SECURE DATA TRANSMISSION IN IOT USING REINFORCEMENT LEARNING AND CIPHERTEXT-POLICY ATTRIBUTE-BASED ENCRYPTION

Ms. HajaBanu Shaikh Mohammed Essa[1], Ms. Zeba Khan[2]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

More and more devices are exchanging data via the Internet of Things (IoT), which increases the quality of IoT facilities. There are various vulnerabilities, security weaknesses, and attack vectors in IoT systems that represent a long-term security issue. To realise the full potential of IoT applications, billions of linked devices must be secured. Information on IoT is vulnerable to threats, assaults, and flaws. To solve IoT-related security, privacy, and vulnerability challenges, a strong security solution is required. Many scientists have previously suggested solutions for security in IoT. Deep Learning is one of the most promising techniques for protecting IoT systems in recent years, and Reinforcement Learning is gaining popularity in this regard. Unlike other deep learning approaches, reinforcement learning may learn the environment with a little amount of data about the parameters to be learnt. Deep reinforcement learning (DRL) methods may be capable of dealing with the aforementioned challenges related with IoT devices in the near future. In this paper, the Deep Reinforcement Learning (DRL) approach is introduced to increase IoT security. This study presents quantum computing for feature selection, and to securely transmit data, the Ciphertext-Policy Attribute-Based Encryption (CP-ABE) method has been built for an IoT system, and its performance is compared with traditional approaches. The suggested solution allowed for safe and scalable data transmission. The accuracy obtained by the proposed system is higher when compared to the other approaches.

[1]Lecturer, Department of Computer and Information, Applied College, Jazan University, Jazan, Kingdom of Saudi Arabia, 45142, PO box 114.
[2]Lecturer, Department of Computer and Information, Applied College, Jazan University, Jazan, Kingdom of Saudi Arabia, 45142, PO box 114.

Email: [1]hshaikh@jazanu.edu.sa, [2]zabumazeedullah@jazanu.edu.sa

DOI: 10.31838/ecb/2023.12.1.107

1. Introduction

A physical and digital world are linked through the Internet of Things (IoT). Using this technology, machines communicate with each other to solve simple to difficult tasks (Olowononi et al. 2020). In the physical and digital world, data is transferred through sensors and actuators. As a result of sensors, data will be collected and analyzed so that a service can be offered to the user. By adding intelligence to machines, it has changed human living significantly, and it will eventually improve the standard of living. The IoT has significantly increased internet usage by connecting all physical devices within a network. IoT smart devices generate health information, geographic information, and security data. IoT can give services of the highest caliber via sharing data. Direct data interchange in the IoT, on the other hand, poses the following issues. Participants in data sharing, for example, find it difficult to trust one another. As a result, determining ways to assure the dependability of shared data may be difficult. Furthermore, due to privacy concerns (Jan et al. 2019), participants are unwilling to provide their personal data in the absence of suitable privacy preservation procedures. In order to ensure data dependability and data privacy, data sharing should provide both data security involves keeping the integrity, legitimacy, and validity of the information while safely preserving and transmitting it. The confidentiality of data can be maintained if only authorised individuals are permitted access (Tang et al. 2018). Protection strategies can be determined by analyzing the expectations, goals, and requirements. Machine learning-based privacy protection has gained a lot of interest with the advancement of AI (Meneghello et al. 2019). Uprety et al. 2021 describe reinforcement learning as a machine learning method that optimizes numerical rewards through interactions with the environment. The human brain communicates with its circumstances, which it then uses to grasp and withstand existence in that context. RL employs the sensory perception system and brain as an example for environmental learning. The process involves an agent examining the entire system in order to comprehend it. It is not practical in many cases, as convergence and finding an optimal policy take a long time. Dimensionality is a curse on traditional RL. As the world becomes more complicated, the RL agent's constraints to learn increases exponentially. One alternative is deep reinforcement learning (DRL), a hybrid of deep learning model and reinforcement learning (RL). This research focuses on how RL has been used to protect IoT technology. The main contribution of the work is,

- This study presented an attack detection approach based on feature selection and reinforcement learning that can effectively pick the optimal set of attributes and maximize the effectiveness of IoT attack identification, especially the prediction of unknown network attacks.
- Introduced a quantum computing-based algorithm for selecting features that quantizes particular functions of conventional feature selection algorithms in order to boost performance by converting conventional feature selection techniques to quantum counterparts.
- A deep reinforcement learning model is developed for the detection and categorization of attack data in an IoT network, which accurately identifies unknown assaults.
- Finally, Ciphertext-Policy Attribute-Based Encryption algorithm is proposed for the encryption of data to securely transmit data within the network.

The organization of the work is as follows, In Section 2 the related works are presented, Section 3 describes the proposed methodology in detail, Section 4 presented the experimental results and Section 5 concludes the work.

Literature survey

To identify and avoid Sybil assaults in an IoT environment, Thuluva et al. (2021) suggested combining the traditional Caesar Cipher Algorithm (CCA) with the lightweight encryption algorithm (LEA) and the Received Signal Strength Indicator (RSSI). The proposed method identifies the dishonest nodes in a specific route by broadcasting the assault on a different node. Furthermore, it stops the onslaught by sending data packets to the designated users in a different manner. To provide identity, confidentiality, and information veracity, the lightweight encryption approach with a 64-bit key is used with AODV as the routing protocol.

Jebri et al. (2021) developed a lightweight secure IoT system using Pseudonym Based Cryptography, Elliptic Curve Cryptography, and Identity Based Encryption. The proposed method covers the great majority of security problems while also allowing for trust registration and anonymous authentication mechanisms. Furthermore, by utilising connection direction anonymity, the solution tackles the intractability problem between IoT nodes. Jian et al. (2021) proposed Hybrid Internet Of Things Data Transmission Security as an alternative to Device

Verification. Every Iot systems were identified and recognized independently.

Kandhoul et al. (2021) proposed a novel defence strategy for GFRSA, A Green Forwarding ratio, OppIoT, and RSA (Rivest, Shamir, and Adleman)-based safe routing mechanism. The next hops is determined by the node's forwarding activity, present energy levels, and information transmission probability estimation. To strengthen the protocol's reliability, messages are encrypted prior transmission through asymmetric cryptography.

Mondal et al. (2022) proposed a secure and energy-efficient healthcare monitoring system to reduce the data gathering system's energy usage and transmission expenses. On the other hand, it also allows safe data transfer by implementing a lightweight security mechanism based on Cipher Block Chaining (CBC). The suggested algorithms proven to be a viable option for safeguarding and prolonging the lifetime of the IoMT network.

Refaee et al. (2022) proposed an IoT architecture for secure and efficient medical data transmission by enhanced routing protocol. To minimize the data dimensionality, K-nearest Neighbor (KNN) imputation and principal component analysis (PCA) are utilised. Using local binary patterns that have been tweaked, the preprocessed data is utilised to extract the features (MLBP). The fuzzy dynamic trust-based RPL (FDT-RPL) strategy significantly enhances data transmission reliability by integrating the butter ant optimization (BAO) and fuzzy dynamic trust-based RPL algorithm for constrained and lossy networks.

Sankar and Karthiga (2020) created a framework for safe data transfer from an IoT wireless body sensor network (WBSN). Authentication, security, and load balancing are the three steps of implementation. For secure data transfer, optimized elliptical curve cryptography (OECC) is used. Load balancing is used to allow numerous users on the network by employing krill herd load balancing (KHLB).

Quantum walks (QWs) were recommended by El-Latif et al. (2021) as a cryptographic strategy to protect the authenticity of images on smart devices. They used QW to permute and integrate the original image into the source images after first replacing the original image. The proposed visually meaningful quantum walks cryptosystem is based on this structure and enables safe visual data transit. Simulation-based tests show how well the presented model performs in relation to visual quality, effectiveness, durability, and key space sensitivity, as well as its capacity to protect intelligent systems.

Sokol et al. 2021 described a technique for decreasing the quantity of data exchanged between a server and a Sensor node, with an emphasis on transmission security, bandwidth, and IoT device resources. Data compression and the replacement of the SSL/TLS cryptographic system with the lightweight cryptography which is based on Vernam cypher concept accomplish necessary reduction. Only device management applications continue to use the traditional SSL/TLS protocol.

An effective method for IIoT outlier identification in energy-efficient secured data transmission was proposed by SakthidasanSankaran and Kim in 2023. The primary goal of this technology is to provide a safe and private data transit mechanism for industrial IoT. The network accomplishes this through categorizing assaults using a Robust Multi-cascaded CNN (RMC-CNN) technique. The information is encrypted by the dynamic honeypot encryption technique using a key generation mechanism.

In order to transmit secure data in IoT based healthcare systems, Kumar et al. (2023) developed a Blockchain-orchestrated Deep Learning technique . A new scalable blockchain design is initially suggested to guarantee data integrity and secure information exchange. ZKP, or zero knowledge proof, is a method used in this design and the InterPlanetary File System (IPFS) for off-chain storage to control the expenses of the data storage, and an Ethereum platform program to handle data security challenges.

Premkumar and SathyaPriya (2022) developed an effective Service Constraint NCBQ (Network Condition and Behavior Quality) Trust Orient Secure Transmission Protocol (SCNCQB-TSTP). The Service Constraint Blockchain Model (SCBM) is used in this technique to safeguard data by classifying services. The service information was encrypted using both specific scheme and key based on the service selected by the user. Furthermore, information has been kept in a block chain for data transit.

Jan et al. (2021) created a successful data protection information embedding technique for a cyber-physical system. CLoG, a novel efficacious edge detector relying on the Canny and Laplacian of Gaussian detectors, was developed in this research and is used to find edge regions in digital pictures. The surreptitious data was encoded in the edges that were discovered. The suggested detector identifies finer edge features than existing detectors, allowing for more information to be disguised in a cover image. As a result, the number of cover images needed to transfer secret data is reduced, fulfilling the needs of resource-constrained systems such as IoT.

Hurrah et al. (2019) developed a security method on several levels based on data concealment and chaos theory. The approach proposed is based on Random Coefficient Selection and Mean Modification (RCSMMA). RCSMMA utilizes a number of discrete cosines. Transform coefficients

drawn at random from two separate blocks to guarantee that information is distributed evenly over the cover image.

2. Methodology

This section provides an in-depth description of the recommended strategy for maintaining information security during communication in IoTs. Data confidentiality is a crucial need that may be fulfilled by using a secure encryption mechanism because data in an IoT network goes via several hops. This is necessary because to the varied variety of IoT networks, applications, and devices that operate and interact via a flood of data and pose a considerable danger of privacy violation due to how easy it is to obtain data there. The challenges presented by a growing number of skilled cyber attackers are forcing attack detection model to depend more and more on automated and intelligent network intrusion detection techniques (Hou et al. 2020). The flow of the presented work is illustrated in Figure 1.

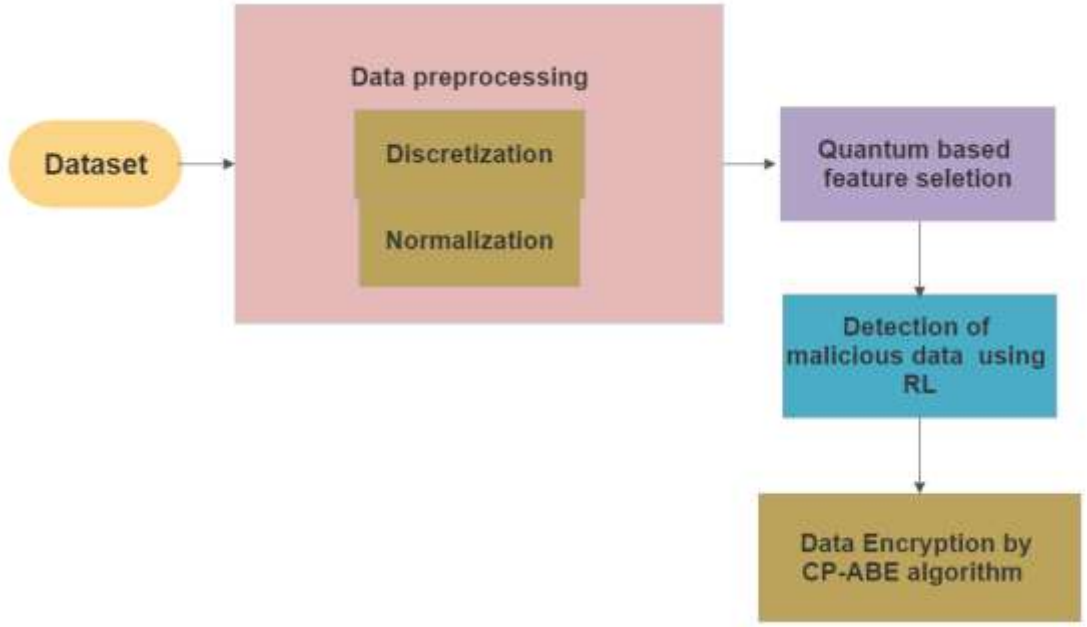

Figure 1 Flow of the proposed work

2.1 Dataset

For the purposes of this research, the NSL-KDD dataset is utilised, although there are several datasets available for network intrusion detection. An enhanced form of the KDD CUP 99 dataset developed by the DARPA intrusion detection assessment programme, consists of NSL-KDD, a dataset that simulates four categories of attacks: root-to-user (U2R), local-to-local (R2L), denial of service (DoS), and probing. NSL-KDD dataset is better suited to evaluating intrusion detection techniques because it eliminates duplicates and duplicate features from KDD Cup 99. As shown in Table 1, the dataset contains the following details.

Table 1 Details of the Dataset

Dataset	**Training set**		**Testing set**	
	Normal	**Attack**	**Normal**	**Attack**
NSL-KDD	67343	58630	12833	9711

2.2 Preprocessing of Data

An anomaly-based intrusion detection method, which uses statistical or machine-learning methods to determine if data is normal or anomalous, considers data preprocessing to be essential. Data representation, quantity, and quality has the considerable impact on the recognition technique's computational performance and accuracy. When working with a dataset that has a significant number of duplicate and/or redundant characteristics, knowledge discovery is likely to be more challenging. To eliminate or decrease

undesired traits, input data should be processed prior to the training process. Techniques used to prepare data include discretizing, normalizing, reducing dimensionality, and sampling.

2.2.1 Feature Discretization

By dividing a continuous collection into recesses and assigning categorical or minimal values to each interval, discretization turns a continuous collection of features into a finite discrete set. Using a discretization technique, the constant characteristics A in S would be divided into d discrete intervals I = {[i_0, i_1], [i_1, i_2],..., [i_{d-1}, id]}, where d0 is the lowest value of A and id is the largest value and $i_j < i_{j+1}$, for j = 0, 1,...,d – 1. The discrete set d is known as a discretization system on feature A, and the set of feature A's cut points is denoted by P = {i_1, i_2,...,i_{d-1}}.Discretization aims to enhance knowledge discovery by simplifying data representation and making some learning algorithms (such as those that only take nominal data) easier to employ.

2.2.2 Feature Normalization

Normalization or standardisation of numerical features refers to the process of handling large variances because the presence of numerical characteristics in training and validation sets may lead to biases in learning algorithms. The min-max normalisation linearly converts a d-dimensional feature vector z = (z_1, z_2, ...,z_d), into a normalised feature vector z^* = (z^*_1, z^*_2, ... , z^*_d), and z_i is translated to the range [0, 1] by the following expression:

$$z_i^* = \frac{z_i - z_i^{min}}{z_i^{max} - z_i^{min}}, \quad i = 1, 2, \dots, d, \tag{1}$$

where $z_i^{max} and z_i^{min}$ are the least and extreme values for the i[th] attribute in z.

The suggested technique makes use of deep reinforcement learning (DRL) networks. Nodes in a network are divided into two types: routers and hosts. The router is the component that connects users to hosts. Hosts are the sources and sinks of network traffic, since routers carry packets from host to host. Due to their significantly greater vulnerability to compromise than routers, hosts should not be trusted.

2.3 Quantum based feature selection

Feature selection can be sped up by quantization, which transforms a conventional approach into a quantum version. Here, feature selection is quantified. An objective function g is maximized by picking a subset of qualities that maximize it. Adding or eliminating a feature i to determine the aim function of the feature set. By requesting Oracle's simplification competence g, one requests the feature set F ∪ {i } or F − {i }. G_f and G_b are executed with diverse quantum classifiers. The prominent machine learning algorithm quantum least-squares support vector machine (LSSVM) is proposed by He et al. (2018). A quantum computer might be used to efficiently implement a support vector machine. Algorithm 1 covers the complete method of feature adding from the candidate feature set C-F to the selected attribute set F that optimizes generalization potential, where D is the complete attribute set and F ⊆ C.

With Grover's algorithm, the attribute i from the candidate attribute set must satisfy g(F ∪ {j }) >gmax, where g(F ∪ {j}) is the adaptability potential of the feature set- F ∪ {j }. As soon as such a feature j is found, gmax is updated to g(F ∪ {j }). A new algorithm is then run until either a high probability of finding feature i is achieved or no feature j is found. Quantum feature selection is performed by forward selection, as described in Algorithm 2. The feature set S is created by starting with an empty set of features. In each iteration, algorithm fea_add (F, C) is used to select feature i∗ from candidate feature set C − F and added to the nominated feature set S. The process concludes when the number of features in S equals p. η ∈ {0, 1}d is a vector size that specifies which characteristics are picked by the algorithm, where value 1 denotes the associated characteristic is considered and rather value is 0 (He et al. (2018).

Algorithm 1 *f ea_add(F, C)*

Input: *D*: the complete feature set; *F*: the chosen feature set where $F \subseteq C$.
Output: the feature index $j*$.
1: $gmax \leftarrow 0$, $j* = 0$
2: **repeat**
3: discover the feature *i*in $C - F$ which achieves $g(F \cup \{j\}) > gmax$using Grover's algorithm, where $g(F \cup \{j\})$ is the simplification ability of the classifier accomplished with the feature set $F \cup \{j\}$;
4: update $gmax = g(F \cup \{j\})$ and $j* = j$;
5: **until** no new jhas been found
6: **return j∗**

Algorithm 2 *f ea_sel_ f or(d, p)*: forward selection-based quantum feature selection

Input: *d*: the quantity of characteristics; *p*: the quantity of characteristics in the finally chosen subset; *C*: the full set
where $C = \{1, 2, \ldots, c\}$;
Output: $\boldsymbol{\eta} \in \{0, 1\}$c: the chosen attribute set.
1: $F \leftarrow \emptyset$, $\boldsymbol{\eta} \leftarrow \mathbf{0}$;
2: **repeat**
3: identify the attribute $j*$ by $j* = f\,ea\_add(F, C)$;
4: $F \leftarrow F \cup \{j*\}$;
5: **until** $|F| = p$
6: Set $\eta k = 1$ for $k \in F$;
7: **return** $\boldsymbol{\eta}$;

The features selected by the quantum algorithm is used for training the reinforcement learning model for the detection of malicious data.

2.4 Reinforcement Learning based attack detection

The suggested system is implemented using a distributed architecture to execute multiple IDS components. Attention learning networks in the core IDS learn to pay attention to agents. An IDS receives learned representations of packets (agent output) and prioritizes them. A central IDS determines whether a packet is malicious or not. Agents are compensated for feedback received from the environment by the central IDS. They learn and update their weights as a result of these rewards. Identifying the most rewarding path is the purpose of reinforcement learning. Generally, it works like this:

$$R_t = \sum_{k=0}^{\infty} \quad \gamma k^{Y} t + k \tag{2}$$

where R_t is the total cumulative, discounted possible return received by the agent at time t, $Y_{t+k}$the rewards of each future time step and discount factorγ ∈ (0; 1). Continually and discretely, the long-term calculation is carried out for future time states. As V^{π} (s) is a measure of state s, and expectation at state s, E (R_t| s_t = s). The value function is determined by the policy π that governs how the agent decides to behave. An ideal value-function with the greatest value among all feasible functions, indicated as $V^*(s) = \max_{\pi} V_{\pi}$ (s), where $\pi^* = \arg\max_{\pi} V^{\pi}(s)$ is the ideal policy that maximises the value of action attainable for state s. Reinforcement learning created a function called Q function that accepts state and action pair inputs and outputs the value of rewards. As a result, we can rewrite the expression as: $\pi^* = Q^*(S, a)$, where Q* is the greatest ideal value for Q. The Bellman equation then yields a recursive formulation for the optimum Q function as:

$$Q^*(s, a) = R(s, a) + \gamma E_{s'}[V^*(s') = R(s, a) + \gamma \sum_{s' \in s} \quad P(s'|s, a)V^*(s') \tag{3}$$

where R(s, a) is the instantaneous projected reward upon action a at state s and $E_{s'}[V^*(s')]$ is the discounted, anticipated, cumulative future reward following the transition to the next state s'. The following is a thorough discussion of several DRL components and core IDS.

Agent: A minimum of one DRL agent per network context is deployed according to the network topology. It is common practice to place IDS agents on routers that are k hops away from the central IDS. In R(k), routers that are directly connected to hosts within k hops of the central IDS are included.

State: The state s is a vector that includes the classification vector (the output of the classifier employed by the agent) and the feature vector. Thresholding is used to create the classification vector from the confidence vector. The confidence vector is made up of the classifier's prediction.

Action: In the context of attack classification, an action is a decision made by an agent during a particular time window. Action vectors are created as a result of this process. This state vector s guides the construction of an action vector as follows: The state vector s is provided to the deep Q-network. A deep Q-network in the agents produces Q-values. A feature vector is classified as an attack if its Q-value is greater than a certain threshold value called the Q-threshold value. The decision vector is created by combining these classifications. A decision vector and a categorization vector are expressed as a logical AND operation by the action vector. Action vector is another term for agent result.

Reward: Rewards are abstract terms used to describe environmental inputs. Depending on an agent's confidence vector, reward vectors r may differ. Classifiers receive positive rewards when their classification results match real-life outcomes. If this is not the case, it will receive a negative reward. Based on the confidence vector, the reward is scaled.

The estimate of squared error of the total of the reward and the discounted estimated optimal future Q-value and old Q-value by the loss function,

$$L(\theta) = E[(TargetQ - Q(s,a;\theta))^2] \quad (4)$$

$$TargetQ = \gamma Q(s'a';\theta) \quad (5)$$

2.5 Secure Data Transmission using Ciphertext-Policy Attribute-Based Encryption

Step 1: A few security settings are input into the CP-ABE technique, resulting in the production of a public key (PK) and a master key (MK).

Step 2: A functional master key file (MKF) and a functional public key file (PKF) are generated using the functional encryption technique, security attributes are used as input

Step 3: Function f(i) is the input to KeyGen(MK,S) of functional encryption, which generates a secret key SK[f(i)], where i = 1, 2, ..., n. F(i) is defined as $f_i(S) = ss_i(KeyGen(MK,S)$, where $ss_i(s)$ is a function that generates the share when (n, k) used to share secrets. PKF and SK [f_i] are sent over the secure channel by the data owner to the data users and the i-th authority, respectively, after the setup algorithm is executed. It uses the precomputed subkeys of a large number, which are used for both encryption and decryption

2.5.1 Encryption Process

A positive integer value is supplied to parameter A, resulting in A≠ k×257, where k varies from 1 to n. Assume the array T, which has integers ranging from 0 to 255, resulting in 256 distinct integers. On the basis of A and T, a new array R is created using linear mapping, as shown in Equation (6):

$$R(i) = mod((A \times (T(i) + 1)), 257) \quad (6)$$

where i varies from 1 to 256. T(i) has values ranging from 0 to 255; positive integers A fulfils A≠ k×257, and k has a value larger than 0. There are two non-integer numbers that are not divisible by 257: (A/257) and (T(i) + 1)/257. This implies that mod ((A (T(i) + 1)), 257) is greater than zero. When R(i) is divided by R(i), then it has a range from 0 to 255, where I is between 1 and 256. To construct the first S-box, transform and multiply R = [R(i)] into a matrix Rb of 2D dimensions. Once the tent-logistic map has been repeated L times, a chaotic sequence of length L is formed. It is possible to increase the intensity of the chaotic series by removing the first (L-256) items, which leads to a novel chaotic series X with length 256. The sorting of X results in an index array J (which consists of J(1), J(2),..., J(256). Regardless of whether the series is periodic or ergodic, it always returns J(i) ≠ J(j), assuming I ≠ j.

Encryption is used on all data to protect data privacy using public and private keys (PK_i, SK_i). MAE calculates the value of each stream and exchanges H_j through MAE (mi) and MAE (H_j). The distributed ledger keeps track of every permitted transaction. Equation (7) provides MAE:

$$MAE(m_i) = \frac{1}{n}\sum_{i=1}^{n} \quad |y_i - f(x_i)| \quad (7)$$

Here, n denotes the entire sum of users, and x_i represents transmission and processing expenses, respectively. Sharing personal data is risky for data providers because of distinct security threats. The problem can be avoided by simply providing the data to the requester with the correct details and protecting the holders' privacy at the same time.

3. Experimental result and discussion

The simulation is done on a computer running Windows7, with an Intel core i7 processor, a CPU frequency of 6.4 GHZ, and the Python programming language used to test the effectiveness of the suggested approach. In this part, we shall assess the proposed network efficiency. The performances of the proposed classifier for attacks detection on IoT are compared withDeep Learning Neural Network (DLNN), Modified Adaptive neuro fuzzy inference system (MANFIS) classifier, Q-learning-based neural network with privacy preservation method (DQ-NNPP), andDeep Learning Modified Neural Network (DLMNN).

A number of parameters were selected for analysis. These included sensitivity, accuracy, specificity, f1 score, encryption time, communication overhead, and decryption times. The accuracy of this approach measures the ability to make accurate predictions. The formulation for accuracy is specified in Equation (8):

$$Accuracy = \frac{T^- + T^+}{T^- + T^+ + F^- + F^+} \quad (8)$$

As defined in Equation (9), sensitivity is calculated by the percentage of true positives, which represents the right identification during testing.

$$Sensitivity = \frac{T^+}{T^+ + F^-} \quad (9)$$

Similarly, specificity is determined as the ratio of true negatives indicated in Equation (10), which indicates the classifier's correct identification.

$$Specificity = \frac{T^-}{T^- + F^+} \quad (10)$$

Finally, the F1-score assessment is essential to find a compromise between accuracy and recall, as shown in Equation (11).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

Table 2 Accuracy of the algorithms

Algorithms	**Accuracy**	
	With feature selection	**Without feature selection**
MANFIS Classifier	86.47	83.39
DLNN	89.45	81.67
DLMNN	91.98	86.56
DQ-NNPP	93.86	89.45
Proposed work	99.65	95.77

Table 3 Sensitivity of the algorithms

Algorithms	**Sensitivity**	
	With feature selection	**With feature selection**
MANFIS Classifier	82.48	79.37
DLNN	83.78	81.49
DLMNN	85.34	82.44
DQ-NNPP	93.58	87.23
Proposed work	96.37	91.45

Table 4 Specificity of the algorithms

Algorithms	**Specificity**	
	With feature selection	**Without feature selection**
MANFIS Classifier	86.49	82.52
DLNN	89.45	84.35
DLMNN	92.63	87.34
DQ-NNPP	94.35	89.67
Proposed work	96.36	92.49

Table 5 F1 score of the algorithms

Algorithms	**F1 score**	
	With feature selection	**Without feature selection**
MANFIS Classifier	89.38	83.89
DLNN	91.46	86.45
DLMNN	93.78	89.75
DQ-NNPP	95.43	92.47

Proposed work	97.48	93.56

The effectiveness of the proposed and existing algorithms is exposed in Table 2 to 5. The accuracy obtained by the proposed work is compared with other algorithms to evaluate the efficiency of the presented work. The accuracy obtained by the proposed system is 99.65% with feature selection and 95.77% without feature selection. By feature selection accuracy is 3.88% higher than without feature selection. Because of inappropriate and duplicate features are removed by the attribute selection procedure, the attribute dimension is reduced, hence improving the accuracy and efficacy of classification, and it also removes noisy data and minimises the overfitting in deep learning. The accuracy obtained by the other algorithms with feature selection are 86.47%, 89.45%, 91.98%, and 93.86% by MANFIS Classifier, DLNN, DLMNN, and DQ-NNPP respectively. The accuracy obtained by MANFIS, DLNN, DLMNN, and DQ-NNPP without feature selection is 83.39%, 81.67%, 86.56%, and 89.45% respectively. The sensitivity of MANFIS, DLNN, DLMNN, DQ-NNPP, and Proposed work with feature selection is 82.48%, 83.78%, 85.34%, 93.58%, and 96.37% and sensitivity without feature selection is 79.37%, 81.49%, 82.44%, 87.23%, 91.45% respectively.

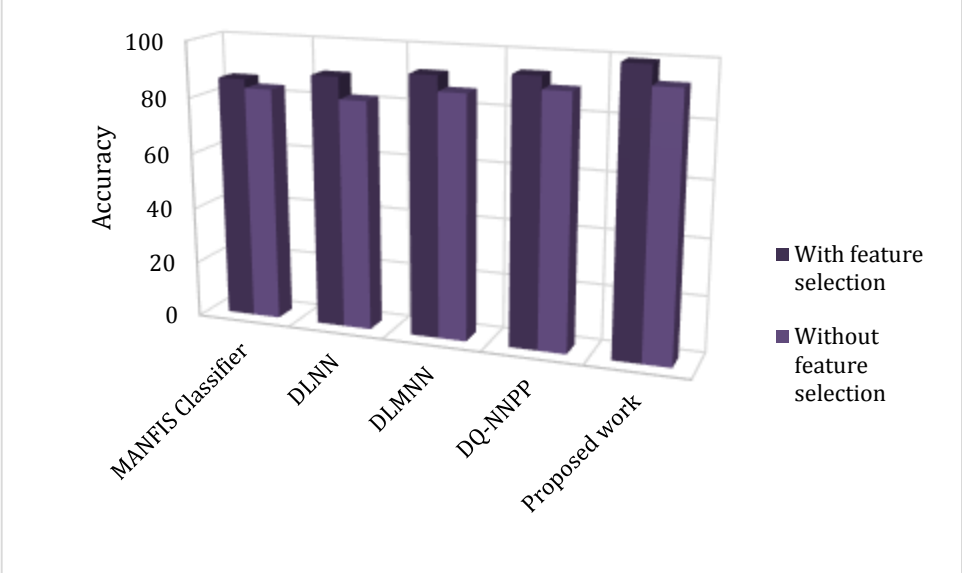

Figure 2 Accuracy of the algorithms

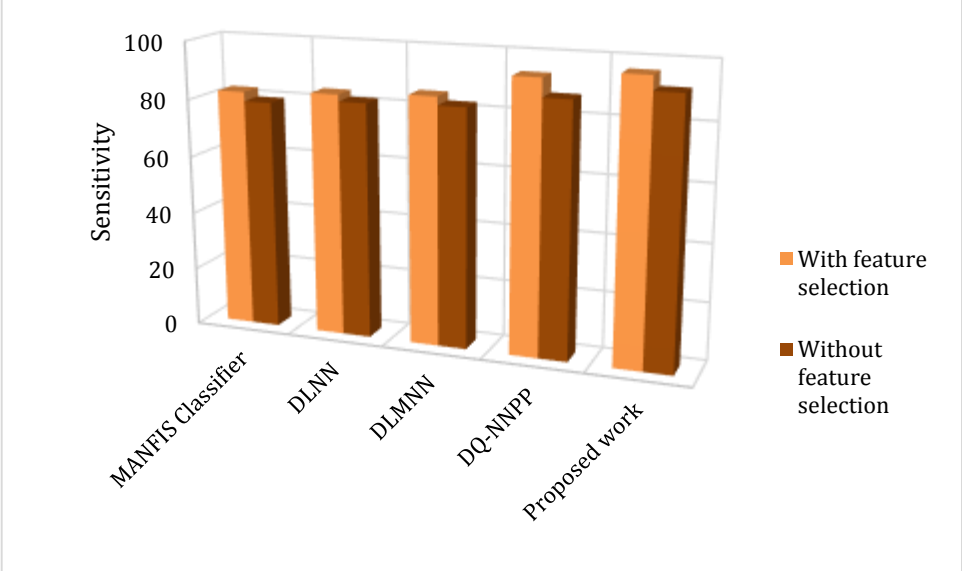

Figure 3 Sensitivity of the algorithms

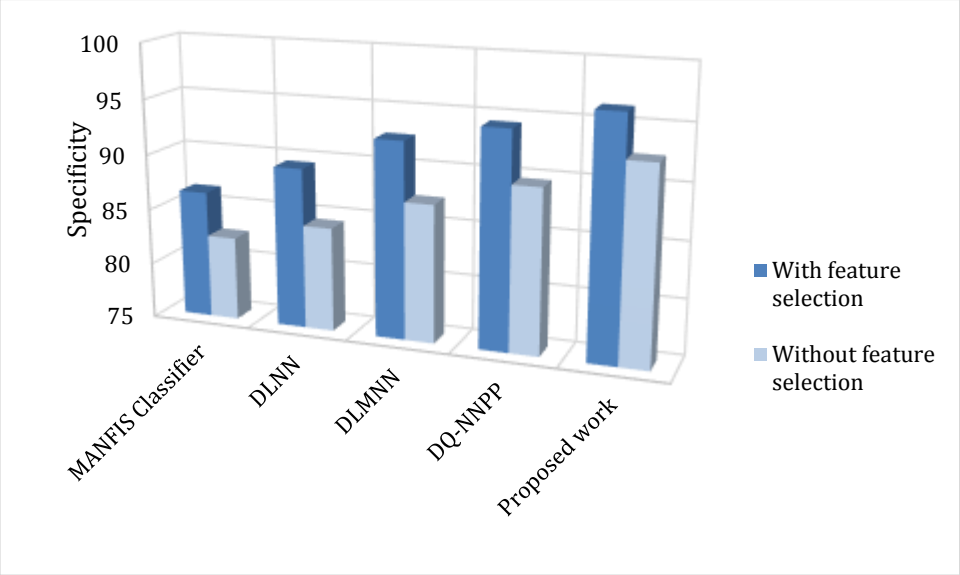

Figure 4 Specificity of the algorithms

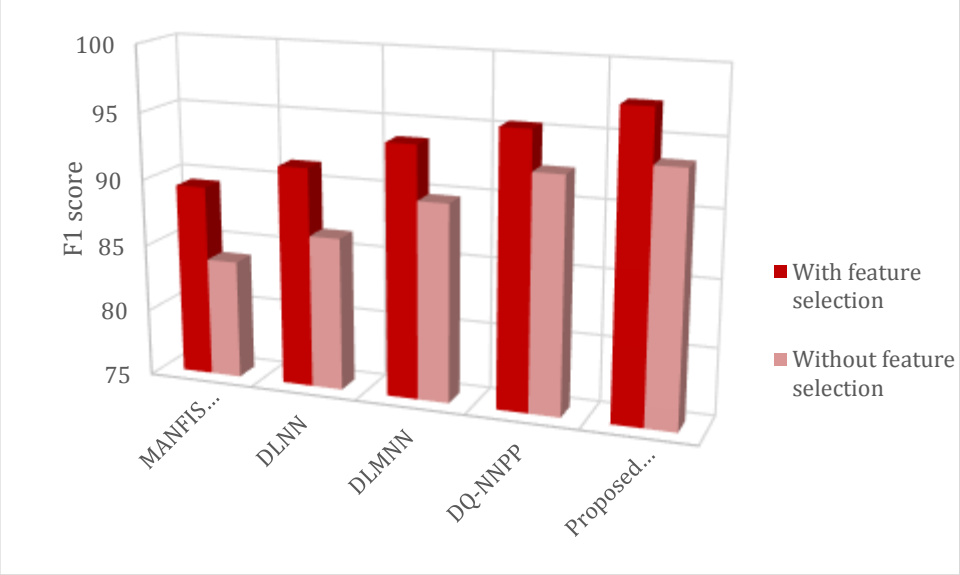

Figure 5 F1 of the algorithms

From the Figures 2 to 5, it is observed that the proposed work performs well in terms of accuracy, sensitivity, specificity and F1 score. The specificity with feature selection for MANFIS, DLNN, DLMNN, DQ-NNPP, and Proposed work is 86.49%, 89.45%, 92.63%, 94.35%, and 96.36% respectively. Without feature selection the specificity is 82.52%, 84.35%, 87.34%, 89.67%, and 92.49% respectively. Also, the F1 score is improved with feature selection which is 89.38%, 91.46%, 93.78%, 95.43%, and 97.48% by the MANFIS, DLNN, DLMNN, DQ-NNPP and the Proposed work. Without feature selection 83.89%, 86.45%, 89.75%, 92.47%, and 93.56% of F1 score. The increase in performance of the classifiers is due to selecting important features from the dataset,which can eliminate unnecessary features to enhance a classifier's classification performance and decrease training and inference time.

Performance of the encryption algorithms

Communication overhead, encryption time, and decryption time were chosen as factors for analysis. On the basis of these parameters, the proposed work was compared to three different conventional techniques, namely, MSCryptoNet, privacy-preserving disease prediction (PPDP), the secure and anonymous biometric based user authentication scheme (SAB-UAS), anddeep Q-learning-based neural network with the privacy preservation method (DQ-NNPP)

The proportion of overall packets (Npack) transferred from node x to node y in a lesser amount of time is defined as communication

overhead (C). The communication overhead formula is stated in Equation (12):

$$\sum_{0}^{N_{pack}} \quad x \rightarrow y \qquad (12)$$

The Encryption Time (E) is the time allotted by the algorithms to convert plain text (P) into cypher text (C) using symmetrical or asymmetrical keys. Equation (13) defines the encryption time formula:

$$E = Time(P \rightarrow C) \qquad (13)$$

Decryption Time (D) is the amount of time required by the algorithms to convert cypher text (P) to plain text (C) utilising symmetric or asymmetric keys. Equation (14) provides the encryption time formula:

$$D = Time(C \rightarrow P) \qquad (14)$$

Table 6 Performance comparison of communication overhead, encryption and decryption time.

Parameters	**SAB-UAS**	**MSCryptoNet**	**PPDP**	**DQ-NNPP**	**Proposed**
Communication Overhead (%)	64.82	65.87	66.92	67.82	69.72
Encryption Time (ms)	66.49	64.28	63.67	61.73	58.83
Decryption Time (ms)	67.76	65.75	64.87	62.47	59.58

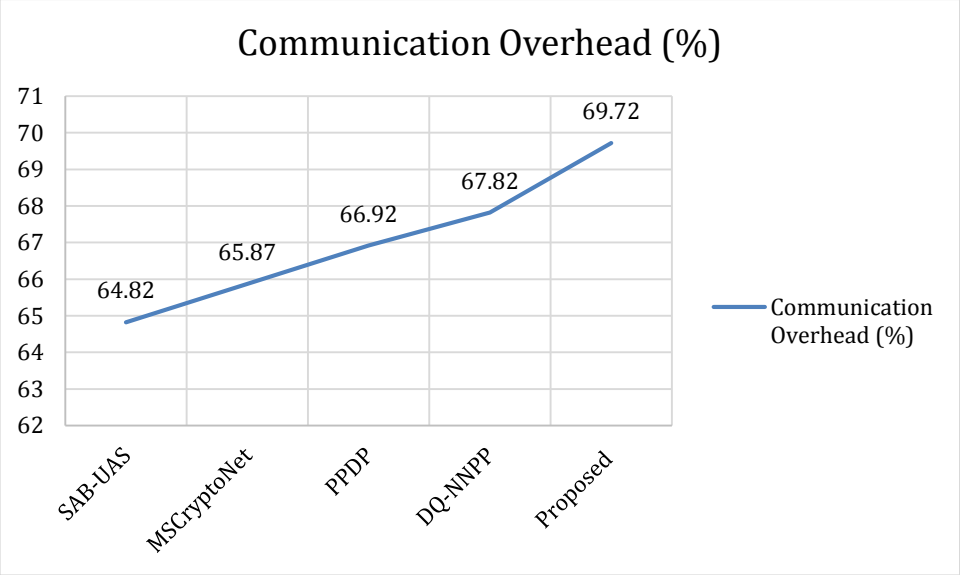

Figure 6 Comparison of communication overhead

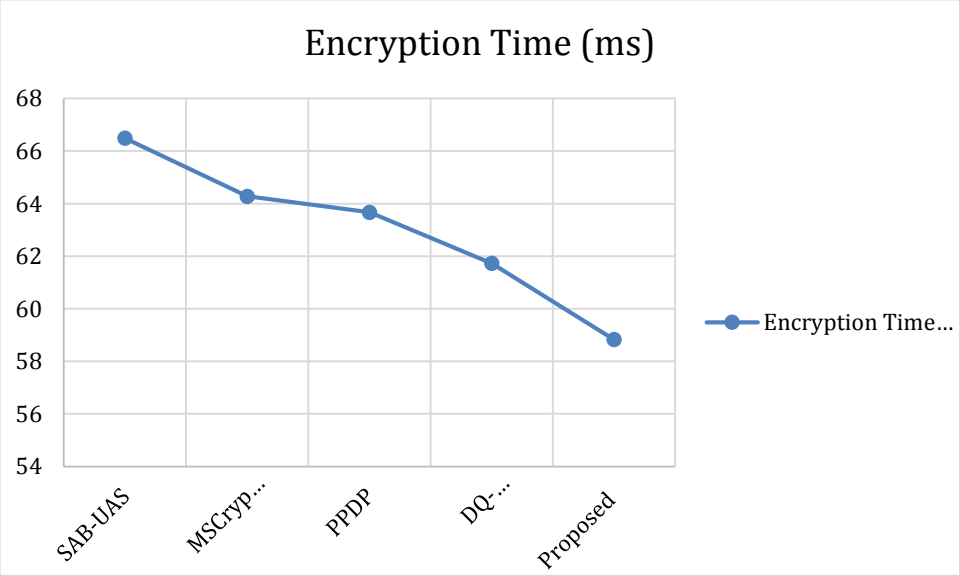

Figure 7 Comparison of encryption time

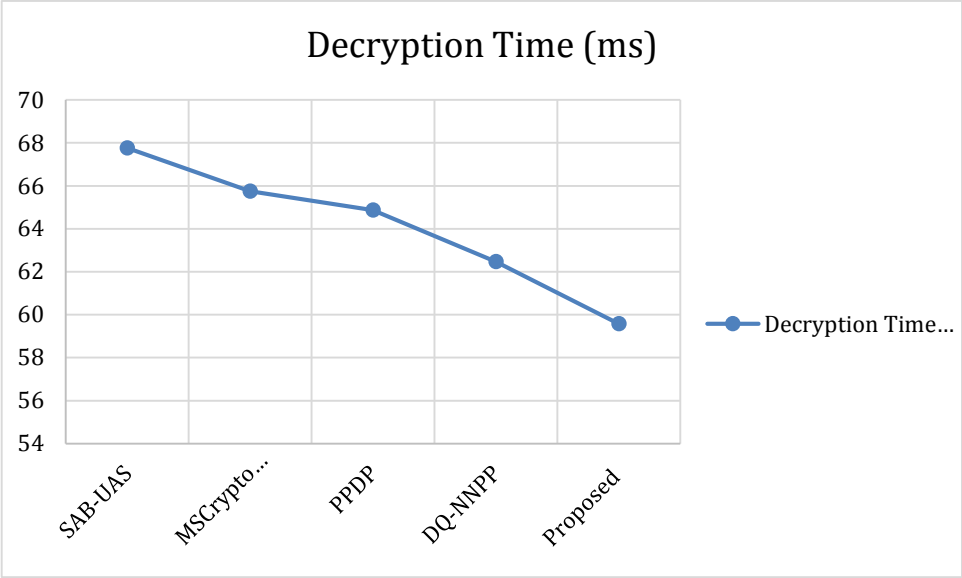

Figure 8 Comparison of decryption time

From the Table 6, the comparison of the proposed encryption algorithm with other algorithms are provided. The graphical representation is presented in the Figures 6 to 8. From the analysis, it is clearly apparent that the proposed work achieves well in all aspects. The communication overhead of the presented work is 69.72 % which is 4.9% higher than SAB-UAS, 3.9% higher than MSCryptoNet, 2.8% higher than PPDP, and 1.9% higher than DQ-NNPP. The encryption time of the proposed system is 58.83 ms which is less compared to other algorithms. The encryption time of SAB-UAS, MSCryptoNet, PPDP, and DQ-NNPP is 66.49 ms, 64.28 ms, 63.67 ms and 61.73 ms respectively. The decryption time of SAB-UAS, MSCryptoNet, PPDP, DQ-NNPP, and Proposed system is 67.76 ms, 65.75 ms, 64.87 ms, 62.47 ms and 59.58 ms respectively.

4. Conclusion

The most difficult concerns in IoT applications are security and privacy. This research developed a reinforcement learning approach for data security and a data encryption method for safely transmitting information over a network, which preserves data privacy when information is being shared. The NSL-KDD dataset is utilised for training the reinforcement learning algorithm, and the key characteristics are chosen from the dataset using the quantum computing model, which improves the trained model's performance. By using Ciphertext-Policy Attribute-Based Encryption, information is securely sent across the network. The sensitivity, accuracy, specificity, F1 score, communication overhead, encryption time, and decryption time of the proposed system are all examined. Experimental results shows that the proposed system performed efficiently and the data is protected and securely transmitted in the network.

5. References

F. Olowononi, D. B. Rawat, and C. Liu, "Resilient Machine Learning for Networked Cyber Physical Systems: A Survey for Machine Learning Security to Securing Machine Learning for CPS," IEEE Communications Surveys and Tutorials, 2020. Early Access, DoI: https://doi.org/10.1109/COMST.2020.3036778 .

Jan, M. A., Zhang, W., Usman, M., Tan, Z., Khan, F., & Luo, E. (2019). SmartEdge: An end-to-end encryption framework for an edge-enabled smart city application. Journal of Network and Computer Applications, 137, 1-10. https://doi.org/10.1016/j.jnca.2019.02.023

Tang, J.; Liu, A.; Zhao, M.; Wang, T. An Aggregate Signature Based Trust Routing for Data Gathering in Sensor Networks. Secur. Commun. Netw. 2018, 2018, 6328504.

Meneghello, F., Calore, M., Zucchetto, D., Polese, M., &Zanella, A. (2019). IoT: Internet of threats? A survey of practical security vulnerabilities in real IoT devices. IEEE Internet of Things Journal, 6(5), 8182-8201. https://doi.org/10.1109/jiot.2019.2935189

Uprety, A., &Rawat, D. B. (2021). Reinforcement learning for IoT security: A comprehensive survey. IEEE Internet of Things Journal, 8(11), 8693-8706. https://doi.org/10.1109/jiot.2020.3040957

Thuluva, A.S.S., Somanathan, M.S., Somula, R. et al. Secure and efficient transmission of data based on Caesar Cipher Algorithm for Sybil attack in IoT. EURASIP J. Adv. Signal Process. 2021, 38 (2021). https://doi.org/10.1186/s13634-021-00748-0

Jebri, S., Ben Amor, A., Abid, M. et al. Enhanced Lightweight Algorithm to Secure Data Transmission in IoT Systems. Wireless PersCommun 116, 2321–2344 (2021). https://doi.org/10.1007/s11277-020-07792-3

Jian, MS., Wu, J.MT. Hybrid Internet of Things (IoT) data transmission security corresponding to device verification. J Ambient Intell Human Comput (2021). https://doi.org/10.1007/s12652-021-03122-y

Kandhoul, N., Dhurandher, S.K. An Efficient and Secure Data Forwarding gh;

Mechanism for Opportunistic IoT. Wireless PersCommun 118, 217–237 (2021). https://doi.org/10.1007/s11277-020-08010-w

Mondal, S., Ghosh, I. & Das, A. Energy efficient and secure healthcare data transmission in the internet of medical things network. MicrosystTechnol (2022). https://doi.org/10.1007/s00542-022-05398-2

Refaee, E., Parveen, S., Begum, K. M., Parveen, F., Raja, M. C., Gupta, S. K., & Krishnan, S. (2022). Secure and scalable healthcare data transmission in IoT based on optimized routing protocols for mobile computing applications. Wireless Communications and Mobile Computing, 2022, 1-12. https://doi.org/10.1155/2022/5665408

Sankar, S., &Karthiga, I. (2020). An IoT-based secure data transmission in WBSN. International Journal of Cloud Computing, 9(2/3), 311. https://doi.org/10.1504/ijcc.2020.10031548

El-Latif, A. A. A., Iliyasu, A. M., &Abd-El-Atty, B. (2021). An efficient visually meaningful quantum walks-based encryption scheme for secure data transmission on IoT and smart applications. Mathematics, 9(23), 3131. doi:10.3390/math9233131

Sokol, I., Hubinský, P., &Chovanec, Ľ. (2021). Lightweight cryptography for the encryption of data communication of IoT devices. Electronics, 10(21), 2567. https://doi.org/10.3390/electronics10212567

SakthidasanSankaran, K., & Kim, B. (2023). Deep learning based energy efficient optimal RMC-CNN model for secured data transmission and anomaly detection in industrial IOT. Sustainable Energy Technologies and Assessments, 56, 102983. https://doi.org/10.1016/j.seta.2022.102983

Kumar, P., Kumar, R., Gupta, G. P., Tripathi, R., Jolfaei, A., &Najmul Islam, A. (2023). A blockchain-orchestrated deep learning approach for secure data transmission in IoT-enabled healthcare system. Journal of Parallel and Distributed Computing, 172, 69-83. https://doi.org/10.1016/j.jpdc.2022.10.002

Premkumar, R., &SathyaPriya, S. (2022). Service constraint NCBQ trust Orient secure transmission with IoT devices for improved data security in cloud using blockchain. Measurement: Sensors, 24, 100486. https://doi.org/10.1016/j.measen.2022.100486

Jan, A., Parah, S. A., Malik, B. A., & Rashid, M. (2021). Secure data transmission in IoTs based on clog edge detection. Future Generation Computer Systems, 121, 59-73. https://doi.org/10.1016/j.future.2021.03.005

Hurrah, N. N., Parah, S. A., Sheikh, J. A., Al-Turjman, F., & Muhammad, K. (2019). Secure data transmission framework for confidentiality in IoTs. Ad Hoc Networks, 95, 101989. https://doi.org/10.1016/j.adhoc.2019.101989

Hou, J, Li, Q, Cui, S, Meng, S, Zhang, S, Ni, Z & Tian, Y 2020, 'Low-cohesion differential privacy protection for industrial internet', The Journal of Supercomputing, vol. 76, no. 11, pp. 8450-8472.

He, Z., Li, L., Huang, Z., & Situ, H. (2018). Quantum-enhanced feature selection with forward selection and backward elimination. Quantum Information Processing, 17(7). https://doi.org/10.1007/s11128-018-1924-8