

A Hybrid Quantum-Inspired Deep Reinforcement Learning Framework for Adaptive Scheduling and Energy-Efficient LoRaWAN in Smart IoT Networks

Shaista Tarannum

Department of Electronics and Communication Engineering, JSSATE, Bengaluru, Karnataka, India | Visvesvaraya Technological University, Belagavi, India | Department of Computer Science and Engineering, MSRUEAS, Bengaluru, Karnataka, India
shaistatarannum123@gmail.com (corresponding author)

Usha Shiramally Mallappa

Department of Electronics and Communication Engineering, JSSATE, Bengaluru, Karnataka, India | Visvesvaraya Technological University, Belagavi, India
ushasm@jssateb.ac.in

Received: 20 September 2025 | Revised: 31 October 2025 | Accepted: 10 November 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.14958>

ABSTRACT

The rapid deployment of Internet-of-Things (IoT) applications in smart city environments has significantly increased the demand for energy-aware and reliable long-range communication solutions. Long-Range Wide-Area Network (LoRaWAN) is one of the most promising IoT technologies and is widely adopted in low-power wide area networks for large-scale deployments. However, LoRaWAN faces scalability issues due to the large number of nodes connected to the same gateway or sharing the same channel. Conventional adaptive data rate and channel allocation strategies often fail to balance scalability, reliability, and energy consumption in highly dynamic LoRaWAN networks. This study, therefore, introduces a quantum-inspired hybrid Double Deep Q-Network (DDQN) and Variational Quantum Circuit (VQC) framework to perform intelligent transmission scheduling and adaptive policy optimization in large-scale LoRaWAN networks. The framework initially leverages a DDQN-based learning agent to select optimal transmission configurations, where it jointly optimizes the spreading factors, channel assignments, and wait actions to reduce energy consumption while maintaining a high Packet Delivery Ratio (PDR). The framework further integrates a quantum-enhanced policy adaptation module that employs a VQC-based decision layer to encode network states into a high-dimensional Hilbert space, enabling improved exploration of the action space and superior adaptation to varying traffic congestion conditions. Comprehensive simulation results show that the proposed hybrid framework outperforms classical Deep Reinforcement Learning (DRL) strategies for adaptive resource scheduling. The results further demonstrate that the hybrid DDQN+VQC achieves up to 28.4% reduction in energy consumption and a 21.7% improvement in PDR under dense IoT deployments. In addition, the proposed hybrid policy positively influences latency performance in LoRaWAN. Overall, the results demonstrate the effectiveness of integrating quantum-inspired policy adaptation with DRL for scalable, energy-efficient, and reliable LoRaWAN optimization.

Keywords-Double Deep Q-Networks (DDQNs); Variational Quantum Circuit (VQC); LoRaWAN optimization; adaptive transmission scheduling; reliability-aware resource allocation; energy efficiency

I. INTRODUCTION

The use of wireless technologies has grown rapidly in recent years, especially the Internet-of-Things (IoT), where countless devices communicate and share data remotely. By 2027, the number of IoT devices, ranging from short-range sensors and smart home gadgets to cellular and wide-area

network systems, is expected to grow to approximately 30.2 billion. This clearly highlights the massive scale and impact of IoT devices on daily life, business operations, and global connectivity [1]. Low-Power Wide-Area Network (LPWAN) technologies have become highly popular for providing cost-effective, energy-efficient, and long-range connectivity for IoT devices. Their low power consumption, extended coverage, and

adequate data rate make them particularly well-suited for IoT-based smart city applications [2, 3].

Some popular technologies of LPWAN include Long-Range Wide-Area Network (LoRaWAN), SigFox, Long-Term Evolution for Machines (LTE-M), and Narrowband IoT (NB-IoT), where LTE-M and NB-IoT operate in licensed bands, offering reliable communication and higher data rates, but their access protocols often lead to frequent retransmissions and higher energy consumption. In contrast, LoRaWAN and Sigfox utilize unlicensed Industrial, Scientific, and Medical (ISM) bands, which lowers deployment costs and supports large private and public networks, but they require advanced interference management. Therefore, to support massive IoT access, these networks must use their channels and radio resources efficiently [4]. However, with the explosive growth of IoT devices, LoRaWAN faces critical challenges related to scalability, dynamic network conditions, energy efficiency, and transmission reliability. These problems are further intensified in dense deployments, where uncoordinated transmissions can result in severe collisions, increased packet delays, and reduced network lifetime [5, 6]. Traditional resource allocation and scheduling mechanisms in LoRaWAN are primarily static and struggle to efficiently adapt to varying traffic conditions [7, 8]. Recently, Deep Reinforcement Learning (DRL) techniques, particularly Deep Q-Networks (DQNs) and Double Deep Q-Networks (DDQNs), have been explored for adaptive transmission scheduling and energy-aware optimization [9, 10]. They have gained significant attention from the research community due to their ability to accomplish stable and high-performance decision making in complex environments, particularly in wireless IoT networks and resource-constrained systems such as LoRaWAN smart cities [11].

However, although DDQNs effectively mitigate temporal-difference overestimation, which is a limitation of DQNs, they still face challenges when dealing with high-dimensional state-action spaces and dynamic scheduling under strict latency and energy constraints. Moreover, while DQNs improve decision-making compared to conventional heuristics, they often suffer from slow convergence, overfitting, and limited generalization, particularly when applied to large state-action spaces. These limitations highlight the need to further enhance DDQN-based frameworks to enable consistent optimization of spreading factors, transmission power, and duty cycles under varying network conditions. In this context, Variational Quantum Circuits (VQCs) [12, 13] have emerged as a promising paradigm in quantum-inspired reinforcement learning, offering improved policy generalization and faster convergence by leveraging quantum state superposition and entanglement to enhance adaptive exploration [14, 15].

Authors in [16] presented a lightweight DDQN-based lightweight scheduling mechanism for energy-efficiency and packet reliability in industrial IoT environments. While effective, the approach lacks support for LoRaWAN-specific challenges such as dynamic data rate adaption, congestion handling, and scalability across large-scale deployments. Authors in [17] presented an auction-based resource allocation and collision control strategy for dense LoRaWAN using a DRL approach integrated with TCP to improve fairness and

channel utilization. While their model effectively reduces packet collisions and enhances throughput, the study considers fixed patterns and does not adapt to varying traffic in large-scale LoRaWAN networks. The study also does not consider quantum-inspired policies for optimizing transmission scheduling under high congestion.

Authors in [18] presented a DRL-based resource allocation framework for dense sliced LoRaWAN networks that aims to enhance network scalability and spectral efficiency. The approach dynamically allocates spreading factors and transmission power to optimize the network capacity while also minimizing collisions. However, the method relies on standard DQN, which often suffers from temporal difference overestimation and slow convergence in highly dynamic traffic scenarios. Moreover, quantum-inspired optimization and varying traffic handling are not explicitly discussed.

Authors in [19] presented a Multi-Agent Reinforcement Learning (MARL) framework to optimize energy efficiency in LoRaWAN-based Wireless Underground Sensor Networks (WUSNs). The approach jointly learns transmission power and spreading factor adaption strategies and achieves significant energy savings under constrained underground environments. However, the focus of the work remains on static, homogenous traffic patterns and lacks mechanisms for scalable resource allocation under dynamically varying network conditions. It should be noted that in the presented approach, energy consumption reduction is mainly driven by learning node-specific channel conditions and optimizing transmission parameters jointly. Author in [20] introduced a foundational DRL-based resource allocation model for LoRaWAN. The study presents a simple single-agent based DQN model, where the state representation is limited to spreading factor, transmission power, and channel selection. The approach does not support varying traffic and also suffers from state-space explosion as network density grows. It also fails to jointly optimize energy efficiency and Packet Delivery Ratio (PDR) while maintaining considerable latency.

Authors in [21] conducted an experimental study using LoRaWAN with reinforcement learning-based scheduling for water quality monitoring systems. The presented approach dynamically adjusts uplink timing and spreading factors to extend sensor lifetime, and it also maintains data fidelity under real-world deployment constraints. While the experimental validation is compelling, the presented work focuses on small-scale, environment-specific deployments, lacks policy generalization under varying traffic, and also relies on fully classical reinforcement learning, which limits its scalability. Authors in [22] presented a frequency-hopping scheduling algorithm for energy efficient LoRaWAN networks using a reinforcement learning approach. Their method dynamically selects channel hopping patterns to mitigate interference and extend battery life in static deployments. However, the work is limited to single-device or small-scale setups and lacks the capacity to adapt to varying traffic loads. Here, the employment of a classical reinforcement learning strategy results in slower convergence in large and complex state-action spaces. Finally, authors in [23] discuss smart city scenarios for LoRaWAN to limit the number of nodes.

After reviewing existing LoRaWAN resource allocation and scheduling techniques, it is observed that the majority of classical reinforcement learning approaches suffer from slow convergence, limited adaptability to varying traffic conditions, and inefficiency in energy–reliability trade-offs, especially in dense and dynamic IoT environments. Therefore, there is a need to design an effective and computationally efficient transmission scheduling framework that can perform adaptive scheduling and resource allocation while retaining balanced optimization across energy consumption, PDR, and latency and also ensuring scalability and reliability for future IoT deployments.

The proposed work addresses the limitations of existing adaptive data rate and DRL approaches in intelligent resource allocation for LoRaWAN and introduces a hybrid quantum–classical optimization framework that integrates VQCs with a DDQN to enhance transmission scheduling. The proposed hybrid DDQN-VQC architecture leverages quantum-enhanced policy adaptation where VQC-based parameterized quantum gates are trained jointly with the DDQN agent. In the proposed work, the quantum component encodes high-dimensional state representations into a quantum Hilbert space and exploits quantum superposition and entanglement to accelerate exploration, improve policy diversity, and enhance decision quality under uncertainty. This allows the framework to achieve better energy efficiency, lower packet loss, and improved scalability compared with classical DRL approaches.

The contributions of this paper are summarized as follows:

- The proposed work designs a hybrid quantum–classical reinforcement learning framework that combines DDQN with a parameterized VQC-based policy network to optimize the spreading factor, transmission power, and duty cycle in LoRaWAN.
- We develop and evaluate a simulation-based framework for assessing energy consumption, reliability, and packet success rates under varying network densities and traffic patterns.
- The experimental evaluation demonstrates that the proposed framework outperforms state-of-the-art DRL-based approaches in terms of energy efficiency, PDR, and adaptability while maintaining scalability in large-scale LoRaWAN scenarios.
- The proposed hybrid policy-based framework is designed to handle varying traffic scenarios by adapting transmission configurations and scheduling policies. The quantum-inspired VQC module also enables faster convergence under varying network conditions, making the framework scalable to mixed traffic environments.

By combining DDQN with quantum-inspired policy adaptation, this work contributes to the advancement of next-generation energy-aware, reliable, and scalable LoRaWAN systems capable of supporting large-scale IoT deployments. The conceptualization of the proposed work is novel in integrating DDQN with VQC for adaptive spreading factor selection and for jointly optimizing energy, reliability (PDR), and delay in dense LoRaWAN networks. It introduces a state

representation capable of capturing battery, queue, and congestion levels for precise action selection. To the best of our knowledge, this is a unique quantum-enhanced DRL framework for energy-aware, scalable LoRaWAN scheduling in smart IoT networks.

II. SYSTEM MODEL

The proposed work considers a LoRaWAN smart city deployment with N end devices, such as $d_i = \{d_1, d_2, \dots, d_N\}$, along with a LoRa gateway G . It is also assumed that the deployment may include more than one LoRa gateways. Each device is assumed to generate packets periodically with data rate R_i . This implies that each IoT end device d_i produces data at a certain rate, which represents the number of packets (or bits) generated per unit time. The proposed work also introduces a set of notations and parameters that are used to model the Artificial Intelligence and Quantum Machine Learning (AI/QML)-driven adaptive framework for scalable and energy-efficient LoRaWAN, as discussed below:

- $E_i(t)$: It refers to the remaining energy of end device d_i at time t , which indicates the residual battery capacity that decreases with each transmission event. The prime reason for selecting the residual energy $E_i(t)$ as a key parameter is that LoRaWAN devices are typically battery-powered; therefore, their lifetime is directly dependent on their remaining energy. By modeling $E_i(t)$, the system can track the energy profiles of LoRaWAN end devices over time and make optimization decisions that extend the network lifetime. Without $E_i(t)$, the system cannot properly evaluate whether the scheduling policy is energy-efficient.
- $T_i(t)$: It represents the transmission power of device d_i . The proposed work considers this parameter to estimate the transmission power level of device d_i at time t . This variable influences the power control in LoRaWAN and helps balance energy consumption and reliability. If the transmission power is too low, packet loss may increase; conversely, if it is too high, the end device may waste energy and increase interference. By considering $T_i(t)$, the proposed work aims to develop an optimization solution that adapts transmission power levels rather than fixing them.
- SF_i : The spreading factor SF_i assigned to device d_i is also considered during the design and modeling of the proposed system. A higher spreading factor enables longer communication range but incurs higher energy cost and increased congestion. In contrast, a lower spreading factor results in faster transmission but reduced coverage. Thus, spreading factor is a key MAC parameter to optimize in LoRaWAN. Therefore, to model the trade-off between coverage and energy, the spreading factor plays a crucial role in LoRaWAN.
- $PDR(t)$: PDR at time t is a key performance parameter and indicator of network reliability. It is defined as the ratio of successfully received packets to the total transmitted packets within a given time window. The proposed system aims to design the learning agent such that it maintains a high PDR while minimizing energy consumption. It serves

as a Quality-of-Service (QoS) constraint in the proposed model, ensuring that energy efficiency does not come at the cost of packet loss.

- $\sigma(t)$: This notation represents the network congestion level at time t , measured as the ratio of collided packets to the total transmitted packets. Since LoRaWAN is based on the ALOHA protocol, collisions increase as the number of devices grows. Therefore, congestion modeling is essential to capture scalability limitations. If not included, the learning model may suggest aggressive scheduling strategies that fail under high collision conditions.

A. Energy-Aware Transmission Model for LoRaWAN

In the proposed system, the energy consumption of each end device is modeled to track and assess the impact of transmission power, spreading factor, and packet size on battery drain. For an end device d_i at time t , the transmission energy is estimated as follows:

$$E_i^{tx}(t) = P_i^{tx} \cdot T_{sym}(SF_i) \cdot L_i \quad (1)$$

Here, P_i^{tx} represents the transmission power of device d_i (W). $T_{sym}(SF_i)$ represents the symbol time (in seconds), which is determined by the spreading factor as $2^{SF_i}/BW$, where BW refers to the channel bandwidth. Additionally, L_i represents the packet length (in symbols). In LoRaWAN, energy is mostly consumed during transmission rather than in idle or receive modes. This model captures three important factors: transmission power, spreading factor, and packet length. With P_i^{tx} , devices can select power levels depending on communication range. LoRa modulation uses a spreading factor SF_i , typically ranging from 7 to 12. A higher spreading factor results in longer symbol duration and, consequently, higher energy consumption. Similarly, a larger packet length L_i increases the number of symbols required, leading to longer transmission time. This formulation aligns with the physical layer specifications of LoRa and captures the trade-off between energy consumption and spreading factor.

The proposed work also models the residual energy after transmission as follows:

$$E_i(t+1) = E_i(t) - E_i^{tx}(t) \quad (2)$$

Since LoRaWAN devices are battery-powered, modeling residual energy is essential for evaluating network lifetime, which is a key metric in IoT systems. In (2), $E_i(t)$ represents the energy of end device d_i at time t . The proposed work also defines energy efficiency as the ratio of successfully delivered packets to the total transmission energy consumed across all devices. The total network energy consumption is computed as follows:

$$E_{total} = \sum_{i=1}^N E_i \quad (3)$$

The primary objective is to minimize the total energy consumption while satisfying the reliability constraint:

$$\min E_{total} \text{ s.t. } PDR \geq \varepsilon \quad (4)$$

Here, ε denotes the minimum acceptable PDR, representing the reliability constraint.

1) Scalability and Reliability Model

The proposed work assumes that the optimal transmission scheduling is designed for dense deployments in LoRaWAN, where network reliability may be constrained by collisions and congestion. In dense IoT networks, there may be hundreds or thousands of devices, and many devices may attempt to send packets at the same time. This aspect is crucial in modeling the proposed system, as without it, the impact of packet collisions on reliability cannot be evaluated as the network scales.

The collision probability for device d_i is modeled as follows:

$$P_{col,i} = 1 - \prod_{j \neq i} \left(1 - \frac{L_j}{T_{frame}}\right) \quad (5)$$

Equation (5) is derived from probability theory considering multiple overlapping transmissions, which is applicable in dense LoRaWAN deployments where many devices share the same channel. Here, $P_{col,i}$ represents the probability that a transmission from device d_i collides with transmissions from other devices. The product is taken over all devices except i . L_j is the transmission duration of the packet sent by device j . In LoRa, this duration depends on the spreading factor, packet length, and other modulation parameters. The term T_{frame} represents the frame duration or observation window, whereas $(1 - L_j/T_{frame})$ computes the probability that device j 's transmission does not overlap with that of device i within the frame. The equation first computes the probability of safe (non-overlapping) transmissions and then derives the probability of at least one collision as its complement. This is modeled as in LoRaWAN deployments many devices share the same channel and also collisions negatively impact the PDR. Therefore, modeling collision probability is essential for designing intelligent scheduling, slot allocation, and adaptive spreading factor or power control policies.

The PDR is defined as: $PDR = N_{success}/N_{tx}$. The proposed work aims to maximize reliability (PDR) while ensuring that the collision probability remains below a predefined threshold θ . This constraint ensures that the system remains scalable, allowing a large number of devices to coexist while maintaining reliable communication.

2) Objective Function

The proposed work aims to jointly minimize the energy consumption and maximize network reliability (PDR) under scalability constraints:

$$\min_{\pi} \mathbb{E}[\alpha \cdot E_{total}(t) - \beta \cdot PDR(t)] \quad (6)$$

subject to:

$$E_i(t+1) = E_i(t) - (P_i^{tx}(s,p) \cdot T_{on}(s)), \forall i \in N$$

$$P_{col,i} < \theta, PDR \geq \varepsilon$$

where:

- $P_i^{tx}(s,p)$ represents the transmission power for a given spreading factor SF_i and transmission power level p .

- $T_{on}(s)$ represents the on-air time, i.e. the duration for which the radio remains active during transmission.
- α and β are weighting factors that balance the trade-off between energy consumption and reliability.

B. Energy-Aware Double Deep Q-Network for Transmission Scheduling

The proposed work initially applies a DDQN to jointly optimize the trade-off between energy consumption and network reliability (PDR) in LoRaWAN. The DDQN-based framework is designed to achieve energy-efficient optimization in smart city environments. The study adopts DDQN instead of Deep Q-Network (DQN) because LoRaWAN environments are highly dynamic, with varying channel conditions, device distances, and residual energy levels of end devices d_i . DDQN helps reduce overestimation bias in Q-value estimation and learns a more stable mapping from states to optimal actions compared to DQN.

The system considers a deployment of end devices $d_i = \{d_1, d_2, \dots, d_N\}$. LoRaWAN operates with two key tunable parameters: spreading factor (SF_i) and channel (CH). The spreading factor SF_i determines the duration of packet transmission and the robustness of signal decoding at the receiver (gateway G), whereas the channel (CH) specifies the frequency band used for transmission. The transmission scheduling problem for massive IoT deployment over LoRaWAN is formulated with the objective of optimizing energy consumption while maintaining a high PDR. The proposed approach uses DDQN to select optimal actions for each device, considering:

- Spreading factor (SF).
- Channel (CH).
- A "wait" option (idle mode to save energy).

The model also incorporates the normalized distance between device i and gateway G , denoted as δ_i , residual energy $E_i(t)$, and a data importance indicator $I_i \in \{0,1\}$, where 0 implies low priority and 1 implies high priority. Additionally, an energy cost function associated with action a is defined as $f(a)$. Therefore, the problem is modeled as a Markov Decision Process (MDP), where the goal is to find the optimal policy π^* that maximizes the expected discounted cumulative reward:

$$\pi^* = \arg \max_{\pi} \mathbb{E}[\sum_{t=0}^T \gamma^t R(s_t, a_t)] \quad (7)$$

where:

- $\gamma \in (0, 1)$ is the discount factor that prioritizes intermediate rewards over long-term gains.
- $R(s_t, a_t)$ captures the trade-off between energy efficiency and successful transmission.

1) State, Action, and Reward Design

a) State Representation (s_t)

The state is encoded as a 51-dimensional feature vector:

$$s_t = [\delta_i, E_i(t), F \in \mathbb{R}^{48}, I_i] \quad (8)$$

The state space includes the residual energy $E_i(t)$ of device d_i at time t , the normalized distance to LoRa gateway δ_i , the frequency of past usage for spreading factor–channel combination $F \in \mathbb{R}^{48}$ (used for collision probability estimation), and a binary data importance flag I_i .

b) Action Space (a_t)

The action space is defined as:

$$\mathcal{A} = \{(SF, CH) | SF \in [7,12], CH \in [0,7]\} \cup \{\text{"wait"}\} \quad (9)$$

It consists of 48 transmission options and 1 idle (wait) action.

c) Reward Function (R)

The reward function balances successful transmission, energy cost, and collision penalty as follows:

$$R(s_t, a_t) = \begin{cases} +1.0 \cdot I_i - 0.05 \cdot (SF - 7), & \text{if successful Tx and } E > 0 \\ -0.01, & \text{if idle (wait)} \\ -1.0, & \text{if collision or } E \leq 0 \end{cases} \quad (10)$$

This formulation:

- Rewards high-priority data delivery.
- Penalizes higher spreading factors due to increased airtime and energy cost.
- Penalizes idle action with a small penalty.
- Discourages collisions and transmissions when energy is insufficient.

2) Deep Double Q-Network Formulation

In the proposed framework the optimal action-value function $Q^*(s, a)$ is approximated using two different neural networks: an online network $Q_{\theta}(s, a)$, updated at every time step, and a target network $Q_{\theta^-}(s, a)$, updated every c steps to minimize overestimation bias.

The DDQN Bellman target is formulated as:

$$y = r + \gamma Q_{\theta^-}(s', \arg \max_{a'} Q_{\theta}(s', a')) \quad (11)$$

The loss function minimized during training stage is:

$$L(\theta) = \frac{1}{B} \sum_{(s,a,r,s') \in \mathcal{B}} [Q_{\theta}(s, a) - y]^2 \quad (12)$$

Here, \mathcal{B} represents a mini-batch sampled from the experience replay buffer of size B .

3) Learning Process Formulation

The DDQN-based energy-aware transmission scheduling follows an ϵ -greedy exploration policy. With probability ϵ , the agent selects a random action under exploration phase, whereas with probability $1 - \epsilon$, it selects the greedy action:

$$a_t = \arg \max_a Q_{\theta}(s_t, a) \quad (13)$$

The exploration rate decays as follows:

$$\epsilon \leftarrow \max(\epsilon \cdot \epsilon_{decay}, \epsilon_{min}) \quad (14)$$

Algorithm 1 describes the DDQN-based energy-aware transmission scheduling for LoRaWAN. The algorithm initializes the networks and the replay buffer, constructs states for each device, selects actions via ϵ -greedy policy, updates energy and rewards, and also trains the network using sampled mini-batches. The target network is periodically updated until convergence to an optimal policy.

Algorithm 1: Energy-Aware Transmission Scheduling in LoRaWAN (DDQN)

1. **Initialize:** Replay buffer D , online network Q_θ , and target network Q_{θ^-}
2. **For each episode:**
 - a. Reset the environment and initialize the states of end devices d_i
 - b. **For each d_i at each step:**
 - Construct state vector s_t
 - Select action a_t via ϵ -greedy exploration policy
 - Simulate transmission outcome (success, collision, or idle)
 - Update residual energy $E_i(t)$
 - Compute reward $R(s_t, a_t)$
 - Store (s_t, a_t, r_t, s_{t+1}) in buffer D
 - Sample mini batch from D and update Q_θ via gradient descent
 - In every c steps, update target network $Q_{\theta^-} \leftarrow Q_\theta$
3. **Repeat:** until convergence

It should be noted that the agent learns the optimal policy $\pi^*(S_t)$ to select the best action $a_t \in \mathcal{A}$ at each step. The DDQN update rule is defined as follows:

$$Q_\theta(s_t, a_t) \leftarrow Q_\theta(s_t, a_t) + \alpha \cdot [r_t + \gamma \cdot Q_{\theta^-}(s_{t+1}, \arg \max_{a'} Q_\theta(s_{t+1}, a')) - Q_\theta(s_t, a_t)] \quad (15)$$

Here, θ represents online network weights, θ^- denotes the weights of the target network, which is synchronized periodically (every 10 episodes). The parameter α represents the learning rate, and γ is the discount factor. Furthermore, r_t denotes the reward and s_t, s_{t+1} represent the current and next states, respectively.

C. Proposed Hybrid Policy for Energy-Efficient Transmission in LoRaWAN

Building on the DDQN-based energy-aware transmission scheduling framework for LoRaWAN in smart city environments, this study further enhances performance by improving transmission decisions under varying network congestion levels. The proposed approach aims to optimize energy efficiency while maintaining a high PDR in dynamic and dense network conditions.

1) State Space (s_t)

The proposed work models the state space for DDQN as follows:

$$s_t = [b_t, q_t, l_t, c_t] \quad (16)$$

Here, b_t denotes the battery level of the end device, q_t represents the normalized queue length (ranging from 0 to 1), l_t indicates the success of the previous transmission, and c_t represents the current network congestion level.

2) Action Space (\mathcal{A})

The action space is defined to capture dynamic traffic scenarios in LoRaWAN:

$$\mathcal{A} = \{a_0, a_1, a_2, a_3\} \quad (17)$$

Here, a_0 denotes transmission on channel 0, a_1 denotes transmission on channel 1, a_2 represents adaptive channel selection based on congestion, and a_3 refers to "wait" or defer transmission. It should be noted that, in this hybrid framework combining DDQN and VQCs, the spreading factor is not fixed. Instead, it is adaptively selected by the VQCs, enabling more flexible and efficient policy adaptation under dynamic network conditions.

3) Hybrid Policy: Double Deep Q-Network and Variational Quantum Circuits

For a given transmission (s_t, a_t, r_t, s_{t+1}) , is computed as described in (11) to obtain the target value y_t . The corresponding loss function is defined as:

$$L(\theta) = \mathbb{E}[(Q_\theta(s_t, a_t) - y_t)^2] \quad (18)$$

In the proposed hybrid policy, a VQC is modeled is utilized for spreading factor selection. The VQC takes a 2D encoded quantum state as input, which is represented as:

$$|\psi(s_t)\rangle = U_{enc}(b_t, c_t)|00\rangle \quad (19)$$

Here, the battery level b_t and congestion level c_t are encoded via R_Y rotations. The variational circuit consists of trainable quantum gates, expressed as:

$$U(\theta_q) = \prod_{l=1}^L R_Y(\theta_{q,l}), R_Z(\theta_{q,l}) \text{ CNOT} \quad (20)$$

Finally, measurement of the quantum state yields the spreading factor decision:

$$SF_t = f_{VQC}(|\psi(s_t)\rangle) \in \{7,9,12\} \quad (21)$$

The reward function is defined as:

$$R_t = \alpha \cdot success_t - \beta \cdot E_t - \gamma \cdot D_t \quad (22)$$

Here, E_t represents the energy consumed, D_t represents the transmission delay, α, β, γ are the design weights, and $success_t$ refers to the packet successfully delivered (1) or lost (0).

4) Training Procedure in Hybrid Policy

During the training procedure, the framework initializes the replay buffer, the policy network Q_θ , the target network Q_{θ^-} , and the VQC weights θ_q . For each episode, the process initializes s_0 . At each time step t , the agent selects an action a_t using the ϵ -greedy exploration policy derived from Q_θ . Further, it uses VQC to select the optimal SF_t . It also interacts with the environment to obtain (s_{t+1}, r_t) , and stores the transition in the replay buffer D . From a sampled mini-batch \mathcal{B} , the loss $L(\theta)$ is

computed and used to update Q_θ . Finally, the target network Q_{θ^-} is updated in every K steps.

It has been observed that Algorithm 1 learns channel selection and the wait action but relies on fixed spreading factors, which limits adaptability under congestion. In contrast, the proposed hybrid strategy employs a quantum-enhanced VQC for spreading factor selection, enabling improved energy efficiency and higher PDR. The following analytical algorithm describes the enhanced hybrid policy integrating DDQN with VQC for efficient transmission and congestion control in LoRaWAN, aiming to balance the trade-off between energy consumption and PDR.

Algorithm 2 presents the hybrid DDQN+VQC-based policy for energy-aware resource optimization in LoRaWAN. It initializes classical and quantum components, constructs state vectors, and selects the actions using ϵ -greedy exploration. In this framework, the VQC layer refines the spreading factor selection, whereas the DDQN module updates Q-values using replayed experiences. Both components are trained iteratively until the hybrid policy converges to an optimized scheduling strategy. The analytical algorithm for hybrid policy-based LoRaWAN resource optimization is as follows.

Algorithm 2: Hybrid Policy-based LoRaWAN Energy-Aware Resource Optimization

Input: Episodes E , steps t , replay buffer D , VQC parameters θ_q

Output: Optimized policy $\pi^*(s)$

1. **Initialize** Q-network Q_θ , target network Q_{θ^-} , VQC circuit, buffer D

2. **For** episode 1 to E do:

Initialize states $s_t = [b_t, q_t, l_t, c_t]$

For $t = 1$ to T do

$\epsilon \leftarrow \text{epsilon_decay}(\text{global_step})$

If $\text{rand}() < \epsilon$ then

$a_t \leftarrow \text{random action}$

Else

$a_t = \arg \max_a Q_\theta(s_t, a)$

$SF_t \leftarrow VQC(b_t, c_t)$

$SF_{t+1}, R_t \leftarrow \text{step_env}(s_t, a_t, SF_t)$

Store transition in replay buffer D

If $|B| > B_{min}$

Sample mini-batch from replay buffer

Compute target:

$y_t = r_t + \gamma Q_{\theta^-}(s', \arg \max_{a'} Q_\theta(s', a'))$

Update: $L(\theta) = \mathbb{E}[(Q_\theta(s_t, a_t) - y_t)^2]$

If $\text{step \% target_sync} == 0$ then

$Q_{\theta^-} \leftarrow Q_\theta$

End for

3. **End for**

Figure 1 illustrates the comparison between the DDQN-based policy and the proposed hybrid DDQN-VQC policy. It has to be noted that the first algorithm of DDQN determines the learned policy through which it reduces unnecessary transmissions when energy is low or the data are of low-

priority. It also helps selecting spreading factor-channel combinations to minimize the collision probability. Algorithm 1 also accomplishes high PDR while prolonging device lifetime. On the other hand, the hybrid approach of DDQN and VQC integrates quantum-enhanced VQC for spreading factor selection, which also achieves energy efficiency and higher PDR.

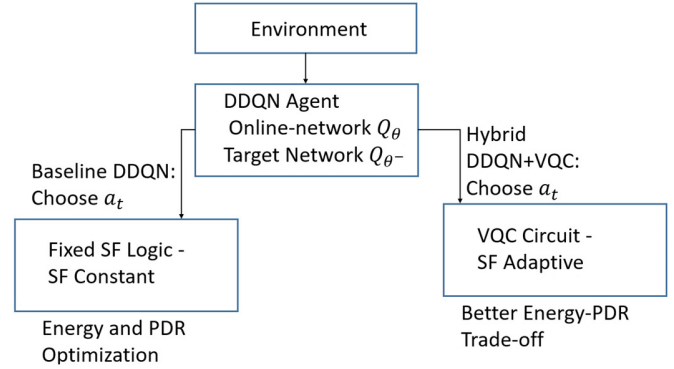


Fig. 1. Comparison of the DDQN-based policy and the proposed hybrid policy.

III. RESULTS AND DISCUSSION

A. Simulation Setup

The proposed energy-aware transmission scheduling framework for LoRaWAN was implemented in Python 3.10 using PyTorch for model training and Matplotlib for visualization. PennyLane was further used to simulate the quantum neural network (VQC) on a classical CPU. Since no large-scale real-world LoRaWAN dataset with diverse traffic and energy profiles is publicly available, the proposed framework is evaluated using a realistic simulation model. Here, the dataset is synthetic, generated for dense LoRaWAN smart city networks. Device positions, traffic patterns, and energy profiles are modeled using a Poisson traffic model ($\lambda = 2$ packets/min), with varying spreading factors, channels, and transmission power levels. Although real-world deployment data are not used, the performance outcomes are validated through comparative analysis with existing DDQN and DQN-based methods reported in the literature, ensuring the reliability of the simulation findings. However, the proposed solution (hybrid DDQN+VQC) can be implemented in real LoRaWAN networks in future experimental studies.

The simulations were conducted on a Windows system with an Intel Core i5-8250U CPU @ 1.60 GHz (up to 1.80 GHz) and 12 GB RAM with an x64-based processor. The simulation parameters are shown in Table I.

B. Performance Analysis for Double Deep Q-Network-Based Energy-Aware Transmission Scheduling

The analysis of PDR with an increasing number of episodes (0–300) shows initial instability from episodes 0–50. As shown in Figure 2, the DDQN exhibits significant fluctuations in PDR, ranging between 0.12 and 0.32, indicating that the agent is in the exploration phase and is still learning optimal transmission strategies. High variance occurs because the agent

explores different combinations of spreading factors, transmission power, and channel usage. However, the PDR gradually increases and stabilizes around 0.25 to 0.35. The model begins to learn energy-aware transmission strategies that avoid redundant retransmissions, minimize collisions, and improve successful packet delivery. During the convergence phase, the PDR shows stable oscillations around 0.30 ± 0.05 . This demonstrates that the DDQN-based LoRa DRL agent has converged to an optimal policy that balances energy efficiency (by avoiding unnecessary transmissions), network reliability (by maximizing packet delivery success), and scalability (by adapting to increasing network density). It clearly shows that the proposed DDQN-based model improves network reliability.

TABLE I. SIMULATION PARAMETERS

| Parameter | Value |
|--------------------------------------|--------------------------------------|
| Number of end devices | 100-500 |
| Simulation area | 2 km × 2 km |
| Gateway count | 1 |
| Transmission power (TP) | 2 dBm – 14 dBm |
| Spreading factors (SF) | 7, 8, 9, 10, 11, 12 |
| Bandwidth | 125 kHz |
| Number of Episodes | 300-500 |
| Frequency band | 868 MHz |
| State size | 51 |
| Discount Factor (γ) | 0.7-0.95 |
| Learning rate (α) | 0.0005 |
| Epsilon decay (ϵ_{decay}) | 0.995 |
| Minimum epsilon (ϵ_{min}) | 0.05 |
| Target update | 10 |
| Memory size | 5,000 |
| Batch size | 64-128 |
| Initial energy | 1.0 J |
| Min_buffer | 2,000 |
| Target_sync | 500 |
| Traffic model | Poisson ($\lambda = 2$ packets/min) |

between 0 and +15 with fewer sharp drops. This indicates that the agent has learned an effective strategy for maximizing performance, leading to improved energy efficiency and packet delivery. The outcome of the DDQN-based adaptive transmission strategy demonstrates successful DRL training, transitioning from unstable and negative rewards in early episodes to more stable and positive rewards in later episodes.

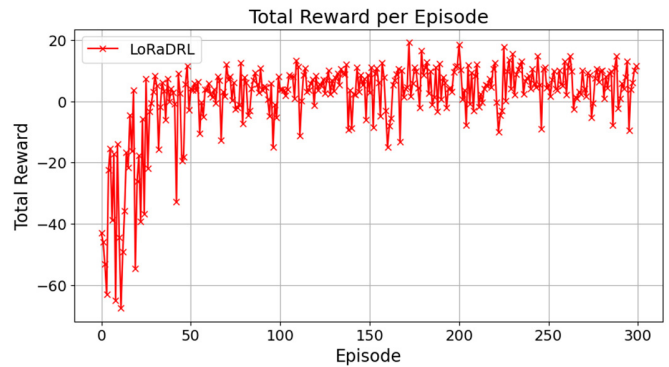


Fig. 3. Total reward computation vs episodes in DDQN-based policy.

Figure 4 shows how the average remaining energy of IoT devices evolves over 300 training episodes using the DDQN-based LoRa DRL scheduling and resource allocation strategy. In the initial phase (episodes 0-50), the average remaining energy is slightly lower (~0.987) because the algorithm is still exploring various transmission configurations (spreading factor, transmission power, and bandwidth). This also leads to higher fluctuations due to exploration-driven randomness. During the learning phase (episodes 50-150), the energy consumption stabilizes, maintaining an average remaining energy between 0.990 and 0.994. Here, the DDQN agent begins to learn an energy-efficient transmission policy for selecting spreading factors and transmission powers that balance energy usage with PDR.

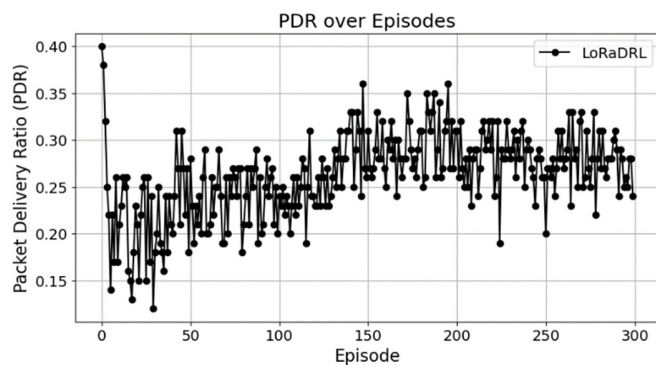


Fig. 2. Analysis of PDR vs episodes in the DDQN-based policy.

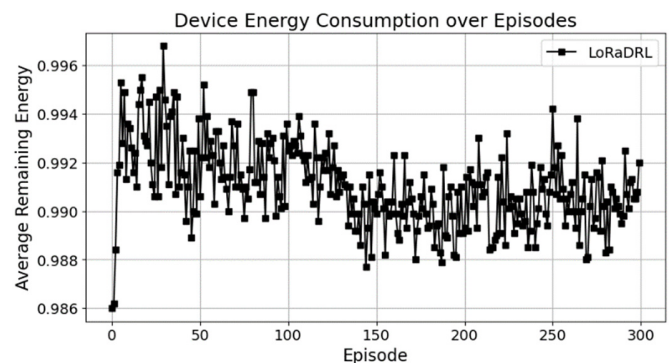


Fig. 4. Average remaining energy vs episodes in the DDQN-based policy.

Figure 3 shows that the DDQN model in the proposed LoRaWAN framework progressively achieves higher rewards. In the initial phase, the total reward is highly negative, fluctuating between -80 and -20, indicating the exploration stage of the agent. During the learning phase, there is a steady upward trend in rewards, moving from negative values toward approximately 0. Here, the agent begins to learn better policies for selecting spreading factors, transmission power, and channels. In the convergence phase, the total reward stabilizes

Finally, during the convergence phase, the DDQN curve becomes relatively stable with minor oscillations around 0.990. This indicates that the DDQN-based agent has successfully optimized energy efficiency, and no further significant improvement occurs. The minor oscillations reflect adaptive policy adjustments due to dynamic network conditions,

although the variations remain minimal. After approximately 150 episodes, the system converges, maintaining an average remaining energy above 98.8%, which highlights the energy-aware nature of the proposed DDQN-based transmission scheduling strategy.

Figure 5 presents the action frequency heatmap, which illustrates the distribution of selected spreading factors and channels. The darker regions correspond to the frequently chosen spreading factor–channel combinations, indicating the agent's preference for configurations that optimize energy efficiency and minimize collisions. The agent predominately selects lower spreading factors (SF0–SF2) with balanced channel usage, particularly favoring channels 0, 1, 4 and 7. This behavior demonstrates the agent's ability to learn an adaptive allocation strategy that improves PDR while maintaining energy efficiency in LoRaWAN.

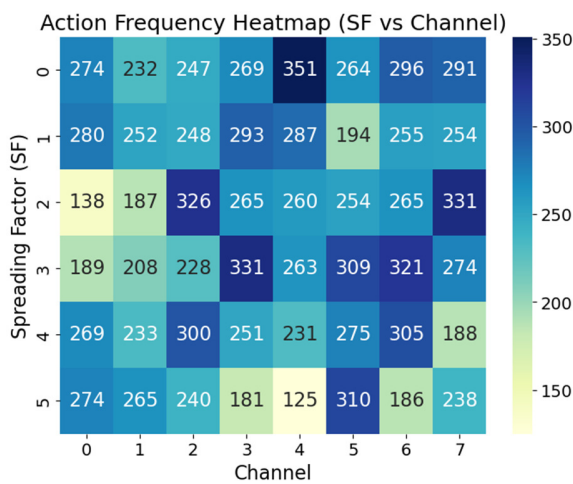


Fig. 5. Action frequency heatmap for the DDQN-based LoRa scheduling framework.

Further the proposed work provides a comparative performance analysis between the hybrid (DDQN+VQC) policy and the classical DQN ($SF = 9$), as shown in Figure 6. The analysis of Figure 6(a) shows that during the initial episodes (0–10), the classical DQN exhibits unstable PDR with significant fluctuations, sometimes dropping close to zero. The hybrid DDQN+VQC method achieves more stable learning early on, showing less volatility. From approximately episode 20 onwards, the hybrid model consistently maintains a slightly higher and more stable PDR (~0.35–0.5) compared to the classical DQN (~0.3–0.45). This demonstrates that integrating VQCs with DDQN improves reliability and stability in packet delivery performance over time.

In Figure 6(b), it is observed that the classical DQN shows high fluctuations in energy consumption, with occasional spikes above 0.10 J per step. In contrast, the hybrid DDQN+VQC model stabilizes earlier and avoids extreme peaks, indicating better exploration and more balanced action selection. In the mid-phase, the hybrid model also achieves more efficient energy usage on average (~0.06 J per step), whereas the classical DQN averages slightly higher (~0.07 J per step).

Lower variance in the hybrid approach suggests better convergence and more consistent decision-making. Although both models stabilize, the hybrid model still shows higher peaks in a few regions due to adaptive dynamic spreading factor selection. However, on average, the hybrid model consumes approximately 8–12% less energy compared to the classical DQN.

The graph in Figure 6(c) compares the average reward per step across episodes for both approaches. During the learning phase, the classical DQN initially shows slightly higher rewards but with high fluctuations. The hybrid model starts with lower rewards but stabilizes quickly after approximately 15 episodes. In the mid-phase, the hybrid DDQN+VQC model catches up and maintains more consistent rewards, whereas fluctuations in the classical DQN remain higher, indicating unstable policy updates. This suggests that the hybrid approach achieves a better balance between exploration and exploitation. In the later phase, both models converge toward similar average rewards (~0.3–0.35 per step). However, the hybrid model exhibits lower variance, producing more reliable performance across episodes. Therefore, the experimental results clearly validate the effectiveness of the proposed DDQN+VQC approach.

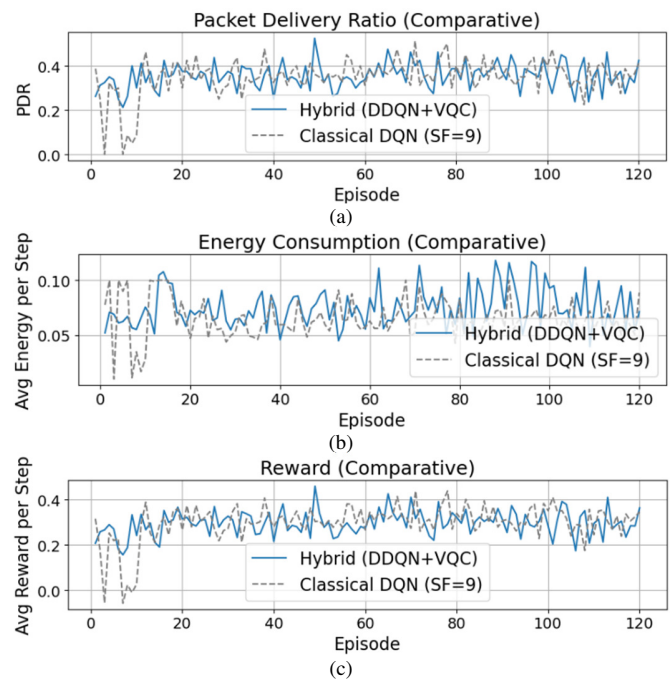


Fig. 6. Comparative analysis of hybrid (DDQN+VQC) vs classical DQN: (a) PDR, (b) energy consumption, (c) reward.

It should be noted that due to the lack of publicly available large-scale LoRaWAN datasets, a full statistical comparison with existing methods could not be performed in this study. The current evaluation is based on controlled simulation scenarios to demonstrate the relative performance gains of the proposed DDQN+VQC framework. Comprehensive statistical benchmarking on real-world datasets will be considered in future work when such datasets become available.

IV. CONCLUSION

In this paper, a hybrid Deep Reinforcement Learning (DRL) framework that combines Double Deep Q-Learning (DDQN) and Variational Quantum Circuit (VQC) is introduced to optimize resource allocation, scheduling, and energy efficiency in Long-Range Wide-Area Network (LoRaWAN)-based Internet-of-Things (IoT) networks. The hybrid approach leverages the robust policy stability of DDQN and the quantum-inspired feature representation capability of VQC, which allows better adaptation to the highly dynamic characteristics of large-scale IoT environments.

The extensive simulation results demonstrate that both the standalone DDQN and the hybrid DDQN+VQC policies achieve consistently better performance, with higher average rewards and lower variance compared to the classical Deep Q-Network (DQN) baseline, as reflected in the reward versus episodes analysis. The hybrid approach learns stable and energy-aware scheduling policies more efficiently, resulting in improved Packet Delivery Ratio (PDR), enhanced energy utilization, and reduced latency across heterogeneous device configurations. Furthermore, the comparative analysis shows that although the classical DQN reaches competitive peak performance, its instability under dynamic network conditions results in policy fluctuations, whereas the proposed hybrid model maintains reliable convergence and ensures consistent Quality of Service (QoS) guarantees.

Future work will focus on incorporating multi-agent DRL to enable cooperative scheduling among multiple gateways and integrating quantum meta-learning for faster policy adaptation in non-stationary environments. The study can also be extended to include real-world datasets and large-scale testbed evaluations to further validate the simulation-based findings.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING SOURCES

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ACKNOWLEDGMENT

Not applicable to this work.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author.

AI USE AND DECLARATION OF GENERATIVE AI USE

During the preparation of this work, the author used ChatGPT in order to look for fundamental theories related to the study. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

- [1] "Ericsson Mobility Report." Ericsson. <https://www.ericsson.com/en/reports-and-papers/mobility-report>.
- [2] F. Cuomo, M. Campo, A. Caponi, G. Bianchi, G. Rossini, and P. Pisani, "EXPLoRa: Extending the performance of LoRa by suitable spreading factor allocations," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications*, Rome, Italy, 2017, pp. 1–8, <https://doi.org/10.1109/WiMOB.2017.8115779>.
- [3] A. Khalifeh, K. A. Aldahdouh, K. A. Darabkh, and W. Al-Sit, "A Survey of 5G Emerging Wireless Technologies Featuring LoRaWAN, Sigfox, NB-IoT and LTE-M," in *2019 International Conference on Wireless Communications Signal Processing and Networking*, Chennai, India, 2019, pp. 561–566, <https://doi.org/10.1109/WiSPNET45539.2019.9032817>.
- [4] M. A. Gava, H. R. O. Rocha, M. J. Faber, M. E. V. Segatto, H. Wörtche, and J. A. L. Silva, "Optimizing Resources and Increasing the Coverage of Internet-of-Things (IoT) Networks: An Approach Based on LoRaWAN," *Sensors*, vol. 23, no. 3, Jan. 2023, Art. no. 1239, <https://doi.org/10.3390/s23031239>.
- [5] K. Lin, M. A. Ullah, H. Alves, K. Mikhaylov, and T. Hao, "Energy Efficiency Optimization for Subterranean LoRaWAN Using a Reinforcement Learning Approach: A Direct-to-Satellite Scenario," *IEEE Wireless Communications Letters*, vol. 13, no. 2, pp. 308–312, Feb. 2024, <https://doi.org/10.1109/LWC.2023.3327833>.
- [6] V. Balyan, "Priority-based resource allocation in layering LoRaWAN," *International Journal on Smart Sensing and Intelligent Systems*, vol. 17, no. 1, Nov. 2024, Art. no. 34, <https://doi.org/10.2478/ijssis-2024-0034>.
- [7] G. Park, W. Lee, and I. Joe, "Network resource optimization with reinforcement learning for low power wide area networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, Sept. 2020, Art. no. 176, <https://doi.org/10.1186/s13638-020-01783-5>.
- [8] E. I. Shayo, A. T. Abdalla, and A. J. Mwambela, "Dynamic multi-frame multi-spreading factor scheduling algorithm for LoRaWAN," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, Feb. 2023, Art. no. 11, <https://doi.org/10.1186/s43067-023-00077-2>.
- [9] P. Maurya, A. Hazra, P. Kumari, T. B. Sørensen, and S. K. Das, "A Comprehensive Survey of Data-Driven Solutions for LoRaWAN: Challenges and Future Directions," *ACM Transactions on Internet of Things*, vol. 6, no. 1, Feb. 2025, Art. no. 6, <https://doi.org/10.1145/3711953>.
- [10] M. Chen, L. Mokdad, J. Ben-Othman, and J.-M. Fourneau, "Dynamic Parameter Allocation With Reinforcement Learning for LoRaWAN," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10250–10265, June 2023, <https://doi.org/10.1109/IJOT.2023.3239301>.
- [11] S. Madhavan, R. Riad Al-Fatlawy, MD. Mujahid Irfan, K. Neeraj, and R. Dineshkumar, "Improving LoRa Low-Power Wide-Area Networks with Double Deep-Q Learning-Based Resource Allocation," in *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics*, Ballari, India, 2024, pp. 1–5, <https://doi.org/10.1109/ICDCECE60827.2024.10549236>.
- [12] A. Núñez and E. María, "Parametrized Quantum Circuits for Reinforcement Learning," Ph.D. dissertation, Universidad de Granada, Granada, Spain, 2024.
- [13] M. Cerezo *et al.*, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, no. 9, pp. 625–644, Sept. 2021, <https://doi.org/10.1038/s42254-021-00348-9>.
- [14] Z. Yan, R. Tanikella, and H. Tabassum, "Optimizing Vehicular Networks with Variational Quantum Circuits-Based Reinforcement Learning," in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops*, Vancouver, Canada, 2024, pp. 1–2, <https://doi.org/10.1109/INFOCOMWKSHP61880.2024.10620888>.
- [15] K. Zhang, C. K. M. Lee, Y. P. Tsang, and C. H. Wu, "Variational Quantum Reinforcement Learning for Joint Resource Allocation of Blockchain-Based Vehicular Edge Computing and Quantum Internet," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 10, pp. 15831–15847, Oct. 2025, <https://doi.org/10.1109/TVT.2025.3568158>.

- [16] S. Gao, Y. Zou, and L. Feng, "A Lightweight Double-Deep Q-Network for Energy Efficiency Optimization of Industrial IoT Devices in Thermal Power Plants," *Electronics*, vol. 14, no. 13, June 2025, Art. no. 2569, <https://doi.org/10.3390/electronics14132569>.
- [17] S. Parween and S. Z. Hussain, "Efficient collision control and auction-based resource allocation mechanism in dense LoRaWAN network via TCP using DRL technique," *International Journal of Information Technology*, vol. 16, no. 7, pp. 4039–4057, Oct. 2024, <https://doi.org/10.1007/s41870-024-01986-9>.
- [18] A. Tellache, A. Mekrache, A. Bradai, R. Boussaha, and Y. Pousset, "Deep Reinforcement Learning based Resource Allocation in Dense Sliced LoRaWAN Networks," in *2022 IEEE International Conference on Consumer Electronics*, Las Vegas, NV, USA, 2022, pp. 1–6, <https://doi.org/10.1109/ICCE53296.2022.9730234>.
- [19] G. Zhao, K. Lin, D. Chapman, N. Metje, and T. Hao, "Optimizing energy efficiency of LoRaWAN-based wireless underground sensor networks: A multi-agent reinforcement learning approach," *Internet of Things*, vol. 22, July 2023, Art. no. 100776, <https://doi.org/10.1016/j.iot.2023.100776>.
- [20] A. Li, "Deep Reinforcement Learning Based Resource Allocation for LoRaWAN," in *2022 IEEE 96th Vehicular Technology Conference*, London, UK, 2022, pp. 1–4, <https://doi.org/10.1109/VTC2022-Fall57202.2022.10012698>.
- [21] J. Mhatre, M. Nguyen, A. Lee, and H. Lee, "Experimental Analysis of LoRaWAN for Optimizing Water Quality Monitoring with Reinforcement Learning-Driven Scheduling," in *2024 IEEE International Performance, Computing, and Communications Conference*, Orlando, FL, USA, 2024, pp. 1–6, <https://doi.org/10.1109/IPCCC59868.2024.10850023>.
- [22] J. Mhatre, A. Lee, and H. Lee, "Frequency Hopping Scheduling Algorithm in Green LoRaWAN: Reinforcement Learning Approach," in *2023 IEEE Conference on Standards for Communications and Networking*, Munich, Germany, 2023, pp. 216–221, <https://doi.org/10.1109/CSCN60443.2023.10453154>.
- [23] M. Catak and B. Celikkaya, "Limits of Number of Motes at Smart City Scenarios for LoRaWAN," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4116–4119, June 2019, <https://doi.org/10.48084/etasr.2693>.

AUTHOR PROFILES

Shaista Tarannum is an Assistant Professor at MSRUAS, Bengaluru, currently pursuing her Ph.D. in Electronics and Communication Engineering at JSSATE, VTU. Her research focuses on the optimization of LoRaWAN networks using Artificial Intelligence and Machine Learning for smart city applications. She holds an M.Tech. in Digital Electronics and Communication and a B.E. in Telecommunication Engineering. With teaching and mentoring expertise in IoT, AI, ML, and Computer Networks, she integrates cutting-edge technologies into both academic and applied research contexts. Her work emphasizes energy-efficient, secure, and adaptive LoRaWAN communication systems, contributing to sustainable urban infrastructure. She has presented at international conferences such as ICASF, ICEET, and ICISAS and published in peer-reviewed journals including Springer and AIP. Recognized for her innovative work, she received the Greenpeace MENA Award at COP28 and was a finalist in the Quantum Innovation Summit 2024. She also led the ISRO START program and conducted faculty training programs in AI-driven academic innovation. A certified AI/ML practitioner with 22 CEUs from Caltech, she is an active member of IEEE and ISTE, committed to advancing interdisciplinary research aligned with the Sustainable Development Goals. She can be contacted at: shaistatarannum123@gmail.com.

Usha Shiramally Mallappa from Bengaluru, Karnataka, India, obtained a B.E Electronics & Communication Engineering degree from Mysore University in 2000. She obtained an M.Tech. in VLSI Design and Embedded Systems from VTU Belgaum in 2011 and awarded a Ph.D. in Optimization and Performance Analysis of Digital Modulators from VTU Belgaum in 2017, and is guiding three research scholars from VTU. She is currently working as an Associate Professor in Electronics & Communication Engineering department at JSS Academy of Technical Education, Bengaluru, Karnataka, India. She is a member of professional bodies such as IEEE, ISTE, and MIE.