation in cloud environments (internal layers) includes two aspects: data processing security and data storage security (Database). Data processing security covers the issues of how to protect user privacy at runtime in a virtualised cloud platform. Data storage security covers the issues of guaranteeing user data privacy when the data is stored in the data centre.

According to the Federal Trade Commission report [5], some issues related to privacy and information security will be



**FIGURE 1** Three-tier layers of Internet of Things Architecture with privacy aspects.

one of the issues and challenges in the near future in the big data and IoT field. IoT devices can generate large amounts of data, even in large quantities, which is very staggering. Data security and data privacy are at risk as hackers may misuse the information. This paper's authors believe that privacy breaches can occur in the following three ways:

- System hacking (including eavesdropping, SQL (Structured Query Language) injections, sniffing, phishing etc.)
- Illegal use by the IoT service providers
- Lack of familiarity or negligence of system users

Privacy risk is that a thief or an intruder can remotely monitor or disrupt a private space. According to our research, the data produced by IoT systems is increasing day by day, and the 'Data Minimising' vacuity is more visible in these types of systems, which can be one of the parameters related to data privacy. It is also possible to define 'data minimisation' in terms of limiting the discovery of knowledge. The authors have outlined many of the advantages and challenges associated with this approach in the following sections. However, researchers explore various innovative methods to mitigate cyberattacks and increase data privacy. Their conceptual framework is based on a strategic aim of ensuring the confidentiality (the guarantee that only authorised users have access to data) of private and personally identifiable information in the application data centre.

In light of problem statements regarding the security and privacy of IoT applications, Leakage of these data can lead to massive security and privacy violations due to improper handling. Thus, illegal access to a large amount of data could be a lucrative resource for hackers. One of the highlighted challenges in mass data is a data breach because of unfair data collection. This paper provides up-to-date and highlights information on the current research progress in IoT security, privacy challenges, and approaches for protecting users' data and the sustainability of IoT-based systems and finally propose a new solution to address some difficulties in this aspect.

## 2 | MATERIALS AND METHODS

This section briefly describes the materials and methods of privacy-based IoT applications in the applying review process. Figure 1 Presents a taxonomy of privacy in three layers of IoT.

This paper highlights the following Analytical Questions (AQ) with clear answers to each of them:

- AQ1: What are the main reasons for choosing IoT as a powerful solution?  
 AQ2: What security/privacy methods have been used in IoT applications?  
 AQ3: What are the research gaps and open questions for future research directions?

A privacy-enhanced federated learning scheme for IoE was presented by Li et al. [6]. Two mechanisms, namely, the

Random Response (RR) mechanism and the Local Adaptive Differential Security (LADP) mechanism, are applied in this approach. The RR is passed to prevent the server from knowing which updates are collected per turn. LADP allows devices to add noise to local updates quickly before sending them to the server. This method applies a Gaussian mechanism to distort each client's local updates. The Gaussian mechanism is the building block of a proprietary algorithm for minimising empirical risk based on stochastic gradient descent. The main strategy is to increase customer privacy at the local level to adapt the approach to an environment without trusted servers. With this approach, greater accuracy is achieved so that privacy is maintained. The weakness of this method is that it depends on the effectiveness of the batch size, learning rate, and machine learning parameters.

A privacy-preserving recommendation mechanism by Blockchain to address problems related to processing and analysing data closely related to users' privacy by Lin et al. [7] is presented. However, a fully distributed model that mitigates the risk of privacy disclosure through centralised data storage has been established to prevent the risk of privacy. A sensitive local hash and local differential privacy were introduced to reduce computational load and provide a strong guarantee of data protection—also, an interplanetary file system with Blockchain combined to improve communication efficiency significantly. The advantage of this mechanism is balancing accuracy and privacy during the recommendation process. The main weakness of this method is that it cannot guarantee accuracy if users' preference features are not taken properly.

Gheisari et al. [3] introduced a three-module framework named 'Ontology-Based Privacy-Preserving' (OBPP) to address three important challenges: heterogeneity, privacy-preserving of generated data, and providing high-level services in IoT-based smart cities. A data storage model and an ontology have been used to solve the problem of heterogeneity. Privacy-preserving IoT devices to find abnormal patterns while addressing service quality. Semantic reasoning rules and, finally, dynamically changing the privacy behaviours of the devices helped privacy rule managers to address the privacy-preserving challenges. The OBPP consists of three modules: 'Ontology', 'Reasoning engine' and 'OBPP procedure' which is located in the Cloud Computing space. It may be more efficient and provide better results in terms of accuracy and could be a good approach in large-scale networks, but in order to support OBPP on increasing numbers of devices, a reliable server is necessary.

Bao et al. [8] proposed an effective, revocable, privacy-preserving fine-grained data sharing with a keyword search (ERPF-DS-KS) scheme to mollify data security and user privacy concerns. This scheme realises effective, fine-grained access control and ciphertext keyword search and offers flexible indirect revocation to malicious data users. A pseudo-identity-based signature system is intended to offer data authenticity. This approach provides lightweight operations for the resource-constrained device in cloud-assisted MIIoT, allowing the cloud to quickly check whether a ciphertext includes the required keyword with minor computational overheads.

Gheisari et al. [9] suggested a context-aware privacy-based approach using Software Defined Networking (SDN) by looking at three major parameters, including applying none static privacy-preserving method, sending sensitive data packets by splitting data into two parts (70%, 30%) through a secure route and a virtual private network (VPN), and context-aware. In this method, IoT-based smart cities are equipped with the SDN paradigm to allow centralised management and flexibility of packets. Moreover, OpenFlow switches and the data plane and control plane are separated in this scenario, and the SDN controller controls the network's data flow to protect privacy. This method addresses three significant drawbacks of literature methods: one static privacy-preserving method for the whole system, sending whole data at once, and not being context-aware. As a result, privacy could be preserved by splitting sensitive data and sending it over a secure route and VPN. The advantages of the proposed method could be better performance in terms of accuracy, affordable overhead, and penetration rate. On the other hand, this approach has some drawbacks: it could not guarantee overhead in increasing the number of devices, and the performance of this solution in resource-constrained devices is inefficient. It could not work for a long time than other methods.

In ref. [10], an intrusion-resilient server-aided attribute-based signature (IR-SA-ABS) approach was presented. The IR-SA-ABS scheme has seven components, including the trusted authority (TA), the attribute authority (AA), the data publisher (which gathers data and generates signatures with server support), the helper device (which updates and refreshes the key), the server, the cloud, and the data subscriber (signature checker). This strategy is intended to regularly update the signature key with the help and support of a helper device, and the signing key is renewed for many rounds periodically within each time period under the control of the helper device. As a result, the system remains safe even if both the helper device and the signing device are compromised. Throughout this operation, the computational overheads of resource-constrained devices are considerably reduced by outsourcing all key updates and refresh activities, as well as the majority of signature creation and verification processes, to a capable server.

Lomotey et al. [11] provided a novel scenario with a combination of policy-based provenance and modelled the peer-to-peer IoT device communication as a graph network also to determine the shortest paths of interconnected IoT devices Floyd's algorithm added to this approach. In the situation that there is no direct link between two IoT nodes, it can be challenging to determine the shortest route for sending data from one to the other. So, Floyd's algorithm-based technique solves this problem. This algorithm finds the shortest path between two points in a directed graph on which the edge weights are positive or negative. Nonetheless, these two approaches were introduced for device and data verification and ultimately to enforce trust and privacy. It is difficult to determine the data source, especially when identifying it is critical to decision-making. So, when an untrusted device track and trace the origin of the data while retaining metadata information of

other devices, it can lead to privacy issues. A dual approach is proposed to verifying and tracing data and devices in an IoT architecture without prejudice to privacy.

Liu et al. [12]. were the first to propose a distributed access control system based on blockchain technology to provide security in the IoT. Fog computing and the alliance chain theory are at the core of this mechanism. The IoT data is encrypted on an edge node using Mixed Linear-Nonlinear Coupled Map Lattice (MLNCML) chaotic systems and the least significant bit (LSB) before being uploaded to the cloud. By providing dynamic and fine-grained access controls for IoT data, the proposed mechanism can provide a solution to the issue of a single point of failure. An edge node can update the model policies by uploading a new model policy to the Blockchain with a timestamp at any given time to facilitate dynamic access control. In this approach, with blockchain technology, alliance nodes can share sensitive data accurately and privately. The weakness of this method is that more complex and requires a lot of computing, and in addition to complexity, it puts a lot of overhead on the system and cannot be used in large-scale networks.

Srivastava et al. [13] proposed a unique blockchain-based IoT paradigm to improve the present IoT-based remote patient monitoring system's security and privacy. With more modern and lightweight cryptographic approaches like the ARX (Addition/Rotation/XOR) encryption scheme, this paradigm allows trustworthy data transfer over the network and cloud storage. The system also includes Ring Signatures, which give crucial privacy aspects such as Signers' Anonymity and Signature Correctness. A double encryption approach is also utilised to make the symmetric key more secure across the network. The Diffie-Hellman key exchange mechanism also protects the public key from an intruder.

Ren et al. [14] proposed a stream data anonymisation method as a privacy-enhancing technique for data collecting by IoT devices. However, privacy enhancement techniques have been proposed for various IoT streaming and media data scenarios. This approach provides a data anonymisation scheme that can anonymise the flow of data before it is stored or transmitted to another system or organisation without compromising user privacy or other confidentiality. So, to anonymise data, the anonymisation engine must be used to define data protection. To show its performance from different angles with three types of data, including context, continuous, and media data considered. The evolution result show this method can keep confidentiality well without significantly affecting the usefulness of data. On the other hand, this work has some drawbacks: in the case of large-scale systems, it adds significant overhead to the system, and for better performance, sensitive data need to anonymise to improve system efficiency.

Bao et al. [15] suggested a lightweight attribute-based searchable encryption (LABSE) system in smart healthcare for resource-constrained devices that achieves fine-grained access control and keyword search while lowering computational overhead. The research attempts to address the essential issue of processing time and energy consumption for resource-constrained devices in smart healthcare. The healthcare



administration centre (HAC) (for developing and monitoring the system), the healthcare cloud service provider (HCSP) (for storing the ciphertext), the implantable/wearable sensor on the patient side, and the mobile terminal device on the doctor's side are four kinds of entities involved in this strategy. A lightweight and anonymous authentication technique is provided to avoid tampering and forgery during Ciphertext transmission. The experimental findings reveal that the LABSE scheme is more efficient in terms of computing time and energy consumption than other state-of-the-art systems.

An architecture for healthcare to preserve privacy by Blockchain approach called BCHealth was proposed by Hossein et al. [16]. The BCHealth architecture allows data owners to control access to their private personal health data to overcome the dilemma of transparency versus access control. Using local storage, BCHealth stores data closest to the data owner instead of a cloud or a health data centre. It introduces the usage of two different chains: the data chain and the access control (policy) chain. Health data is stored in the data chain, and access control policies are stored in the chain with access control policies for the patient. Additionally, BCHealth has an emergency alarm system that alerts corresponding medical staff if immediate action is required based on patient health status. A new modified BC (BlockChain) network and Proof-of-Authority (PoA) consensus algorithm have been provided to improve the system's performance and scalability. This system preserves the privacy of patient data by storing hashes of the data as transactions in the chain. By doing so, delays will be reduced, and the bandwidth used will be reduced. In order to improve the BC network's scalability and throughput, BCHealth uses a new clustering approach. This method can benefit from storing the hash of the data on the Blockchain and using clustering to reduce delay and storage requirements, both of which are advantages. The downside of this approach is that it takes longer and has more overhead than centralised solutions.

To address privacy, efficiency, and usability concerns, Bao et al. [17] first describe an efficient technique, PH-ABE-DS, that achieves full policy concealing by implementing access control using the inner product. Furthermore, the authors devised an effective indirect revocation method that allows the cloud and users to update the ciphertext and secret user key with minimal storage and computational overhead. On this principle, the EA-PH-ABE-DS approach, by utilising edge computing, considerably minimises the overheads of the described resource-constrained devices. As a result, one of the advantages of this technique is that, via comprehensive theoretical and experimental comparisons, both suggested two schemes demonstrate their superiority over the most recent comparable works in terms of functionality and performance and are secure and reliable.

Qu et al. [18] work focussed on a framework for solving the identified fog computing issues based on blockchain-enabled federated learning (FL-Block). This blockchain-driven global learning platform offers local updates for end devices that miners verify. FL-Block enables autonomous machine learning by using the Blockchain's consensus

mechanism to achieve consensus and to maintain a global model without relying on any centralised authority. Fog servers generate and store global updates based on local updates being sent by end devices. The block generation efficiency could be guaranteed since only the global update pointer is saved on-chain, and a distributed hash table (DHT) is used to save the data. However, FL-Block ensures decentralised privacy protections by preventing single points of failure. The evolution result show this method can keep accuracy, efficiency, and resistance to poisoning attacks. But there are some disadvantages associated with this work, such as computation and communication costs.

Decentralised applications such as Distributed Ledger Technologies and Blockchain have emerged as excellent applications for the secure exchange of information in a decentralised manner employing privacy-preserving approaches such as zero-knowledge protocols. Singh et al. [19] provide the advanced zero-knowledge ledger by replacing their range-proof approach with the most efficient range-proof technique based on the enhanced inner product-based zero-knowledge proofs. This approach combines numerous range-proofs into a single range-proof, making the present zero-knowledge ledger system more efficient than the earlier one. The performance improved by replacing the current zkLedger-based auditing system's range of proof with the most efficient range proof technique based on the improved inner product based on zero-knowledge proof. This method could be more efficient in terms of the need for low computational power and memory for resource-constrained devices. On the other side, the system's drawback is that it contains some extra calculations that take a certain amount of time to complete.

## 2.1 | A brief summary of the evaluation of solutions

This sub-section presents a short analysis of the security/privacy-based IoT applications shown in refs. [3, 6–19]. IoT applications generate enormous amounts of data. There is always the risk of data leakage and security issues; those mentioned applications covered these problems and proposed state of the art of security/privacy algorithms. According to the review and classification of articles [7, 12, 13, 16, 18, 19], blockchain technology is used to increase and provide system security and privacy in data transfer. In short, a blockchain can act as a distributed ledger with digitally signed data and is auditable; any changes made therein can be traced back to the original data, ensuring security. This shows that blockchain technology is more reliable in securing data. Moreover, other applications use different types of security and privacy. For instance, Ontology and Anonymisation have been used for privacy in articles [3, 14]. On the other hand, due to the mass production of data in the IoT, security and privacy are challenged. The evolution result shows the mentioned methods have some limits in large-scale mode, including dependency on a reliable server, depending on the user's preference, recourse-constraint, the complex and overhead and computational cost. Therefore, in

addition to the literature methods, the data minimisation method can play an important role in providing security and privacy. Table 1 categorised research based on its security/privacy factors. The majority of research presented accuracy, encryption, and energy, as seen in this table.

### 3 | PROPOSED METHOD

The following text provides a comparative table to compare the proposed method to previous methods based on parameters such as cost, efficient data management, and efficient data storage. Thereafter, more detailed information is provided to clarify the proposed method.

A general comparison of the proposed method with the previous methods is shown in Table 2. According to this comparison, the proposed method could be more effective at preserving privacy than other methods. Another issue that can be a critical challenge for the Internet of Things systems in the long term is efficient data storage and management, which can be solved by the proposed method.

#### 3.1 | Data minimisation

Data minimisation is an efficient technique for ensuring privacy and can play a significant role in the rapidly expanding IoT. Privacy protection might be enhanced by proactively decreasing the quantity of data gathered and processed. As a result, interactions and processes related to software, information, communication technologies and system architecture should be left unidentified. Consistently identifiable and viewable personal information should be maintained to a

minimum wherever feasible. Limiting data release in accordance with the idea of data minimisation and purpose limitation IoT devices should minimise the quantity of data that leaves devices by converting raw data into aggregated data and removing raw data after the data necessary for processing has been gathered. As a general rule, deletion should occur at the closest point of raw data acquisition and, if feasible, immediately on the device.

Solution Data Minimisation seems to have received less attention. This type of solution can solve many problems of mass data. From the authors' point of view, Solution Data Minimisation needs further investigation and research.

The proposed data minimisation process is presented in Figure 2.

This method starts by gathering data on the system, that is, sensor data. After gathering data in the next step, the pre-processing phase starts, and for the process of transforming raw data into numerical features, Feature Extraction applies to the process. Processed data is an input of Creating Dataset step that is an important part of preparing a dataset for the machine learning mechanism. The resulting dataset and model process (to recognise certain types of patterns in the machine learning phase, a model over a set of data is essential) provides data on the process of running training data, and this training phase optimises the data to find certain patterns of data. After analysing the data, minimised and structured data is produced. This approach aims to extract the 'essential information' of the produced data as its knowledge without decreasing accuracy. Collecting essential data is not mean losing other data. The model remains accurate even though fewer data are collected.

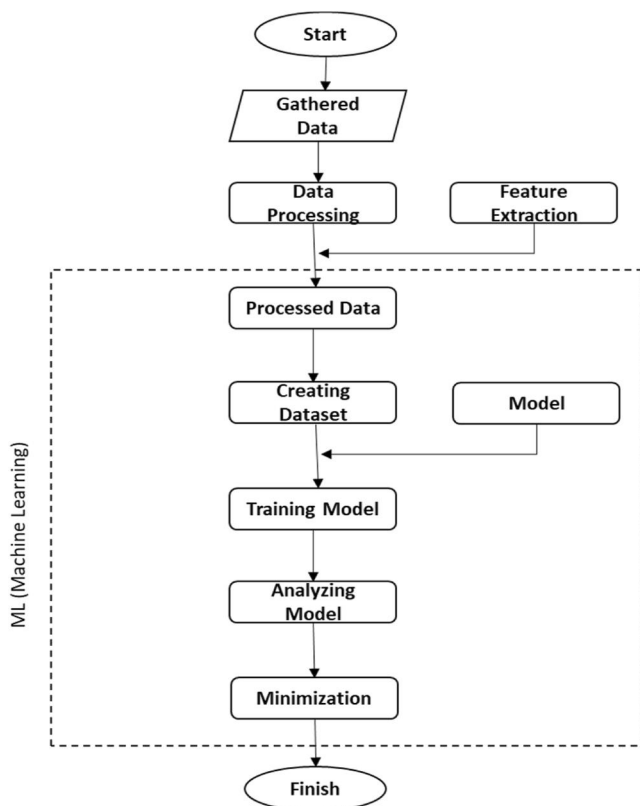
One of the efficient ways to prevent the accumulation of additional data is to collect and store essential data. In addition to reducing costs, data minimisation also reduces data volume.

Research	Scalability	Energy	Accuracy	Overhead	Data minimising	Encryption
Li et al. [6]	*		*		*	
Lin et al. [7]			*	*		*
Gheisari et al. [3]	*	*	*	*		*
Gheisari et al. [9]		*	*	*		
Lomotey et al. [11]	*	*		*		
Ren et al. [14]		*				*
Liu et al. [12]			*			*
Hossein et al. [16]	*		*		*	*
Qu et al. [18]		*	*			*
Bao et al. [10]				*		*
Singh et al. [19]				*		
Bao et al. [15]		*		*		*
Bao et al. [17]		*		*		*
Srivastava et al. [13]			*			*
Bao et al. [8]		*	*	*		

**TABLE 1** Comparison of the existing evaluation factors in studies that have been discussed

**TABLE 2** The comparison of literature papers and the proposed method

Research	Cost effective	Data management (efficient)	Storage (efficient)	Low latency	Non complex computing	Energy efficiency	Accuracy
Li et al. [6]		*					*
Lin et al. [7]			*		*		*
Gheisari et al. [3]			*			*	*
Gheisari et al. [9]				*		*	*
Lomotey et al. [11]					*	*	
Ren et al. [14]	*			*		*	
Liu et al. [12]							*
Hossein et al. [16]	*	*	*	*			*
Qu et al. [18]	*					*	*
Bao et al. [10]	*			*	*		
Singh et al. [19]			*		*		
Bao et al. [15]	*				*	*	
Bao et al. [17]	*		*		*	*	
Srivastava et al. [13]							*
Bao et al. [8]	*					*	*
Proposed	*	*	*	*	*	*	*

**FIGURE 2** Workflow of the proposed method.

It is impossible for a business to continue collecting and storing data indefinitely, as the cost of data storage is always going to be an expense [20]. The risks associated with too much data are also high (especially personally identifiable data). The consequences of data loss and breaches must also be considered. When sensitive personal information is leaked, a solution can easily collapse, or criminal charges may be brought. Moreover, it would be even more painful to lose data in the first place that did not even necessary.

## 4 | RESULTS AND DISCUSSION

This section discusses a statistical analysis of various studies on privacy-based IoT applications. So to answer the following context addresses the deterministic analytical questions mentioned in the previous section. Based on Table 1, the accuracy criteria are the most important aspect of privacy. A discussion of the benefits, problems, and shortcomings associated with the proposed method is also included.

### 4.1 | AQ1: What are the main reasons for choosing IoT as a powerful solution?

IoT is an important technology that is set to improve people's lifestyles by some great valuable benefits, including (a) Data

can be collected from many different sources through sensors, (b) Automate the process to reduce workload, (c) savings in time and resources will increase efficiency, (d) It helps manage important tasks because of a lack of time.

#### 4.2 | AQ2: What security/privacy methods have been used in IoT applications?

IoT's major security and privacy concerns are heterogeneity, authentication, and identification. In general, user data could be exposed to many difficulties and vulnerable to cyber-attacks due to poor security and privacy. Moreover, massive data is more profitable in this type of network. Several methods have been introduced in recent years, including Blockchain-based solutions, Machine Learning, Ontology, data anonymisation etc.

#### 4.3 | AQ3: What are the research gaps and open questions for future research directions?

In terms of security/privacy, user data could be a profitable resource for attackers to steal data. Due to the low level of security and privacy in many parts and layers of IoT, applications can be exposed to many difficulties, such as unauthorised access, Denial of service (DoS) attacks, and blocking wireless signals. Therefore, it is difficult to implement unless security/privacy issues are addressed. Strengthening encryption by using the latest developments is beneficial to enhancing cybersecurity. These technologies are incorporated into the various layers of the network. Thus, the unreadable data formats protected by encryption means that hackers cannot read and misuse them. Also, new methods such as Blockchain, Machine Learning, Data Minimisation, and Data Encryption can greatly improve privacy problems to ensure security and privacy. As we know, data generation with the expansion of IoT networks can be staggering, and having a suitable solution seems obvious. However, methods based on Data Minimisation can be a great help to these networks for privacy. The limitation of data leakage can be achieved directly by minimising data. Clearly, the less data collection, storing, and sharing by smart devices, the easier it will be to protect users' personal information. Processing the gathered data with Data Minimisation models could potentially help prevent prediction bias or other forms of misusing. It leads to more awareness of data accuracy, data privacy, data management, and efficient data storage. 'Data minimisation' can also mean limiting the scope of knowledge discovery.

The authors want to develop a method for addressing the problems of data minimisation using machine learning in the near future. Machine learning for diagnosis and therapy raises concerns about data privacy. As IoT gadgets and other devices generate more types of data, machine learning's ability to identify and analyse patterns in this data poses a substantial challenge to implementing and enforcing the principle of data minimisation. Virtually every form of data gathering and

processing may be rationalised as 'relevant to the target approach'. As a result, such a broad scope of data might evade the controls to decrease the risk that the data reduction principle intends to provide.

Looking for limits of this strategy, maybe a system's major difficulty while using data minimisation is identifying what sort of information is essential to accomplish the task. A data minimisation project can be undertaken using a variety of approaches. Some people may desire to improve efficiency by removing redundant, obsolete, or insignificant information that has no prospective value. Others, on the other hand, might just strive to minimise their storage footprint.

## 5 | CONCLUSION AND FUTURE WORK

Generally, this paper takes a precise look at the security and privacy-based models glimpsing into any current IoT-based systems using known secure cryptographic tools and technology and introducing efficient approaches to overcome some difficulties of security and privacy in IoT. This paper proposes a method for privacy protection with data minimisation based on background knowledge and an in-depth literature evaluation.

A careful evaluation of the existing research demonstrates that some achievements have been achieved, including:

- Secures both sensitive and public content
- Utilisation of various confidential controls to protect against the unauthorised use of information
- Reduction in operational charges

And some challenges remained

- Optimisation and reducing mass-produced data for gathering essential data
- Providing consumer confidence
- Requirement of high-standard solutions due to the increasing scale of data

In light of the existing challenges, the proposed method has the potential to address a wide range of issues. Using a minimisation process can further enhance privacy degree by collecting and aggregating data that is essential. This approach enables applications to truly minimise the amount of data collected without decreasing accuracy that, as a result, can support the mentioned indicators (Indicators of Table 2) well. Thus fulfilling the data privacy requirement set out.

Since the proposed method is in its early stages, the authors' main future direction for this work is to implement the system in a testable system to provide some guarantees of security and performance in the real world. After comparing and analysing the performance of the proposed method with other existing methods, it is expected that it can perform more optimally in preserving privacy than the previous methods.



## AUTHOR CONTRIBUTIONS

Mahdi Safaei yaraziz wrote the first draft of the manuscript, and all authors commented on earlier versions of the manuscript. All authors contributed to the idea and design of the study. The authors prepared the documents and collected and analysed the data. The final manuscript was read and approved by all authors.

## CONFLICTS OF INTEREST

The authors are not affiliated with any organisation having a direct or indirect financial interest in the subject dealt with in the manuscript.

## DATA AVAILABILITY STATEMENT

No data was used for this short communication.

## CONSENT TO PARTICIPATE

Informed consent was obtained from all individual participants included in the study.

## CONSENT FOR PUBLICATION

Participants' informed consent to publish their data and photos.

## ORCID

Mahdi Safaei Yaraziz  <https://orcid.org/0000-0002-1675-3619>

Mehdi Gheisari  <https://orcid.org/0000-0002-5643-0021>

Yang Liu  <https://orcid.org/0000-0003-2486-5765>

## REFERENCES

1. Fatchi, N., et al.: An automata algorithm for generating trusted graphs in online social networks. *Appl. Soft Comput.* 118, 108475 (2022). Elsevier. <https://doi.org/10.1016/j.asoc.2022.108475>
2. Hasan, M.: State of IoT 2022: Number of connected IoT devices growing 18% to 14.4 billion globally. *IoT Analytics*. <https://iot-analytics.com/number-connected-iot-devices/> Accessed 18 May 2022
3. Gheisari, M., et al.: OBPP: an ontology-based framework for privacy-preserving in IoT-based smart city. *Future Generat. Comput. Syst.* 123, 1–13 (2021). Elsevier. <https://doi.org/10.1016/j.future.2021.01.028>
4. Schneier, B.: IoT security: what's plan B? *IEEE Security & Privacy* 15(5), 96 (2017). IEEE. <https://doi.org/10.1109/MSP.2017.3681066>
5. Federal Trade Commission. [https://www.ftc.gov/system/files/document/s/reports/federal-trade-commission-2020-privacy-data-security-update/20210524\\_privacy\\_and\\_data\\_security\\_annual\\_update.pdf](https://www.ftc.gov/system/files/document/s/reports/federal-trade-commission-2020-privacy-data-security-update/20210524_privacy_and_data_security_annual_update.pdf) Accessed 24 May 2021
6. Li, Z., et al.: RR-LADP: a privacy-enhanced federated learning scheme for Internet of everything. In: *IEEE Consumer Electronics Magazine*. 10(5), 93–101 (2021). IEEE. <https://doi.org/10.1109/MCE.2021.3059958>
7. Lin, L., Tian, Y., Liu, Y.: A blockchain-based privacy-preserving recommendation mechanism. In: *2021 IEEE 5th International Conference on*

8. Cryptography, Security and Privacy (CSP). 74–78 (2021). IEEE. <https://doi.org/10.1109/CSP51677.2021.9357604>
9. Bao, Y., et al.: Efficient, revocable, and privacy-preserving fine-grained data sharing with keyword search for the cloud-assisted medical IoT system. *IEEE J Biomed Health Inform* 26(5), 2041–2051 (2022). IEEE. <https://doi.org/10.1109/JBHI.2021.3100871>
10. Gheisari, M., et al.: A context-aware privacy-preserving method for IoT-based smart city using Software Defined Networking. *Computers & Security*. 87, 101470 (2019). Elsevier. <https://doi.org/10.1016/j.cose.2019.02.006>
11. Bao, Y., Qiu, W., Cheng, X.: Efficient and fine-grained signature for IIoT with resistance to key exposure. *Internet of Things J.* 8(11), 9189–9205 (2021). <https://doi.org/10.1109/JIOT.2021.3055861>
12. Lomotey, R.K., et al.: Data verification and privacy in IoT architecture. In: *2019 IEEE World Congress on Services (SERVICES)*. 66–71 (2019). IEEE. <https://doi.org/10.1109/SERVICES.2019.00026>
13. Liu, Y., Zhang, J., Zhan, J.: Privacy protection for fog computing and the Internet of things data based on Blockchain. *Cluster Comput.* 24(2), 1331–1345 (2021). Springer. <https://doi.org/10.1007/s10586-020-03190-3>
14. Srivastava, G., Crichigno, J., Dhar, S.: A light and secure healthcare blockchain for IoT medical devices. In: *Canadian Conference of Electrical and Computer Engineering (CCECE)*. 1–5 (2019). IEEE. <https://doi.org/10.1109/CCECE.2019.8861593>
15. Ren, W., et al.: Privacy enhancing techniques in the Internet of things using data anonymisation. *Inf. Syst. Front* (2021). Springer. <https://doi.org/10.1007/s10796-021-10116-w>
16. Bao, Y., Qiu, W., Cheng, X.: Secure and lightweight fine-grained searchable data sharing for IoT-oriented and cloud-assisted smart healthcare system. *Internet of Things* 9(4), 2513–2526 (2021). IEEE. <https://doi.org/10.1109/JIOT.2021.3063846>
17. Hossein, K.M., et al.: BCHealth: a novel blockchain-based privacy-preserving architecture for IoT healthcare applications. *Comput. Commun.* 180, 31–47 (2021). Elsevier. <https://doi.org/10.1016/j.comcom.2021.08.011>
18. Bao, Y., et al.: Fine-grained data sharing with enhanced privacy protection and dynamic users group service for the IoV. *IEEE trans Intell Transp Syst.* 1–15 (2022). IEEE. <https://doi.org/10.1109/TITS.2022.3187980>
19. Qu, Y., et al.: Decentralized privacy using blockchain-enabled federated learning in fog computing. *Internet of Things* 7(6), 5171–5183 (2020). IEEE. <https://doi.org/10.1109/JIOT.2020.2977383>
20. Singh, R., et al.: Privacy-preserving ledger for blockchain and Internet of Things-enabled cyber-physical systems. *Comput. Electr. Eng.* 103, 108290 (2022). Elsevier. <https://doi.org/10.1016/j.compeleceng.2022.108290>
21. Sadowski, J.: When data is capital: datafication, accumulation, and extraction. *Big Data & Soc.* 6(1), 205395171882054 (2019). Sage. <https://doi.org/10.1177/2053951718820549>

**How to cite this article:** Safaei Yaraziz, M., et al.: Recent trends towards privacy-preservation in Internet of Things, its challenges and future directions. *IET Circuits Devices Syst.* 1–9 (2022). <https://doi.org/10.1049/cds2.12138>