Complex & Intelligent Systems (2023) 9:3655–3679
https://doi.org/10.1007/s40747-021-00489-5

ORIGINAL ARTICLE

A noise-based privacy preserving model for Internet of Things

Shelendra Kumar Jain[1] · **Nishtha Kesswani**[1]

Received: 11 September 2020 / Accepted: 27 July 2021 / Published online: 25 August 2021

Abstract
With the ever-increasing number of devices, the Internet of Things facilitates the connection between the devices in the hyper-connected world. As the number of interconnected devices increases, sensitive data disclosure becomes an important issue that needs to be addressed. In order to prevent the disclosure of sensitive data, effective and feasible privacy preservation strategies are necessary. A noise-based privacy-preserving model has been proposed in this article. The components of the noise-based privacy-preserving model include Multilevel Noise Treatment for data collection; user preferences-based data classifier to classify sensitive and non-sensitive data; Noise Removal and Fuzzification Mechanism for data access and user-customized privacy preservation mechanism. Experiments have been conducted to evaluate the performance and feasibility of the proposed model. The results have been compared with existing approaches. The experimental results show an improvement in the proposed noise-based privacy-preserving model in terms of computational overhead. The comparative analysis indicates that the proposed model without the fuzzifier has around 52–77% less computational overhead than the Data access control scheme and 46–70% less computational overhead compared to the Dynamic Privacy Protection model. The proposed model with the fuzzifier has around 48–73% less computational overhead compared to the Data access control scheme and 31–63% less computational overhead compared to the Dynamic Privacy Protection model. Furthermore, the privacy analysis has been done with the relevant approaches. The results indicate that the proposed model can customize privacy as per the users' preferences and at the same time takes less execution time which reduces the overhead on the resource constraint IoT devices.

Keywords Internet of Things · Privacy · Privacy preservation · Sensitive information · Privacy threats in IoT

Introduction

The amalgamation of various technologies like sensor communications, cloud computing, Internet of Things (IoT), artificial intelligence, machine and deep learning plays a vital role in the smart world [1]. IoT is a prevailing technology capable of morphing human lives by providing ease and smartness in varied conventional application domains. As shown in Figs. 1, 2, and 3, IoT is a hybrid environment that is a combination of many technologies such as sensing, data storage, data analytics, and connectivity of things. Further, IoT extends the capabilities of the physical things [2].

IoT applications like smart city, smart healthcare systems, smart building, smart transport and smart environment [3], industrial, agriculture, supply chain management [4], smart retail, location-based services, etc. may deal with sensitive data such as health information, financial information [5], location footprints, Personally Identifiable Information (PII) [6], data of personal life, etc. Data deluge from billions of entities producing information is a significant threat to privacy [7] (Fig. 4).

Privacy is the right of individuals, which helps them keep their information secret and have control over their information [8]. Privacy preservation is an important aspect that must be considered in every existing logical and physical system to reduce the possibilities of privacy breaches. Ensuring Information privacy is an increasing concern for government, business, consumer, and likewise [9]. In IoT-based networks, personal information is collected from smart devices, and weak privacy measures can misuse sensitive information. If this personal information is stolen, then results can be detrimental [10].

Some of the significant privacy challenges in IoT are as follows:

✉ Shelendra Kumar Jain
shelendra23@hotmail.com

[1] Department of Computer Science, Central University of Rajasthan, NH-8, Bandar Sindri, Dist-Ajmer, Rajasthan 305817, India

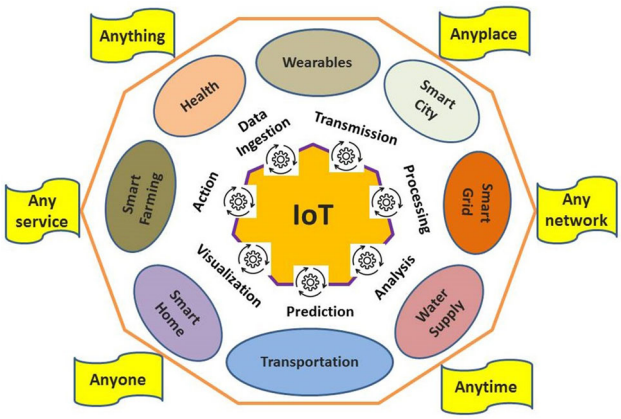

Fig. 1 Introduction to the Internet of Things and applications

(1) What private data are sensed, where is this data stored, how and who uses the data? [10]
(2) Automate the process of identification of sensitive and non-sensitive data.
(3) How to allow users to control and manage their data, maintain user's anonymity, and preserve the data integrity in each phase of the data's life cycle? [5]
(4) Implementation of efficient mechanism that is suitable for pervasive infrastructure and resource-constrained IoT devices [11].

Many researchers have emphasized that privacy and security are the most challenging problems in IoT because of the risk associated with leakage of the user's private information from several IoT services [12]. Data protection by design and by default (or privacy by design) is crucial to address privacy and protection of data [13]. Users will accept IoT-based systems only if they are secure, trustworthy, and privacy is preserved [8]. Users must be equipped with tools to retain their anonymity in an IoT-based connected world [7]. Thereby, in an IoT environment, an efficient and well-planned strategy is necessary to preserve privacy. The novelties and contributions of this paper as follows:

(1) A Multilevel Noise Mechanism has been proposed for data collection to ensure privacy preservation in the Internet of Things environment.
(2) A user preferences-based data classifier has been proposed to classify sensitive and non-sensitive data in the Internet of Things environment.
(3) Noise Removal and Fuzzification Mechanism has been proposed for data access to ensure privacy preservation in the Internet of Things environment.

The remainder of this paper is organized as follows: "Related work and motivation" describes related work and motivation. "Adversary model and design objectives" presents adversary model and design objectives. The noise-

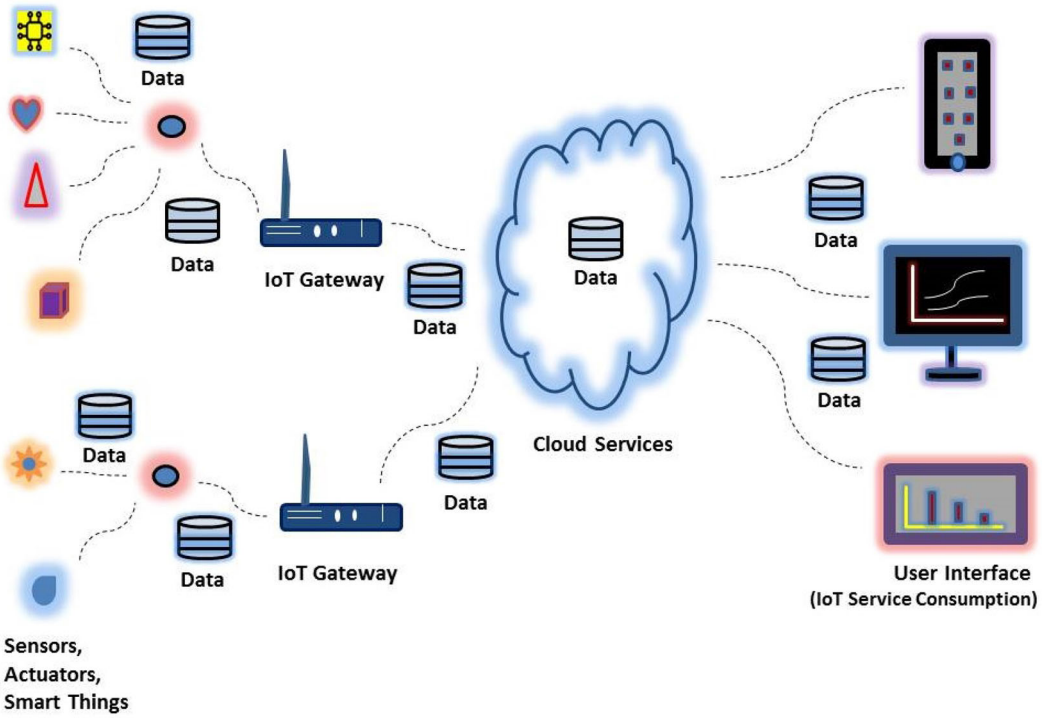

Fig. 2 Major components of Internet of Things

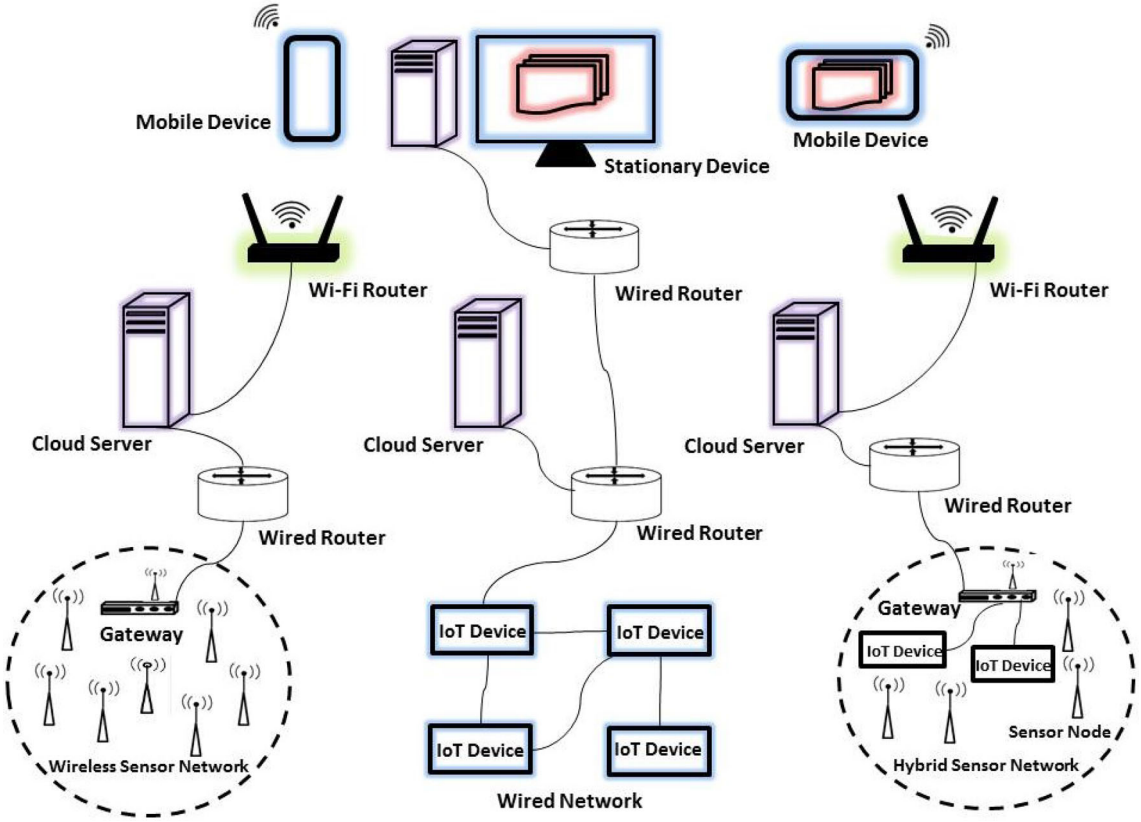

Fig. 3 A typical architecture the IoT environment

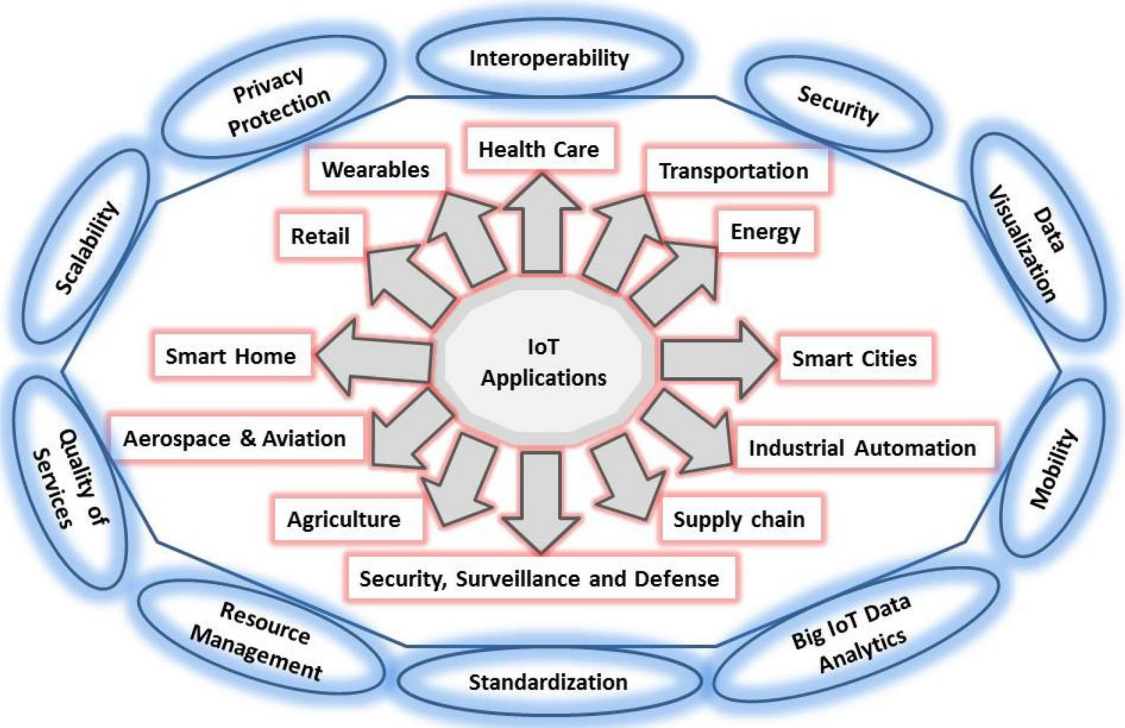

Fig. 4 Application domains and research challenges of IoT

based privacy-preserving model is described in "Noise Based Privacy Preserving model". The experiments and results are given in "Experiments and results", and "Limitations and future scope" concludes the paper.

Related work and motivation

The consumer's trust can be enhanced by privacy preservation in IoT, and it can be achieved by fulfilling the privacy requirements at data generation, storage, usage, and sharing [10]. Ziegeldorf et al. [14] analyzed the privacy issues, discussed the evolving features and trends in IoT, and classified privacy threats. According to the survey [15], more research needs to be done to ensure security and privacy for the IoT paradigm's success. With the miniature power sources, small memory, limited processing capability, and incredibly resource-constrained IoT devices [16], User privacy and data protection, authentication and identity management, trust management, policy integration, authorization and access control, end-to-end security, etc. are security and privacy challenges in the IoT that need to be addressed (Tables 1 and 2).

The personal data collection and usage of these data are challenges to individual privacy in the IoT [17]. Corcoran [18] has introduced different privacy classes and outlined some ideas for improved privacy framework for IoT, such as; data should be protected at the data source. For the mitigation of heavy computation constraints due to cryptographic operations in the sensors used in medical applications, Moosavi et al. [19] proposed a Secure and Efficient Authentication and Authorization (SEA) Architecture perform authentication and authorization on behalf of the medical sensors by the distributed smart e-health gateways. SEA architecture is based on the fact that various heavy-weight security protocols and certificate validation efficiently can be handled by smart e-health gateway and the remote end-user because both have sufficient resources.

Appavoo et al. [20] proposed a privacy-preserving model to prevent service providers from revealing sensed values, sensor types, and user preferences. The proposed work can be considered as a simple form of functional encryption. A case of a semi-trusted service provider has been considered. In this work, the author represented privacy loss (Eq. 1) in the form of mutual information [21].

$$I(\Psi, V; \delta) = H(S, V) - H(\Psi, V|\delta), \tag{1}$$

Ψ, V, and δ are the random variables for the set of sensors that can be utilized, the set of sensed values, and the set of outcomes for the trigger conditions, respectively. H (S, V) represents the maximum information that can be predicted for sensors and their values.

Turgut and Boloni [22] have concentrated on the value and cost of data exchange in IoT with the other types of cost. They described an exciting relationship between the value of information and the cost of privacy (customer's benefit from Eq. 2 and business benefit from Eq. 3) for the IoT paradigm's success. The definition of the notations used in these equations is given in Table 4.

$$\eta_{\text{service}} - \sigma_{\text{privacy}} - \sigma_{\text{hardware}}^{\text{user}} - \sigma_{\text{payment}} > 0, \tag{2}$$

$$\rho_{\text{information}} - \sigma_{\text{hardware}}^{\text{business}} + \sigma_{\text{payment}} > 0. \tag{3}$$

As a notion that trust can be directly related to privacy [23], Butun [24] mapped privacy and trust relation by integrating multi-dimensional relationship of the sensitivity level of PII items, privacy, and trust (Eq. 4).

$$\Gamma(\phi; \varepsilon, \Omega, \pi) = \frac{1}{1 + e^{-(-\varepsilon(\phi - \pi\Omega))}}. \tag{4}$$

Jayaraman et al. [25] introduced privacy-preserving IoT architecture and data ingestion scheme in which produced IoT data are split into R parts, where R is the number of servers. If a jth datum produced by an IoT device is D and the number of servers is three (R=3), then it will be split into data addends, namely α_{1j}, α_{2j} and α_{3j}, where

$$D_j = \sum_{i=1}^{R} \alpha_{ij}. \tag{5}$$

Along with privacy-preserving IoT architecture, Jayaraman et al. also proposed a privacy pre serving data access scheme based on the Paillier cryptosystem's homomorphic properties (Tables 1, 2).

The Dynamic Privacy Protection (DPP) model [26] is designed to ensure mobile device user privacy. DPP model generates a privacy protection plan to determine the security mode for each data or data package. In this model, privacy protection levels are classified based on privacy weight. Total privacy weight $\mathbb{P}$ is calculated using Eq. (6). In this equation, $N^e(D_i)$ is the number of data or data packages (D_i) that use higher-level security mode, and $N^n(D_i)$ is the number of data or data packages that use lower-level security mode. If values of binary function s(i) = 1, then encryption will be used and if s(i) = 0 then non-encryption will be used.

$$\mathbb{P} = \sum_{S_{(i)=1}} N^e(D_i) \times W^e(D_i) + \sum_{S_{(i)=0}} N^n(D_i) \times W^n(D_i). \tag{6}$$

Many researchers have tried to address security and privacy issues in the Internet of Things. Several privacy preservation techniques for IoT have been proposed, but to

Table 1 Basic concepts used for privacy preservation in the various existing frameworks/approaches

Framework/approach	Basic concept for privacy preservation	References
Lightweight encryption algorithm	Encryption/decryption	[27]
Privacy-preserving IoT architecture	Encryption/decryption	[25]
DPP model	Selectively encrypt data	[26]
EPIC	Differentially Private (DP) obfuscation mechanism	[28]
Privacy-preserving model	Trust evaluation	[29]
Privacy-preserving trust model	Functional encryption/decryption	[20]
Information relevance model	Contextual privacy perception framework	[30]
Interaction-based privacy protection management framework	Restricting the non-authorized operations and neutralizing the execution of non-authorized operations	[31]
Privacy monitoring framework	Informative event, access log analyzer, obfuscation	[32]
Privacy preserving communication protocol	Chaos-based cryptographic scheme and message authentication codes	[33]
Balance privacy-preserving data aggregation model	Slicing and mixing technology	[34]
Privacy preserving scheme	Identity-based Encryption (IBE) and symmetric encryption	[35]

the best of our knowledge, only a little research work has been carried out to ensure end-to-end privacy, i.e., privacy preservation in all the layers in the IoT ecosystem, along with implementation and detailed results analysis. Also, Many proposed privacy-preserving frameworks are based on cryptographic operations. Many of the existing frameworks have not included data classifier mechanisms and user customization-based privacy preservation. Many of the existing work on IoT privacy has not considered the trade-off between privacy and quality-of-service in the practical scenario. This paper has addressed these issues, presents a systematic flow of IoT data, and implements and analyzes the Noise-Based Privacy-Preserving model (NBPPM model). The proposed model's novelty is that it ensures data privacy with fair efficiency at all the layers (edge layer, middleware, and application layer) of the IoT ecosystem.

Adversary model and design objectives

This section is focused on various privacy threats associated with IoT. In the adversary model, it has been assumed that an adversary is well equipped to monitor communication channels. Any malicious insider at the data storage level (such as a rogue administrator) can access sensitive and non-sensitive data, analyze data and make inferences to gain advantages. An unauthorized user can access sensitive data at the application level, and a service provider can access user data to provide services to the user.

As an example of inference threat in IoT based healthcare application, let us assume a universal set of sensors in IoT is $X = \{s_1, s_2, s_3, \ldots s_n\}$ where n is number of sensors in the IoT based system and a universal set of location of these sensors is $L = \{l_1, l_2, l_3, \ldots l_n\}$. A set for data produced by the sensors in set X is $D = \{d_1, d_2, d_3, \ldots d_n\}$. If a set of different m kinds of diseases is $Y = \{y_1, y_2, y_3, \ldots y_m\}$. An adversary well equipped with tools and malicious intention can draw fruitful inferences by employing following inference rules in the inference attack:

$$
\begin{aligned}
R_1 &: (d_1 \pm a_1) \wedge (d_2 \pm a_2) \wedge (d_3 \pm a_3) \wedge \ldots \wedge (d_n \pm a_n) \rightarrow y_1 \\
R_2 &: (d_1 \pm b_1) \wedge (d_2 \pm b_2) \wedge (d_3 \pm b_3) \wedge \ldots \wedge (d_n \pm b_n) \rightarrow y_2 \\
R_3 &: (d_1 \pm c_1) \wedge (d_2 \pm c_2) \wedge (d_3 \pm c_3) \wedge \ldots \wedge (d_n \pm c_n) \rightarrow y_3 \\
&\vdots \\
R_j &: (d_1 \pm k_1) \wedge (d_2 \pm k_2) \wedge (d_3 \pm k_3) \wedge \ldots \wedge (d_n \pm k_n) \rightarrow y_m
\end{aligned}
$$

where $a_1, \ldots a_n$, $b_1, \ldots b_n$, $c_1, \ldots c_n$ and $k_1, \ldots k_n$ are constants used to form specific ranges for the derivation of a useful inference rule. For example, through the above inference rules, an eavesdropper can infer patient disease, which may be private information for the patient, and through location set L, linkage-based attack can be performed, i.e., $\{(d_1, l_1), (d_2, l_2), (d_3, l_3), \ldots (d_n, l_n)\}$. It can result in physical, mental, economic, and social exploitation of the victim.

Security and privacy threats in IoT

An overview of the major security and privacy threats [14,38–40] in the IoT environment is mentioned in Table 3.

Problem definition and design objectives

The critical research problem is defined as developing a systematic model to ensure end-to-end privacy against various threats for resource-constrained IoT environments. As

Table 2 Key parameters, challenges, important findings of the existing studies

Framework/approach	Key parameters or building block	Challenges/issues for that solution is proposed	Important findings	References
Lightweight encryption algorithm	Hash function SHA-3 Symmetric key cryptosystem	Needs of a practical strategy to prevent the inside attack	The lightweight encryption algorithm that protects the communication among the sensor-node and the Sharemind system preserve the patient data privacy if the three data servers in the Sharemind system do not collude	[27]
Privacy preserving IoT architecture	Data ingestion scheme splits the IoT data into n (number of servers) parts	Lack of control over IoT devices Privacy loss over IoT devices, storage infrastructure, applications, and related communications Developing techniques that can ensure privacy in the IoT data collection, storage, and retrieval	Innovative schemes for privacy-preservation of the IoT data	[25]
DPP Model	Privacy weight Dynamic programming Selectively encrypt data, based on the requirements and constraints of the associated hardware or software	User's privacy Violation when different data are combined Without incurring unrealistic performance overheads, ensuring the security of data in transit and at rest	Uses the content-oriented approach to selectively encrypt data for privacy protection	[26]
EPIC	Utility optimal differential privacy mechanism	Protecting from the traffic analysis attacks due to resources constrain	A privacy-preserving traffic obfuscation framework Adversaries cannot link any traffic flow to a particular smart home	[28]
Privacy-preserving trust model	Trust and uniformization models	Minimizing the privacy-loss in the presence of untrusted service providers	A lightweight approach to functional encryption	[20]
Privacy-preserving model	Based on simple threshold detection Direct interactive trust, friend recommendation trust and historical trust Dynamic self-adjusting trust evaluation approach	How to build a trust model that can prevent non trusted objects from accessing private data	A lightweight strategy to access control for privacy-preservation Privacy protection problem is transformed into a simple judgment problem	[29]
Information relevance model	Consumer's privacy sensitivity as the summation of their privacy concerns Population privacy sensitivity	To treat privacy uniformly is unfair and socially inefficient by which a substantial proportion of the population remains unsatisfied by a common-policy	Acknowledged the existence of individual differences with respect to unique security and privacy protection needs Contribute to quantifiable means to measure and evaluate the customized privacy	[30]

Table 2 continued

Framework/approach	Key parameters or building block	Challenges/issues for that solution is proposed	Important findings	References
Privacy monitoring framework	Informative events and access log analyzer Average response time	For the broader adoption of cloud computing, the necessity of proper privacy and security mechanisms to control the sensitive information committed to cloud service providers by users	The framework provides a mechanism that enables cloud customers to track details, such as what happens to their data, where data is stored, and who accesses their data	[32]
Privacy preserving communication protocol	Symmetric encryption scheme	Eavesdroppers can aggregate the traffic information to profile a household RFID tags, sensors, actuators, and central Controller are known for limited computing capabilities and not capable of carrying out complicated computing operations	A lightweight secure and privacy-preserving communication protocol that leverages chaos-based encryption and Message Authentication Codes (MAC)	[33]
BPDA Model	Slicing and mixing technology	Sensitive information that sensor nodes gathered is inclined to be leaked for the hostile environment	Good performance in terms of privacy-preserving efficacy and communication overhead and increases the lifetime of the network	[34]
Privacy preserving scheme	IBE scheme and symmetric encryption	The exchanged data including sensitive and critical information are sent via an insecure channel Content and contextual privacy requirements must be satisfied	Privacy preservation solution for E-health fulfilling privacy requirements	[35]
IoTp	Data Masking technique and Distributed Approach	Lack of end to end privacy Linking data collected from sensors	Ensures privacy at data collection, data store and data access	[36]
SPT	Data splitting with the data obfuscation	Computational constraint, storage cost, and battery power are the major issues Ensures data privacy in the IoT ecosystem through lightweight data collection and data access protocols in the resource-constrained IoT ecosystem Ensure end to end data privacy efficiently	Ensure the data privacy with lightweight approaches for the resource constrained IoT devices	[37]
Noise based privacy-preserving model	Multilevel noise treatment Fuzzification	Privacy breach in data collection, storage and Retrieval Practical and feasible privacy preservation strategies	ensures privacy preservation according to the user's preferences Less burden on the resource constraint IoT devices	Proposed

Table 3 Overview of various security and privacy threats in the IoT

Threats at different levels				
IoT Node (Sensing device)	Gateway	IoT network	Cloud level	Application level
Hacked IoT node	Single point attack	Sniffing attack	Malicious administrator (insider threat)	Lack of user control over their data
Lack of control by legitimate user over IoT node	Inferences	Traffic analysis	Single point attack	Malicious apps
Privacy violating interaction and presentation	Linkage	Sybil attack	Inferences	User unawareness
Lifecycle transitions	Hacked gateway		Linkage	Profiling
Tracking			Lack of user control	Permission escalation
Inventory attack				Collusion attack
				Tracking

the components of IoT such as sensors, actuators, etc. have limited computing capabilities and are not suitable for performing complex computing operations [33], our objective was to plan and develop a model against privacy threats and incorporate privacy preservation characteristics such as to safeguard sensitive information, data access control, query privacy, and user-based privacy customization. Along with privacy preservation, our main objective was to reduce the computational overhead for resource-constrained IoT environments.

Noise based privacy preserving model

This section presents the proposed noise-based privacy-preserving model. The methodology with the structural diagram and detailed functioning of all modules involved in the NBPPM model have been described.

Overview

Let us assume a typical IoT environment consists of IoT devices, middleware, data storage, and user devices with apps that consume service providers' services. The components of the NBPPM model are shown in Fig. 5. Data produced from a source device must be protected in-transit, in-process, and at rest from an intruder that may exist between a source device and a legitimate user device. This goal is achieved in the proposed NBPPM model by incorporating noise while data move from the data source to data storage and denoising the noise at the user device. The proposed model also incorporates the fuzzification mechanism for privacy customization. Thus, the proposed NBPPM uses twofold privacy preservation using noise and fuzzification.

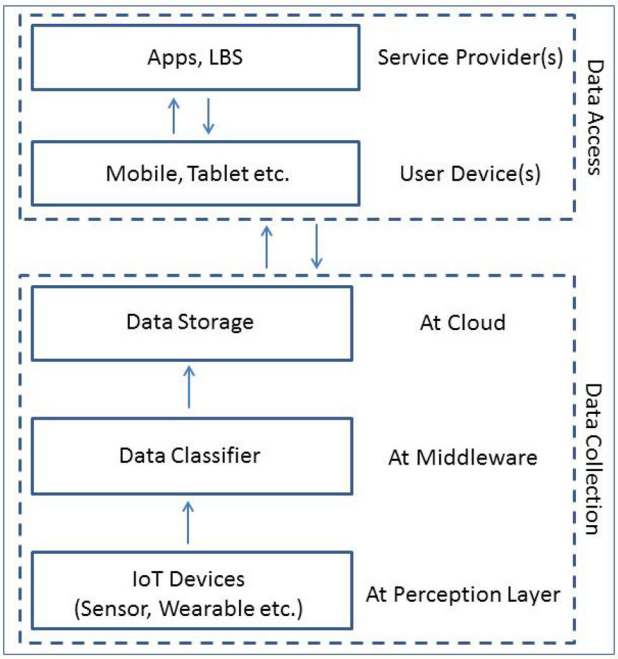

Fig. 5 Overview of the core components in IoT for the NBPPM model

Methodology

The proposed NBPPM model's fundamental modules are the data classification module, multilevel noise treatment module, and noise removal and fuzzification module. In this subsection, each module has been described comprehensively. The overall layout of the proposed model is shown in Fig. 6.

As shown in the proposed methodology's flowchart (Fig. 7), level 1 noise is added to all types of data (i.e., sensitive and non-sensitive data). After the level 1 noise addition, data splitting is performed on each data. A data classifier synchronized with the user customization setting performs data

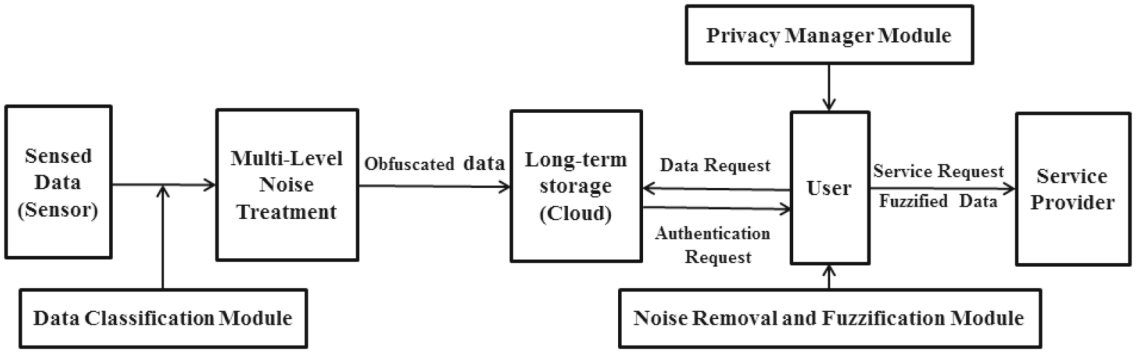

Fig. 6 Overall layout of the proposed model

classification according to the user preferences. If the data attribute is sensitive, then data addends proceed for level 2 and level 3 noise addition. If the data attribute is non-sensitive, then data addends proceed for level 3 noise addition (Algorithm 1). All of these noised data addends are stored in the data repository (i.e., Cloud Storage). An authenticated user can access noised data addends using valid credentials. At the user end, data addends are de-noised using the noise removal process (Algorithm 2). Further, if a service provider requests users' data to provide services, the service user can supply fuzzified data (based on the user privacy preferences) to the service provider (Fig. 9).

Data classification module

A data classification mechanism is a necessary step before incorporating a privacy protection mechanism. The data classification mechanism acts as a classifier to categorize data into two classes: sensitive and non-sensitive data class. One of the major issues for data classification is who and how it is decided which data attribute is sensitive and non-sensitive. The data owner is the best entity that can decide the sensitivity of his/her data for an IoT environment. In our proposed data classification mechanism, a data owner can customize his/her data privacy by setting attribute sensitivity to sensitive and non-sensitive mode at the application level, and from the application level, it will be synchronized with the data classifier module. Depending upon the sensitivity of the data, it is treated to multiple levels of noise. Further, at this point, an alternative policy can also be adopted for the data classification by considering an application-specific scenario, i.e., an IoT environment in which some of the data owners cannot judge data sensitivity correctly or may not have any knowledge about the data sensitivity. In this case, a predefined data sensitivity can be added. This predefined data sensitivity can be decided according to specific IoT applications and the General Data Protection Rules and Regulations of the particular country. For instance, in the IoT healthcare system, blood glucose level, heart rate, respiration rate, blood pressure, body temperature can be put in the sensitive category of data, and room temperature and humidity can be considered under the non-sensitive data category. A hybrid policy can also be deployed, combining predefined data classification and user-defined privacy preferences. Therefore, a user can change predefined settings according to his/her personal privacy preferences in the IoT ecosystem.

Multilevel noise treatment

In the multilevel noise treatment module of the NBPPM model, noise acts as a private key for the user. A random number generation algorithm is used to generate and divide noise into sub-noises. Let P be the generated noise; then P will be divided into three sub-noises P_1, P_2, and P_3 through a random number generation algorithm at the user end. Each sub-noise P_1, P_2, and P_3 is privately shared with the Data-Source, middleware, and data storage server, respectively.

Data splitting and multilevel noise treatment are two critical steps of the NBPPM model, as shown in Fig. 7. Each datum sensed D in the IoT environment is treated with sub-noise P_1 at level 1 from an operator, picked out from the operator table for the sensed data of particular attribute type F_i (Table 5). Operator selection for level 1 sub-noise is based on modulo operation with the Data Identifier, i.e., from Qth position, where N=9 for Table 5. After the treatment of level 1 noise, resultant data is split into three data addends, namely X, Y, and Z. Data classifier module checks data addends X, Y, and Z for sensitivity. If these data addends are part of a sensitive attribute type data, then each of the data addends will be treated with level 2 and level 3 sub-noises. If the data addends are parts of a non-sensitive attribute, then each data addend will pass through level 3 sub-noise treatments only. For instance, as shown in Fig. 7, the sensed data D are treated with noise P_1 at level 1, and then resultant data are split into three data addend, namely $(X, Y, Z)_{F_i}$. Then data classifier checks the sensitivity of attribute type F_i. If the F_i is sensitive attribute type, then $(X, Y, Z)_{F_i}$ will be treated with noise P_2 and P_3 resulting into $(A, B, C)_{F_i}$ and $(K, L, M)_{F_i}$,

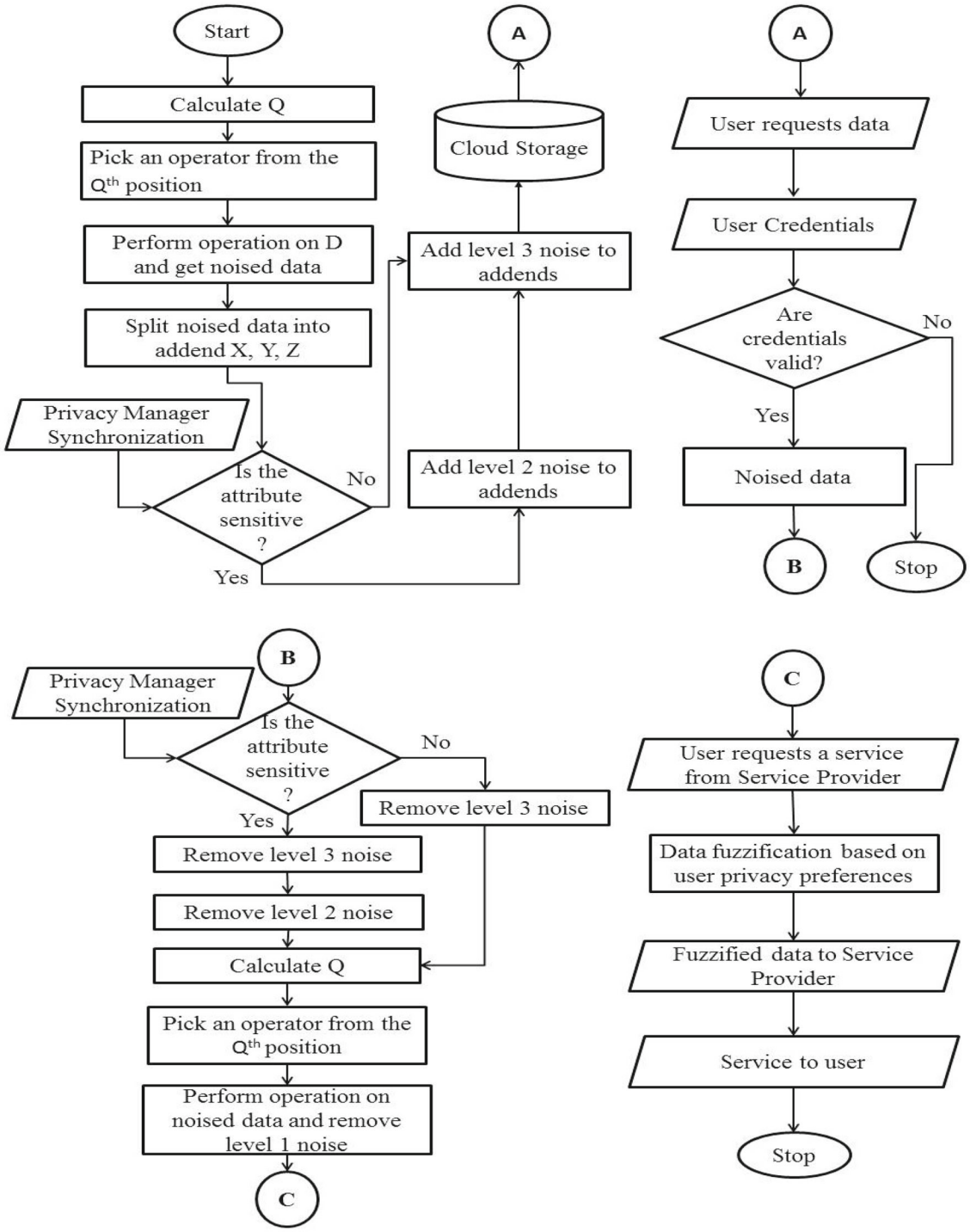

Fig. 7 Flowchart of the proposed methodology

respectively. If F_i is non-sensitive, then $(X, Y, Z)_{F_i}$ will be treated with noise P_3 resulting into $(K', L', M')_{F_i}$. Both $(K, L, M)_{F_i}$ and $(K', L', M')_{F_i}$ are stored in the long-term storage or the cloud (Fig. 8).

Noise removal and fuzzification

Noise removal at the user device is a reverse mechanism of the Multilevel Noise treatment mechanism. In an IoT environment, the user requests a service from the service provider. In order to provide the service, user data are requested from

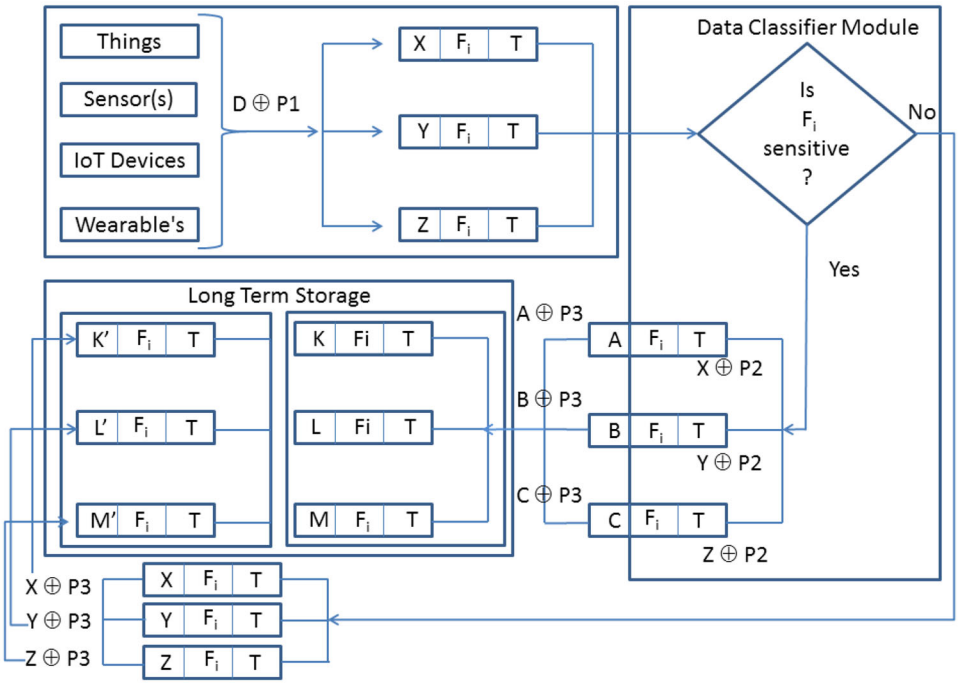

Fig. 8 Multilevel noise treatment methodology of the NBPPM model

Algorithm 1 Multilevel Noise Treatment Mechanism

Input: *Sensed IoT parameter (D), Data Identifier (D_{ID}), Timestamp (T), Attribute Type (F_i), N*
Output: *Noised IoT data addends*

1: Start
2: $Q \leftarrow D_{ID} \bmod N$
3: For F_i attribute Pick an operator $\oplus_i$ from Q^{th} position from operator table
4: $D \leftarrow D \oplus_i P_1$
5: Generate two data addends X and Y randomly
6: $Z \leftarrow D - X + Y$
7: Forward X, Y, Z to level 2 through secure channel
8: **if** $Sensitivity(F_i) == True$ **then** ▷ Senstivie Data
9: $A \leftarrow X \oplus_2 P_2$
10: $B \leftarrow Y \oplus_2 P2$
11: $C \leftarrow Z \oplus_2 P2$
12: Forward A, B, C to level 3 through secure channel
13: $K \leftarrow A \oplus_3 P_3$
14: $L \leftarrow B \oplus_3 P_3$
15: $M \leftarrow C \oplus_3 P_3$
16: **else** ▷ Non-sensitive Data
17: Forward X, Y, Z to level 3 through secure channel
18: $K' \leftarrow X \oplus_3 P_3$
19: $L' \leftarrow Y \oplus_3 P_3$
20: $M' \leftarrow Z \oplus_3 P_3$
21: **end if**
22: Stop

the service provider. As shown in Fig. 9, the user accesses the requested data from long-term data storage through valid user credentials. In the proposed NBPPM model, the authentication mechanism is incorporated to verify user validity through username and password. A valid user can access noisy data through a secure channel, and then the noise removal process is initiated through sub-keys, which act as the private key for the user. The process of Noise removal and fuzzification is shown in Algorithm 2.

Privacy is ensured through the fuzzification process when data are transferred between the user and the service provider. A sub-module, termed as privacy manager shown in Fig. 9, plays a vital role in user privacy customization. A fuzzifier sub-module is synchronized with user privacy preferences. A user can set his/her privacy preferences for a particular service, and accordingly, the fuzzifier decides the quality for data to be sent to access a service.

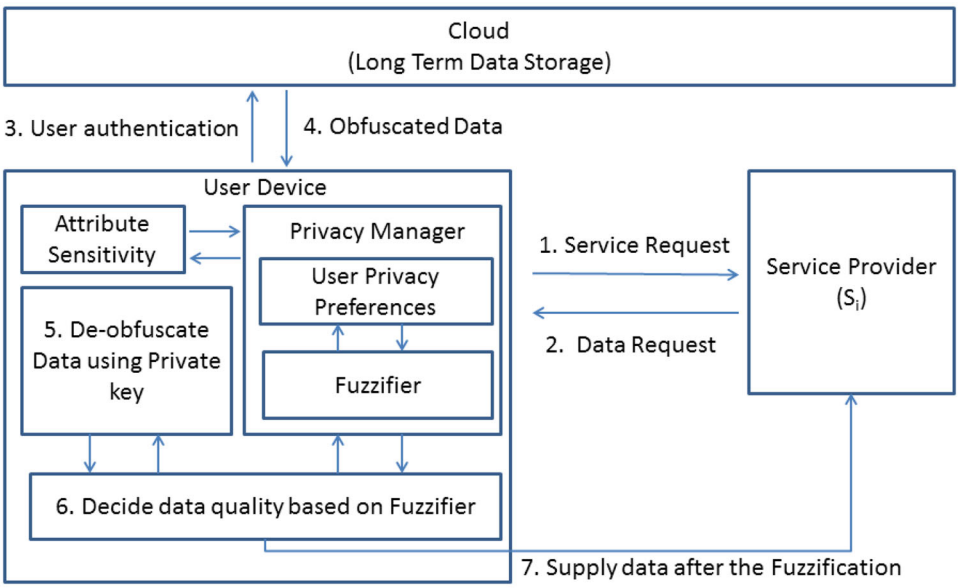

Fig. 9 Noise removal and fuzzification methodology of the NBPPM model

Algorithm 2 Noise Removal and Fuzzification Mechanism

Input: *Data request for IoT parameter (D) containing Data Field Name Identifier (F_i), timestamp and unique username*
Output: *Fuzzified IoT parameter (D')*

Start
if ($Username$ and $Password$) == True) **then**
  if Sensitivity (F_i) == True **then**
    User receives K, L, M
    $A \leftarrow K \ominus_3 P_3$
    $B \leftarrow L \ominus_3 P_3$
    $C \leftarrow M \ominus_3 P_3$
    $X \leftarrow A \ominus_2 P_2$
    $Y \leftarrow B \ominus_2 P_2$
    $Z \leftarrow C \ominus_2 P_2$
    $S \leftarrow X + Y + Z$
    $Q \leftarrow D_{ID} \bmod N$
    Pick operator $\oplus$ from Q^{th} position from operator table
    $D \leftarrow S \ominus_1 P_1$
  else if Sensitivity (F_i) == false **then**
    User receives K', L', M'
    $X \leftarrow K' \ominus_3 P_3$
    $Y \leftarrow L' \ominus_3 P_3$
    $Z \leftarrow M' \ominus_3 P_3$
    $S \leftarrow X + Y + Z$
    $Q \leftarrow D_{ID} \bmod N$
    Pick operator $\oplus$ from Q^{th} position from operator table
    $D \leftarrow S \ominus_1 P_1$
  end if
end if
$SL \leftarrow$ data sensitivity level
$F \leftarrow$ required level of fuzzification with respect to SL
$D' \leftarrow D \pm F$
Forward D' to S_i
Stop

Table 4 Summary of notations

Symbol	Meaning
D	Sensed IoT parameter
D_{ID}	Data identifier
N	Total operators in a row in operator table
T	Timestamp
F_i	Attribute Identifier
P	Noise
P_1	Sub-noise 1
P_2	Sub-noise 2
P_3	Sub-noise 3
X, Y, Z	Data addend at level 1
A, B, C	Data addend at level 2
K, L, M	Data addend at level 3 (Sensitive)
K', L', M'	Data addend at level 3 (non-sensitive)
$\oplus_1$	An operator from operator table
$\oplus_i$	An operator to add noise at ith level
$\ominus_i$	Reverse operator of $\oplus_K$
D'	Fuzzified data
S_i	ith service to user
η_{Service}	Value of the service received by the user
σ_{Privacy}	Cost of the privacy loss
$\sigma_{\text{Hardware}}^{\text{User}}$	User share in the cost of hardware and related service
σ_{Payment}	Payment made by the user for the service
$\rho_{\text{information}}$	Value of information collected by the provider
$\sigma_{\text{Business}}^{\text{Hardware}}$	Share of the business for hardware and maintenance cost
Γ	Trust value
ϕ	User privacy preference coefficient
ξ	Sensitivity coefficient for personally identifiable information items
Ω	Privacy coefficient
π	System trust coefficient
$(t_j)_{F_i}$	Execution time to access jth content of F_i attribute type
ω	Computational time

Table 5 An example of an operator Table

Attribute	0	1	2	3	4	5	6	7	8
F_1	+	−	*	+	*	−	+	−	*
F_2	*	−	+	−	*	+	*	−	+
F_3	−	*	+	−	+	*	−	*	+

A comprehensive overview of the functioning of the fuzzifier is as follows. As already defined, a universal set X over sensor domain as $X = \{s_1, s_2, s_3, \ldots s_n\}$. A user can set the sensitivity level for the data attribute of a sensor node (s_i) that senses the specific parameter value. Two fuzzy sets $\tilde{A}$ and $\tilde{\lambda}$ are defined as follows:

$\tilde{A}$ = 'Sensitive data' and $\tilde{\lambda}$ = 'Obfuscation quantity'. (7)

Membership function of $\tilde{A}$ and $\tilde{\lambda}$ are $\mu_{\tilde{A}}$ and $\mu_{\tilde{\lambda}}$, respectively, where $\mu_{\tilde{A}} \in [0, 1]$ and $\mu_{\tilde{\lambda}} \in [0, 1]$. Value of the membership function $\mu_{\tilde{A}}$ may be provided through an interface for the user. Value of $\mu_{\tilde{A}}$ indicates the level of the data sensitivity. Value of $\mu_{\tilde{\lambda}}$ indicates about the level of obfuscation. Membership value of the $\mu_{\tilde{\lambda}}$ will be decided through the value of $\mu_{\tilde{A}}$. i.e., $\mu_{\tilde{\lambda}}$ depends on $\mu_{\tilde{A}}$ and an illustrative example of the relationship between $\mu_{\tilde{A}}$ and $\mu_{\tilde{\lambda}}$ may be as follows (Eq. 7 and Table 6):

$$\mu_{\tilde{\lambda}} = f(x, \mu_{\tilde{A}}) = \begin{cases} 0 & \mu_{\tilde{A}} = 0 \\ \mu_{\tilde{A}} + c_1, & 0.1 \le \mu_{\tilde{A}} \le 0.4, 0.1 \le c_1 \le 0.4 \\ \mu_{\tilde{A}} + c_2, & 0.4 < \mu_{\tilde{A}} < 0.7, 0.1 \le c_2 \le 0.3 \\ 1 & \text{otherwise} \end{cases},$$

Table 6 An example for sensitivity level of data and corresponding level of data obfuscation

x	S_1	S_2	S_3	S_4	S_5	.	.	.	S_n
$\mu_{\tilde{A}}(x)$	0.8	0.2	0.6	0.3	0	.	.	.	0.1
$\mu_{\tilde{\lambda}}(x)$	1	0.3	0.8	0.6	0	.	.	.	0.3

SequenceNo.	X_acceleration	Y_acceleration	Z_acceleration	Activity
1	1667	2072	2047	1
2	1611	1957	1906	1
3	1601	1939	1831	1
4	1643	1965	1879	1
5	1604	1959	1921	1
6	1640	1829	1940	1
7	1607	1910	1910	1
8	1546	2045	1910	1
9	1529	2049	1972	1
10	1637	1978	1945	1
11	1596	2046	1866	1
12	1590	2006	1978	1
13	1601	1966	1957	1
14	1542	2003	1959	1
15	1598	2027	1941	1
16	1511	2258	1983	1
17	1555	1980	2023	1
18	1508	2468	1934	1
19	1580	1697	2005	1
20	1627	2073	1992	1
21	1592	2130	2063	1
22	1634	2088	1991	1
23	1638	2102	1916	1
24	1593	2123	1948	1
25	1542	2133	2034	1
26	1601	2015	2042	1
27	1613	1938	1936	1
28	1644	1974	2000	1
29	1642	1933	2046	1
30	1605	1925	2011	1
31	1586	1998	2066	1
32	1577	2032	2108	1
33	1598	1980	2066	1
34	1561	1942	2092	1
35	1628	1935	2142	1
36	1694	1965	2052	1
37	1627	1922	2081	1
38	1598	1950	2117	1
39	1612	1952	2075	1
40	1630	1958	2024	1

Fig. 10 Snapshot of the activity recognition dataset

where $x \in X$ and c_1 and c_2 can be fixed within a range and used to add the required quantity of the noise.

Experiments and results

The Noise-Based Privacy-Preserving Model has been presented comprehensively in "Noise Based Privacy Preserving Model". This section presents the experimental setup, findings of the experiment, performance evaluation, security, and privacy analysis to show how privacy can be protected through the proposed model.

	CALORIES BURNED	AVERAGE HEART RATE (bpm)	AVERAGE STEPS (per min.)	TOTAL STEPS	STRIDE LENGTH (cm)
2	445	135	132	8432	93
3	942	154	149	15804	74
4	514	128	128	9382	82
5	786	159	152	13682	89
6	872	151	142	16755	79
7	841	162	169	14188	74
8	993	158	159	18962	71
9	900	161	166	15399	69
10	1051	169	162	18048	72
11	726	115	115	15899	76
12	469	143	119	7601	87
13	851	151	150	11796	93
14	132	108	110	2783	78
15	1319	155	142	19682	87
16	1693	151	142	26930	83
17	951	109	108	14481	76
18	355	130	111	6201	86
19	401	140	121	6794	86
20	240	141	124	3793	84
21	410	129	110	6443	83
22	981	142	138	15574	84
23	430	137	122	7421	87
24	732	124	108	9028	83
25	320	128	122	3976	82
26	990	151	133	15570	84
27	1396	154	143	19813	86
28	732	148	138	11619	84
29	767	144	126	11299	87
30	578	111	116	12161	76
31	791	113	109	16485	76
32	319	107	108	6716	79
33	329	102	121	5944	93
34	686	123	114	13441	80
35	960	140	156	15936	84
36	504	146	149	8351	81
37	609	141	155	10105	84
38	985	148	145	16352	79

Fig. 11 Snapshot of the activity tracker dataset

Experimental configurations

The proposed multilevel noise function mechanism, data classification mechanism, and noise removal and fuzzification mechanism are implemented in NetBeans IDE 8.2 [45] for Java. SQLite version 3.21.0 [46] as a backend and SQLiteStudio 3.1.1 [47] is used to manage SQLite database. Proposed mechanisms are executed on the two different types of datasets. The first dataset is Activity Recognition from a Single Chest-Mounted Accelerometer [48] dataset. This dataset is collected from a wearable accelerometer mounted on the chest. Accelerometer data are collected from 15 participants performing 7 activities. The sampling frequency of the accelerometer was 52 Hz. Each record in a file contains a sequential number, x acceleration (attribute F_1), y acceleration (attribute F_2), z acceleration (attribute F_3), and label for activity attributes. The second dataset is collected from the activity tracker, a hand-wearable device, and contains three-axis Accelerometer, Detached PPG Cardio Tachometer, Infrared Wear Sensor. This activity tracker can continuously track Heart Rate, Steps, Distance, and Calories Burned parameters. It is assumed that data collected from a wearable accelerometer mounted on the chest and activity tracker device are sensitive for the user. Different results for various cases are recorded for findings and performance anal-

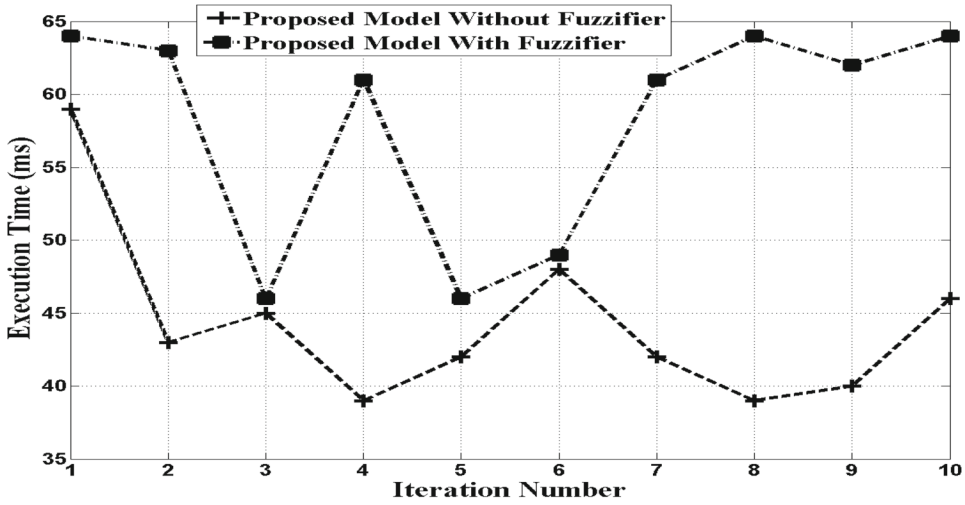

Fig. 12 Comparative execution time of the proposed model with fuzzifier and without fuzzifer

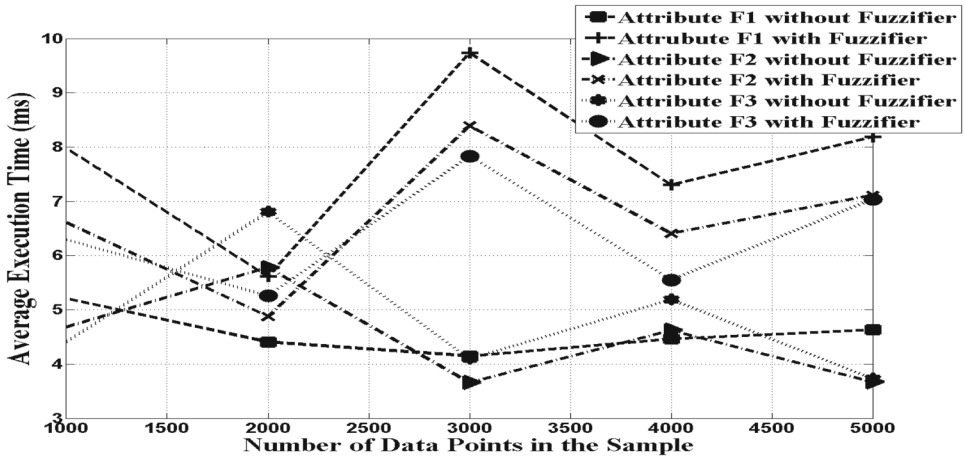

Fig. 13 Comparative analyses of the average execution time of the proposed model without fuzzifier and with fuzzifier

ysis. Initial simulation input parameters for the model are IoT parameter (D), Data Identifier (D_{ID}), Timestamp (T), Attribute Type (F_i), Total operators in a row in the operator table (N).

Results and discussion

The execution time is the time to access all the contents of a sample dataset. As shown in Eq. 8, the average execution time is the average of the total time to access the N_c number of contents of a specific attribute F_i. $(t_j)_{F_i}$ is the execution time to access jth content of F_i attribute type.

$$(\text{Average execution time})_{F_i} = \sum_{j=1}^{N_c} \frac{(t_j)_{F_i}}{N_c}. \tag{8}$$

A sample from the activity tracker dataset has been taken and calculated the execution time. Figure 12 shows the comparative execution time of the noise removal without fuzzification and with the fuzzification mechanism in the proposed model. It can be observed from the figure that noise removal with the fuzzification mechanism requires more execution time than noise removal without fuzzification. The sample sizes of 1000–5000 records (data points) have been taken from the Single Chest-Mounted Accelerometer dataset and calculated the average data execution time. The snapshots of the different data are shown in Figs. 10 and 11. Figure 13 presents the comparative average execution time of the noise removal without fuzzification and with the fuzzification mechanism in the proposed model for each data attribute F_1, F_2, and F_3. A sample of the data before and after the noise treatment is shown in Table 7, and a sample of the data without fuzzification and with the fuzzification after the noise removal is shown in Table 8. As shown in Table 8, all the data of a specific attribute type are treated with a fixed amount of noise. It gives a fixed amount of difference with all data of a particular attribute type, but it is not necessary to treat data with the fixed amount of noise. Every data of the particular attribute type may be treated with different random noises, and the resultant varying difference may enhance privacy.

Table 7 A sample of the data before and after the noise treatment

Data before noise addition			Data after noise addition								
			Data addends of F_1			Data addends of F_2			Data addends of F_3		
F_1	F_2	F_3	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3
1667	2072	2047	346	959	535	1166	320	759	1353	1273	48,747
1611	1957	1906	180	1289	39,004	635	185	1360	1320	304	505
1601	1939	1831	1073	1120	−369	945	712	455	680	85	1239
1643	1965	1879	476	92	40,705	960	297	48,066	706	320	1076
1604	1959	1921	750	1309	−282	385	962	835	1416	830	45,977
1640	1829	1940	1465	701	−303	1307	649	43,967	468	1352	293
1607	1910	1910	917	882	−19	1171	674	238	1252	719	45977
1546	2045	1910	977	114	37,757	138	217	1913	107	1098	928
1529	2049	1972	1081	864	−193	1294	1315	48814	278	1380	487
1637	1978	1945	880	931	−1	329	1385	437	1046	542	47,235
1596	2046	1866	1084	1429	37,585	544	154	1571	1027	858	204
1590	2006	1978	626	1171	16	897	1345	−63	929	496	726
1601	1966	1957	814	1123	38,286	1116	1199	47,033	734	192	1254
1542	2003	1959	751	1249	−285	548	1147	531	696	1147	47,330
1598	2027	1941	545	934	342	484	1126	49263	916	833	365
1511	2258	1983	735	1427	−478	1441	450	540	453	1018	48,302
1555	1980	2023	412	868	37793	1365	351	487	685	1278	283
1508	2468	1934	1303	234	194	1209	452	60,237	1117	469	521
1580	1697	2005	1373	998	−618	195	87	1588	1084	970	48,269
1627	2073	1992	1009	176	39,688	960	567	769	152	261	1802
1592	2130	2063	430	1416	−31	903	219	1181	821	647	768
1634	2088	1991	695	828	39,525	1071	195	51,132	136	478	1600
1638	2102	1916	186	445	1180	1240	472	613	1440	707	45,951
1593	2123	1948	1005	741	70	120	753	52,400	345	1294	482
1542	2133	2034	651	214	850	1338	806	162	953	437	49,658
1601	2015	2042	72	767	39,384	387	1229	622	1293	1057	−85
1613	1938	1936	1247	986	−397	302	1438	46,908	958	1402	−251
1644	1974	2000	1404	1366	−953	632	1006	509	269	1460	48,469
1642	1933	2046	1412	1320	38516	1386	416	354	1294	369	606
1605	1925	2011	474	991	363	260	953	885	289	1059	836
1586	1998	2066	468	1177	38,203	852	72	49,224	236	1244	809
1577	2032	2108	1107	938	−295	885	82	1288	844	776	51,278
1598	1980	2066	1217	915	−311	238	105	49,355	1060	1115	64
1561	1942	2092	851	1071	−188	646	793	676	379	1394	50725
1628	1935	2142	1425	149	39,324	520	503	1135	558	735	1072
1694	1965	2052	1272	103	542	969	1354	47,000	271	997	957
1627	1922	2081	1013	1156	−369	1384	1326	−615	102	701	51420
1598	1950	2117	343	1266	38,539	361	137	1675	455	1103	782
1612	1952	2075	1436	95	304	217	1166	742	482	1232	534
1630	1958	2024	672	932	39,344	764	163	48,221	953	769	525

Table 7 continued

Data before noise addition			Data after noise addition								
			Data addends of F_1			Data addends of F_2			Data addends of F_3		
F_1	F_2	F_3	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3
1609	1973	2005	989	838	−45	330	867	999	1308	507	48,508
1600	1983	2014	515	508	800	731	185	48,857	85	1293	809
1608	1969	2024	230	1004	547	422	919	801	1398	669	48,731
1612	1957	2019	262	1200	39,036	794	342	1044	334	924	984
1605	1943	2018	1273	470	85	1160	1146	46,467	552	765	874
1640	1917	2037	914	1156	−257	1160	217	713	401	305	50,417
1610	1949	2077	424	328	39,696	1258	1170	−256	1082	560	658

Table 8 A sample of the data without fuzzification and with the fuzzification after the noise removal

Data after noise removal and without fuzzification			Data after noise removal and with fuzzification		
DA_1	DA_2	DA_3	DA_1	DA_2	DA_3
1667	2072	2047	1662.5	2074.5	2044
1611	1957	1906	1606.5	1959.5	1903
1601	1939	1831	1596.5	1941.5	1828
1643	1965	1879	1638.5	1967.5	1876
1604	1959	1921	1599.5	1961.5	1918
1640	1829	1940	1635.5	1831.5	1937
1607	1910	1910	1602.5	1912.5	1907
1546	2045	1910	1541.5	2047.5	1907
1529	2049	1972	1524.5	2051.5	1969
1637	1978	1945	1632.5	1980.5	1942
1596	2046	1866	1591.5	2048.5	1863
1590	2006	1978	1585.5	2008.5	1975
1601	1966	1957	1596.5	1968.5	1954
1542	2003	1959	1537.5	2005.5	1956
1598	2027	1941	1593.5	2029.5	1938
1511	2258	1983	1506.5	2260.5	1980
1555	1980	2023	1550.5	1982.5	2020
1508	2468	1934	1503.5	2470.5	1931
1580	1697	2005	1575.5	1699.5	2002
1627	2073	1992	1622.5	2075.5	1989
1592	2130	2063	1587.5	2132.5	2060
1634	2088	1991	1629.5	2090.5	1988
1638	2102	1916	1633.5	2104.5	1913
1593	2123	1948	1588.5	2125.5	1945
1542	2133	2034	1537.5	2135.5	2031
1601	2015	2042	1596.5	2017.5	2039
1613	1938	1936	1608.5	1940.5	1933
1644	1974	2000	1639.5	1976.5	1997

Table 8 continued

Data after noise removal and without fuzzification			Data after noise removal and with fuzzification		
DA_1	DA_2	DA_3	DA_1	DA_2	DA_3
1642	1933	2046	1637.5	1927.5	2043
1605	1925	2011	1600.5	2000.5	2008
1586	1998	2066	1581.5	2034.5	2063
1577	2032	2108	1572.5	1982.5	2105
1598	1980	2066	1593.5	1944.5	2063
1561	1942	2092	1556.5	1937.5	2089
1628	1935	2142	1623.5	1967.5	2139
1694	1965	2052	1689.5	1924.5	2049
1627	1922	2081	1622.5	1952.5	2078
1598	1950	2117	1593.5	1954.5	2114
1612	1952	2075	1607.5	1960.5	2072
1630	1958	2024	1625.5	1975.5	2021
1609	1973	2005	1604.5	1985.5	2002
1600	1983	2014	1595.5	1971.5	2011
1608	1969	2024	1603.5	1959.5	2021
1612	1957	2019	1607.5	1945.5	2016
1605	1943	2018	1600.5	1919.5	2015
1640	1917	2037	1635.5	1951.5	2034
1610	1949	2077	1605.5	1932.5	2074
1633	1930	2076	1628.5	1870.5	2073
1573	1868	2058	1568.5	1946.5	2055
1568	1944	1959	1563.5	2047.5	1956
1576	2217	2059	1571.5	1983.5	2056

Figure 14 presents the findings of the comparative average execution time of the noise removal without fuzzification in the proposed model, and data access control scheme [25] for each data attribute F_1, F_2, and F_3 of the Single Chest-Mounted Accelerometer dataset. Figure 15 presents the findings of the comparative average execution time of the noise removal with fuzzification in the proposed model and data access control scheme [25] for each data attribute F_1, F_2, and F_3 of the Single Chest-Mounted Accelerometer dataset. It is clear from both of these figures that our proposed noise removal mechanism requires less execution time than the data access control scheme [25].

The findings presented in Fig. 16 are the comparative execution time of the noise removal without fuzzification in the proposed model, data access control scheme [25] and data access time of the DPP model [26]. Next, Fig. 17 presents the comparative execution time of the noise removal with fuzzification in the proposed model; data access control scheme [25] and data access time of the DPP model [26].

Algorithmic behavior and performance evaluation

IoT components, such as sensors, actuators, etc., have limited computing capabilities and are not suitable for performing complex computing operations. The comparative analysis shown in Table 9 indicates that the proposed model without fuzzifier has around 52–77% and 46–70% less computational overhead than the data access controls scheme and DPP model, respectively. The proposed model with the fuzzifier has around 48–73% and 31–63% less computational overhead than the data access controls scheme and DPP model, respectively. As the critical research problem to develop a systematic model to ensure end-to-end privacy against various threats for resource-constrained IoT environments and the main objective of the proposed NBPPM model, this analysis of the computational overhead for resource-constrained IoT environments shows the efficiency of the NBPPM model.

$$\Gamma(\phi; \varepsilon, \Omega, \pi, \omega) = \frac{1}{1 + e^{-(-\varepsilon(\phi - \pi\Omega))}} * \frac{1}{\omega}. \tag{9}$$

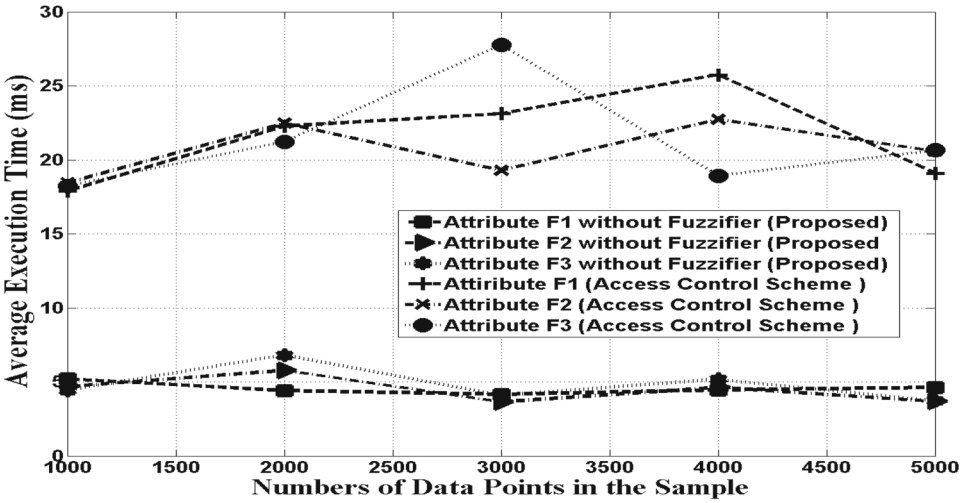

Fig. 14 Comparative analyses of the average execution time of the proposed model without fuzzifier and data access control scheme [25]

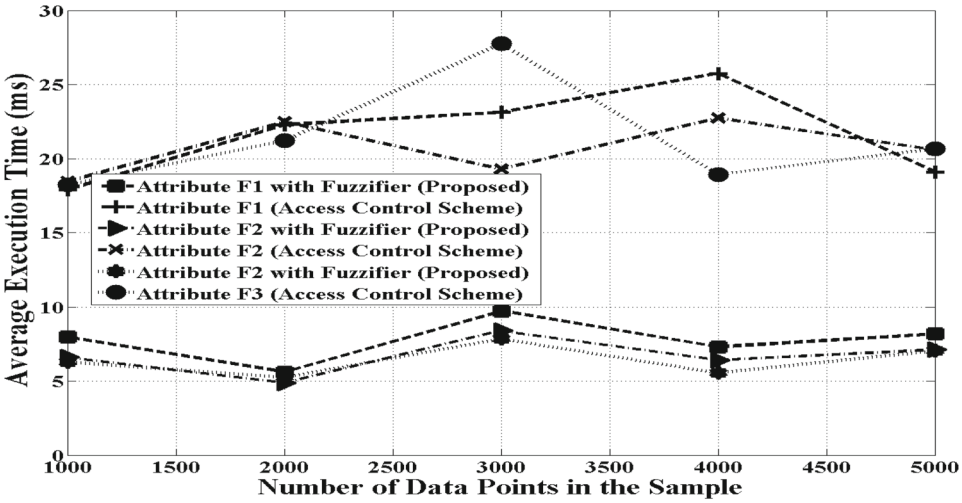

Fig. 15 Comparative analyses of the average execution time of the proposed model with fuzzifier and data access control scheme [25]

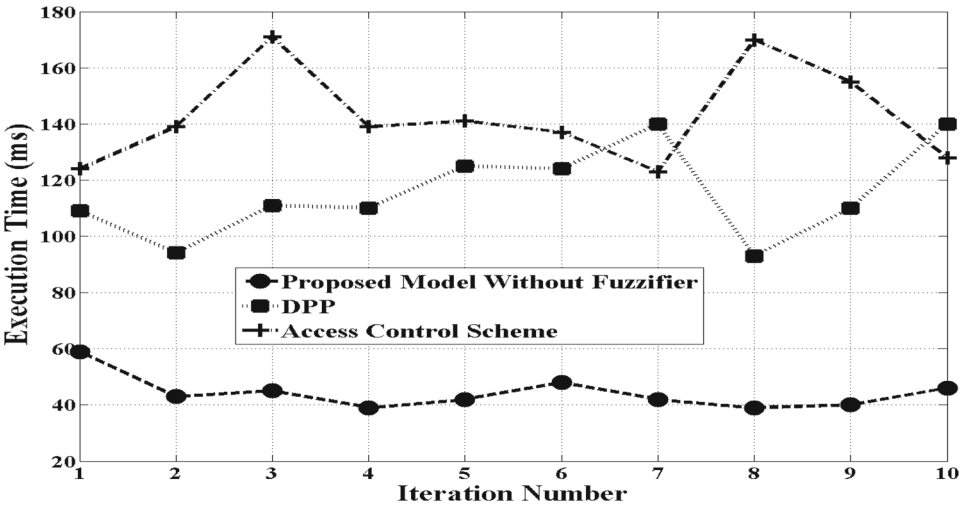

Fig. 16 Comparative analyses of the execution time of the proposed model without fuzzifier, data access control scheme [25] and DPP model [26]

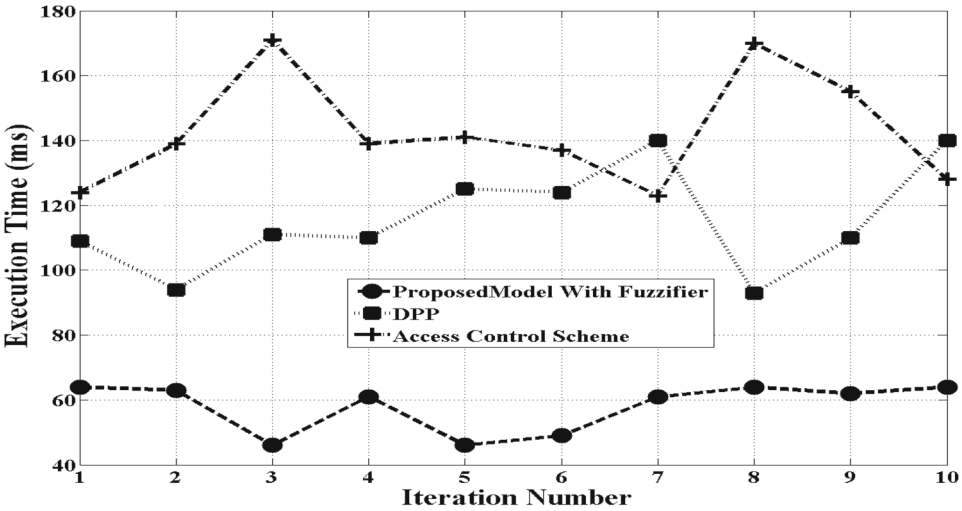

Fig. 17 Comparative analyses of the execution time of the proposed model with fuzzifier, data access control scheme [25] and DPP model [26]

Table 9 Comparative analysis of the computational overhead

Iteration number	% Decrement in the computational overhead (in the proposed model without fuzzifier)		% Decrement in the computational overhead (in the proposed model with fuzzifier)	
	In comparison of the data access control scheme (%)	In comparison of the DPP model (%)	In comparison of the data access control scheme (%)	In comparison of the DPP model (%)
1	52.419	45.871	48.387	41.284
2	69.064	54.255	54.676	32.978
3	73.684	59.459	73.099	58.558
4	71.942	64.545	56.115	44.545
5	70.212	66.400	67.375	63.200
6	64.963	61.290	64.233	60.483
7	65.853	70.000	50.406	56.428
8	77.058	58.064	62.352	31.182
9	74.193	63.636	60.000	43.636
10	64.062	67.142	50.000	54.285

As an IoT environment deals with a massive amount of data, it is crucial to consider computational time for performance measurement. The integrated multi-dimensional relationship of sensitivity levels of personally identifiable information items, privacy, and trust (Eq. (4)) allowed the author to devise Eq. (9). It is believed that the efficiency of a privacy-preserving algorithm increases as there is a decrement in the computational time (ω) of that algorithm, and trust value will increase with the increment in the effectiveness of the privacy-preserving algorithm. In terms of computational time, Noise Removal and Fuzzification Mechanism's efficacy can be seen in the comparative analysis with the given existing mechanism. In particular, the findings presented show that computational time is less in our noise removal mechanism compare to the encryption-based mechanisms. A privacy customization feature has been incorporated for the user, and comparative analysis with this feature also shows better performance. The experimental results presented in "Results and discussion" validate the feasibility and applicability of the novel NBPPM model for privacy preservation in the real-world and resource constraint environment of the internet of things.

Security and privacy analysis

The proposed NBPPM model ensures security and privacy through Multilevel Noise Treatment and Fuzzification. The privacy of the data is ensured by adding noise. The noise is sub-divided into three sub-keys as described in "Multilevel noise treatment". Sub-noise P_1, P_2 and P_3 is privately shared with the Data-Source, middleware and data storage server, respectively. The proposed Multilevel Noise Treatment Mechanism stores sensed IoT parameter D as noisy data addends. At the data-source, every sensed parameter is converted into noisy data and then split into meaningless noisy data addends, so it is difficult to know original data without

Table 10 An instance of data before and after privacy preservation in NBPPM model

	Attribute F_1			Attribute F_2			Attribute F_3		
Sensed data	1667			2072			2047		
Data addend	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3	DA_1	DA_2	DA_3
Source device	280	893	469	1100	254	693	1287	1207	48,681
Middleware	295	908	484	1115	269	708	1302	1222	48,696
Long-term storage	346	959	535	1166	320	759	1353	1273	48,747
User device (after denoising and before fuzzification)	1667			2072			2047		
User device after fuzzification	1662.5			2074.5			2044.0		

the sub-key P_1 and the operator used to treat the source data with the noise P_1. Further, in the proposed model, a user-customized data classifier is employed to protect sensitive data with a higher level of privacy preservation. At middleware, complexity increases for an eavesdropper to know the original sensed parameter due to the requirement of sub-noise P_1, P_2 and the operators used to treat the source data with the noise P_1 and P_2. At long-term data storage (such as cloud), it is extremely complex due to the requirement of all three sub-noises and operators used to treat the source data with the noise P_1, P_2 and P_3. A comprehensive status of an instance of data at different levels within the NBPPM model, i.e., data before and after privacy preservation, is shown in Table 10.

Furthermore, an attacker could use vulnerabilities such as a weak credential mechanism to gain access to the data. If a user requests data through sending a data request for IoT parameter (D) (containing data field identifier (F_i), timestamp, and unique username); the authentication mechanism is used in our proposed model to authenticate the user, and thereby the non-legitimate user cannot access the sensitive data. An access control list (ACL) maintains for usernames and their credentials. Even if, at this level, an eavesdropper succeeds in accessing noisy data addend, then privacy will still be preserved since noisy data addends are meaningless. After the successful authentication, only a legitimate user will be able to access noisy data addends. Our proposed Noise Removal and Fuzzification Mechanism also provides flexible and dynamic ways to preserve privacy through the privacy manager module. A user can customize his/her sensitive attributes and level of the sensitivity of their data. Based on the privacy customization, a user-specific privacy preservation environment will be created by the fuzzifier module. A comparative analysis of different frameworks for privacy preservation in IoT is presented in Table 11.

Applicability in real life applications

The proposed NBPPM model can be used in all real-life IoT applications, especially in the application domains where data sensitivity is high. This subsection illustrates a real-life example of the NBPPM model in the IoT-based healthcare system. A typical IoT-based healthcare system involves patient (s), doctors (s), hospital (s), and IoT-based service (s). In this IoT ecosystem, a patient is the user of the IoT-based healthcare system. A patient can be equipped with sensors (that sense the patient's health parameters), and with a mobile app, a patient can be enabled to use IoT-based healthcare services. Doctor and hospital act as a service provider. A hospital may use third-party services like cloud services to store a massive amount of the produced IoT data. In this scenario, patient's data are sensitive because of the sensing of health-related parameters. The sensitivity of these health-related parameters may vary from patient to patient, i.e., some patients may want to keep their data private because, for them, sensed health parameters are highly sensitive and for some patients, sensed health parameters are less sensitive. In this situation, the proposed NBPPM model may play a significant role in preserving privacy. An NBPPM model-based IoT healthcare system; preserves user privacy at different levels of the IoT ecosystem, as described in "Security and privacy analysis".

Limitations and future scope

Several different modifications, experiments, and analyses have been left for the future due to the study's broad research scope. Future work may focus on in-depth analysis of the particular mechanisms with new proposals to try different enhanced strategies. The following subsection emphasizes the potential future scope for improvement and research directions.

Table 11 Comparative analysis of different frameworks for privacy preservation in IoT

Framework/approach	Available features		Major aspects/threat(s) covered	References
	Data classifier	Customized user privacy option		
Lightweight encryption algorithm	×	×	Inside attack	[27]
Privacy preserving IoT architecture	×	×	Data breach in collection, storage and retrieval	[25]
DPP model	✓	Configured privacy weight	User privacy violation	[26]
EPIC	×	×	Traffic analysis attack, side channel attack	[28]
Privacy-preserving model	✓	✓	Non trusted objects	[29]
Privacy-preserving trust model	×	×	Sensing node identity, sensed value, user preferences	[20]
NBPPM model	✓	✓	Privacy breach in data collection, storage and retrieval	Our proposed model

Applicability and scope of the proposed solution with emerging domains

Evolutionary computations, i.e., Genetic Algorithms (GA) based obfuscation mechanism, could be applied in the proposed models. Crossover and Mutation phases can play a significant role in suppressing sensitive information in the IoT ecosystem, and managing the mutation phase to regenerate information can be a challenging step. Still, it will be interesting to develop and analyses the behavior of these kinds of optimization techniques.

Machine Learning has the potential for real-time automation, intelligent processing, and analysis of the high volume of data. The data classification mechanism of the proposed Noise-Based Privacy-Preserving Model can exploit this predictive-power of Machine Learning. This predictive-power may assist in identifying sensitive information in the IoT ecosystem and will reduce human intervention. In the future, Machine Learning-based mechanisms may be incorporated in the proposed model and analyze behavior.

Accountability is an important feature that can enhance every privacy preservation mechanism by rendering control over sensitive personal information. A procedure that keeps the history of all logs (such as a chain of all paths where sensitive data are traveled and the details of the data accessing entity) can be incorporated with the proposed models but will increase computational and space overhead. A future study can be conducted to incorporate this aspect.

There is a tradeoff between Quality of Service (QoS) provided to the user and users' data consumed by the service provider. In the Noise-Based Privacy-Preserving Model, a sub-module (Privacy Manager) plays a crucial role in customizing user privacy, and privacy is ensured through the fuzzification process when it is transferred between the user and the service provider. Here is the scope to further optimize the membership function for the specific application and case study.

Along with the study's future scope, the following are emerging domains where proposed solutions can be employed:

Edge computing and fog computing

As edge computing and fog computing, both paradigms move the computational capabilities closer to the data source, and these computing technologies may move data intelligence and data analytics near the IoT ecosystem's data sources. In future work, such approaches may be adopted in our proposed privacy-preserving model to distribute trust with the enhanced privacy protection in the IoT ecosystem.

It will be interesting to develop and study the architectural integration of edge and fog computing in the privacy-

preserving IoT ecosystem, privacy-preserving edge and fog data processing, and management of edge and fog nodes in the frameworks.

Blockchain

An adversary can infer significant information about the users from blockchain-based IoT networks. These systems need specific privacy-protection plans to preserve personal and device privacy. A critical perspective that causes privacy leakage in the blockchain network is address reuse. Public addresses of blockchain users are open to anyone in the network, and an adversary can easily access these addresses through internet access. A perfect anonymous transaction in the blockchain is unlikely without any particular privacy-preservation plan. Also, linking attacks can be performed over distributed ledger that contains a copy of transactions [49]. The proposed model can be applied to preserve privacy in this scenario, and it is further a future scope of the study for these use cases.

Fifth Generation technology

The lately emerging Fifth Generation (5G) technology is expected to transform every area of life by connecting everything, everywhere, by employing IoT devices. However, massively interconnected devices and high-speed data communication will bring the challenge of privacy and energy insufficiency. 5G industries and organizations require privacy-preservation for their endurance and competency. Moreover, billions of devices supposed to communicate using the 5G network will spend a considerable amount of energy while confined energy-resources. Hence, energy-optimization is a future challenge confronted by 5G industries that need to be addressed [50]. In this case, our proposed privacy-preserving model can be integrated with 5G technology, and it will be interesting to study improved privacy with the energy resource optimization in this specific use case.

Autonomous vehicles

The emergence of complex cyber-physical systems (CPS) such as an autonomous vehicle is equipped with different sensors and intelligent logic to provide advanced auxiliary services. Due to their sensor and inboard intelligence, such vehicles gather, analyze, and capitalize upon an unprecedented quantity of fine-grained data and cooperate in real-time with various stakeholders. However, such valuable data can significantly impact data-driven economies of scale, which raises questions concerning privacy and integrity-dependent situations [51]. Our proposed study's future scope is the measurement of real-time performance with autonomous vehicles and should cover a study of the level of the balance between privacy preservation and quality of service in this specific use case.

Conclusion

The NBPPM model has been presented to address critical issues of privacy preservation in the IoT ecosystem. The proposed model ensures end-to-end privacy preservation in the IoT environment. The NBPPM is a robust and flexible model that ensures privacy preservation according to the user's preferences. The performance of the proposed NBPPM model has been evaluated in terms of computation overheads. Our experimental results show that the computational cost in NBPPM is reasonably less in the practical scenarios. In this article, the feasibility of the proposed model has been demonstrated for the IoT's resource-constrained environment. An exciting future work of the NBPPM model may be incorporating accountability procedures at the appropriate levels to enhance control over personal information in the IoT environment. The outcomes of this work may have a significant effect on IoT-based industries.

Acknowledgements The authors would like to thank the Ministry of Human Resource Development, Government of India, for the Global Initiative of Academic Networks (GIAN). GIAN network provides us general ideas for the IoT ecosystem. Furthermore, the authors are thankful to the UCI Machine Learning Repository [48] that helps to fulfill the necessary data required in the experiment. Finally, the authors wish to thank reviewers for their precious time and comments. This research was supported by UGC: Maulana Azad National Fellowship formulated and funded by the Ministry of Minority Affairs [Grant F11/2015MANF/(SAWebsite)].

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Solanki A, Nayyar A (2019) Green internet of things (g-IoT): Ict technologies, principles, applications, projects, and challenges. In: Handbook of research on big data and the IoT, IGI Global. pp 379–405
2. Virkki J, Chen L (2013) Personal perspectives: individual privacy in the IoT. Adv Internet Things 3(02):21
3. Krishnamurthi R, Kumar A, Gopinathan D, Nayyar A, Qureshi B (2020) An overview of IoT sensor data processing, fusion, and analysis techniques. Sensors 20(21):6076
4. Jain SK, Kesswani N (2019) Smart judiciary system: a smart dust based IoT application. In: International conference on emerging technologies in computer engineering. Springer, pp 128–140
5. Perera C, Ranjan R, Wang L, Khan SU, Zomaya AY (2015) Big data privacy in the Internet of Things era. IT Prof 17(3):32–39
6. Jain SK, Kesswani N (2019) Privacy threat model for IoT. In: International conference on Internet of Things and connected technologies. Springer, pp 278–293
7. Roman R, Zhou J, Lopez J (2013) On the features and challenges of security and privacy in distributed internet of things. Comput Netw 57(10):2266–2279
8. Porambage P, Ylianttila M, Schmitt C, Kumar P, Gurtov A, Vasilakos AV (2016) The quest for privacy in the Internet of Things. IEEE Cloud Comput 3(2):36–45
9. Smith HJ, Dinev T, Xu H (2011) Information privacy research: an interdisciplinary review. MIS Quart 35(4):989–1016
10. Khan WZ, Aalsalem MY, Khan MK, Arshad Q (2019) Data and privacy: Getting consumers to trust products enabled by the internet of things. IEEE Consum Electron Magaz 8(2):35–38
11. Ahmed AIA, Ab Hamid SH, Gani A, Khan MK et al (2019) Trust and reputation for internet of things: fundamentals, taxonomy, and open research challenges. J Netw Comput Appl 145:102409
12. Sen AAA, Eassa FA, Jambi K, Yamin M (2018) Preserving privacy in internet of things: a survey. Int J Inf Technol 10(2):189–200
13. Fabiano N (2017) Internet of things and blockchain: legal issues and privacy. the challenge for a privacy standard. In: 2017 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), IEEE, pp 727–734
14. Ziegeldorf JH, Morchon OG, Wehrle K (2014) Privacy in the Internet of Things: threats and challenges. Secur Commun Netw 7(12):2728–2742
15. Abomhara M, Køien GM (2014) Security and privacy in the Internet of Things: current status and open issues. In: (2014) international conference on privacy and security in mobile systems (PRISMS). IEEE 1–8
16. Sennan S, Ramasubbareddy S, Luhach AK, Nayyar A, Qureshi B (2020) CT-RPL: cluster tree based routing protocol to maximize the lifetime of Internet of Things. Sensors 20(20):5858
17. Caron X, Bosua R, Maynard SB, Ahmad A (2016) The Internet of Things (IoT) and its impact on individual privacy: an Australian perspective. Comput Law Secur Rev 32(1):4–15
18. Corcoran PM (2016) A privacy framework for the internet of things. In: (2016) IEEE 3rd world forum on Internet of Things (WF-IoT). IEEE 13–18
19. Moosavi SR, Gia TN, Rahmani A-M, Nigussie E, Virtanen S, Isoaho J, Tenhunen H (2015) Sea: a secure and efficient authentication and authorization architecture for IoT-based healthcare using smart gateways. Procedia Comput Sci 52:452–459
20. Appavoo P, Chan MC, Bhojan A, Chang E-C (2016) Efficient and privacy-preserving access to sensor data for internet of things (iot) based services. In: 2016 8th International conference on communication systems and networks (COMSNETS). IEEE 1–8
21. Sankar L, Rajagopalan SR, Poor HV (2013) Utility-privacy tradeoffs in databases: an information-theoretic approach. IEEE Trans Inf Foren Secur 8(6):838–852
22. Turgut D, Boloni L (2017) Value of information and cost of privacy in the Internet of Things. IEEE Commun Mag 55(9):62–66
23. Daubert J, Wiesmaier A, Kikiras P (2015) A view on privacy and trust in IoT. In: 2015 IEEE international conference on communication workshop (ICCW), IEEE. pp 2665–2670
24. Butun I (2017) Privacy and trust relations in Internet of Things from the user point of view. In: 2017 IEEE 7th annual computing and communication workshop and conference (CCWC). IEEE 1–5
25. Jayaraman PP, Yang X, Yavari A, Georgakopoulos D, Yi X (2017) Privacy preserving Internet of Things: from privacy techniques to a blueprint architecture and efficient implementation. Future Gener Comput Syst 76:540–549
26. Gai K, Choo K-KR, Qiu M, Zhu L (2018) Privacy-preserving content-oriented wireless communication in Internet-of-Things. IEEE Internet Things J 5(4):3059–3067
27. Yi X, Willemson J, Nait-Abdesselam F (2013) Privacy-preserving wireless medical sensor network. In: 2013 12th IEEE international conference on trust, security and privacy in computing and communications, IEEE. pp 118–125
28. Liu J, Zhang C, Fang Y (2018) Epic: a differential privacy framework to defend smart homes against internet traffic analysis. IEEE Internet Things J 5(2):1206–1217
29. Chen Z, Tian L (2017) Privacy-preserving model of IoT based trust evaluation. IEICE Trans Inf Syst 100(2):371–374
30. Zhou W, Piramuthu S (2015) Information relevance model of customized privacy for IoT. J Bus Ethics 131(1):19–30
31. Samani A, Ghenniwa HH, Wahaishi A (2015) Privacy in Internet of Things: a model and protection framework. Procedia Comput Sci 52:606–613
32. Shabalala M, Tarwireyi P, Adigun M (2014) Privacy monitoring framework for enhancing transparency in cloud computing. In: 2014 IEEE 6th international conference on adaptive science and technology (ICAST). IEEE, pp 1–7
33. Song T, Li R, Mei B, Yu J, Xing X, Cheng X (2017) A privacy preserving communication protocol for IoT applications in smart homes. IEEE Internet Things J 4(6):1844–1852
34. Zhang C, Li C, Zhao Y (2015) A balance privacy-preserving data aggregation model in wireless sensor networks. Int J Distrib Sens Netw 11(6):937280
35. Boussada R, Elhdhili ME, Saidane LA (2017) Privacy preserving solution for Internet of Things with application to ehealth. In: 2017 IEEE/ACS 14th international conference on computer systems and applications (AICCSA). IEEE, pp 384–391
36. Jain SK, Kesswani N (2020) Iotp an efficient privacy preserving scheme for Internet of Things environment. Int J Inf Secur Privacy (IJISP) 14(2):116–142
37. Jain SK, Kesswani N, Agarwal B (2020) Security, privacy and trust: privacy preserving model for Internet of Things. Int J Intell Inf Database Syst 13(2–4):249–277
38. Berrehili FZ, Belmekki A (2016) Privacy preservation in the Internet of Things. In: International symposium on ubiquitous networking. Springer, pp 163–175
39. Liang H, Wu D, Xu J, Ma H (2015) Survey on privacy protection of android devices. In: 2015 IEEE 2nd International conference on cyber security and cloud computing. IEEE, pp 241–246
40. Al-Gburi A, Al-Hasnawi A, Lilien L (2018) Differentiating security from privacy in Internet of Things: a survey of selected threats and controls. In: Computer and network security essentials. Springer, pp 153–172
41. What personal data is considered sensitive? (2018) European commission–European commission. https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-

data/what-personal-data-considered-sensitive_en. Accessed: 05-09-2020
42. Article 4(13), (14) and (15), Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e1489-1-1. Accessed: 05-09-2020
43. Article 9, Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e2051-1-1. Accessed: 05-09-2020
44. Recitals (51) to (56) of the GDPR, Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e40-1-1. Accessed: 05-09-2020
45. Netbeans.org. NetBeans IDE 8.2. https://netbeans.org/downloads/8.2/. Accessed: 27-02-2019
46. Hipp R et. al SQLite (Version 3.8.10.2) [Computer software]. SQLite Development Team. https://www.sqlite.org/download.html. Accessed: 24-10-2017
47. SQLiteStudio 3.1.1. https://sqlitestudio.pl/index.rvt?act=download. Accessed: 30-11-2016
48. UCI machine learning repository. https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer. Accessed: 15 Feb 2019
49. Hassan MU, Rehmani MH, Chen J (2019) Privacy preservation in blockchain based IoT systems: integration issues, prospects, challenges, and future research directions. Future Gener Comput Syst 97:512–529
50. Humayun M, Jhanjhi N, Alruwaili M, Amalathas SS, Balasubramanian V, Selvaraj B (2020) Privacy protection and energy optimization for 5G-aided industrial Internet of Things. IEEE Access 8:183665–183677
51. Karnouskos S, Kerschbaum F (2017) Privacy and integrity considerations in hyperconnected autonomous vehicles. Proc IEEE 106(1):160–170