RL-based Energy-Efficient Data Transmission over Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network

W. A. Nelson, S. R. Yeduri, A. Jha, A. Kumar, L. R. Cenkeramaddi

Abstract: **The lifetime of a UAV-assisted wireless network is determined by the amount of energy consumed by the UAVs during flight, data collection, and transmission to the ground station. Routing protocols are commonly used for data transmission in a communication network. However, because of the mobility of UAVs, using a routing protocol with a single communication technology results in higher delay and more energy consumption in a UAV-assisted wireless network. To overcome this, we propose two reinforcement learning (RL) algorithms, Q-learning and deep Q-network (DQN), for energy-efficient data transmission over a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. We consider BLE, LTE, Wi-Fi, and LoRa for communication over a UAV-GS link. The RL algorithms take any random network as input and learn the best policy to output the network with less energy consumption. The reward/penalty is chosen in such a way that the network with the highest energy consumption is penalized and the one with the lowest is rewarded, thereby minimizing total network energy consumption. Based on learning, it creates a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network by assigning the best communication technology to a UAV-GS link. Further, we compare the performance of proposed RL algorithms with a rule-based algorithm and random hybrid scheme. In addition, we propose a theoretical framework for constructing hybrid network for both free space and free space multipath path loss models. We demonstrate the performance comparison of the proposed work with the conventional shortest path routing algorithm in terms of network energy consumption and average network delay using extensive results. Finally, the effect of the velocity of the UAV and the number of packets on the performance of the proposed framework is analyzed.**

G.1 Introduction

Technological advances and ease of regulations in today's world have helped unmanned aerial vehicle (UAV) technology to serve the ever-growing demand of modern applications. UAVs are used in a number of applications that include precision agriculture [1], space exploration [2], providing connectivity to terrestrial networks [3], and public safety [4]. Their ability to maneuver and access remote locations has made them suitable for remote sensing and disaster management applications [5]. UAV swarms with high precision control have recently paved the way for UAV light-shows in various events thereby reducing pollution due to explosives [6].

To ensure safe and secure UAV operations, UAV communication plays a vital role. UAVs are equipped with various communication modules such as Bluetooth Low Energy (BLE), Long Term Evolution (LTE), Wireless Fidelity (Wi-Fi), and Long Range (LoRa). These communication technologies ensure that UAVs are transmitting the required control and telemetry data to the ground control stations (GS). UAV health and trajectory are monitored by continuously sending signals to the GS and if required, to other UAVs.

Various works have been proposed in this regard. A dual radio internet-of-things (IoT) architecture has been proposed in [7] for the application of wildlife monitoring systems. The proposed approach in [7] leverages BLE in low power wide area network (LPWAN) based on the proximity among the wildlife animal herd. Finally, an analytical model for energy consumption has been presented to evaluate the performance of the proposed approach [7]. It has been shown that the proposed dual radio network improves energy efficiency when compared to a network utilizing LPWAN alone [7]. In [8], the authors have proposed an indoor hybrid RF/PLC/VLC communication system to switch the device connections among the RF, PLC, and VLC in order to improve the sum rate capacity. Further, the transmit power minimization problem has been formulated and analyzed in [9]. However, no literature is presented focusing on the hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network for energy-efficient data transmission.

Further, many works have been presented in the literature focusing on the application of reinforcement learning (RL) in UAV-assisted wireless networks. A deep learning (DL) algorithm has been proposed in [10] for the collection of data in a UAV-assisted wireless network. In [11], several DL-based artificial intelligence (AI) methods such as point learning, multi-agent deep deterministic policy gradient, and federated DL have been proposed to solve the optimization problem of energy efficiency in a UAV-assisted wireless network. In [12], a combination of echo state learning and RL is used to solve the joint flight control and spectrum access problem in TeraHertz-band for UAV-assisted wireless networks. A deep reinforcement learning (DRL) approach has been proposed in [13] for resource allocation in terms of throughput, bandwidth, and power consumption in UAV-assisted wireless networks. In [14], the authors utilize the stochastic learning automata (SLA) algorithm to perform joint optimization of channel and relay selection in UAV-aided device-to-device (D2D) networks. However, none of these works focused on utilizing RL al-

gorithms for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. Motivated by this, we propose two RL algorithms for energy-efficient data transmission over a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. The main contributions of this paper are:

- A hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network is formed with free space (FS) and free space multipath (FSMP) energy models.
- Analytical expressions corresponding to network energy consumption and average network delay are derived for the data transmission over the hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network.
- Further, two RL algorithms namely, Q-learning and deep Q-network (DQN) are proposed for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. The performance of the proposed RL algorithms is compared with the rule-based algorithm and random hybrid scheme.
- Through extensive numerical results, we show that the proposed RL algorithms result in energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. We also compare the performance of the proposed hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network in terms of the network energy consumption and average network delay with conventional shortest path routing algorithm considering individual communication technology.

The remaining sections of this article are structured as follows: Section G.2 discusses the system model and problem formulation. The proposed theoretical framework for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network is presented in Section G.3. The analytical expressions corresponding to network energy consumption and average network delay for FS and FSMP models are derived in Section G.4. Section G.5 describes the proposed RL algorithms for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. Evaluation metrics are described in Section G.6 and the numerical results are discussed in Section G.7. Finally, Section G.8 concludes the paper by providing a summary and possible future research directions.

G.2 System model and problem formulation

We consider a UAV-assisted wireless network wherein, N UAVs are deployed randomly over B^2 area over l layers of height $h_1, h_2, \cdots, h_l$, respectively, as shown in Fig. G.1. These UAVs collect the data and send it to the GS which is situated on the ground at $(B/3, B/3)$. We consider GS to have access to all communication technologies to collect the data and process it. Here, the data arrival at each UAV is assumed to follow the Poisson process. After a successful transmission of data from one location, the UAV moves to another location randomly to collect the data.

Table G.1: Summary of important parameters used in this work.

Notation	Definition
C	Cost of the network in terms of energy consumption
N	Number of UAVs present in the network
h_i	Height of the i^{th} layer
l	Total number of layers
B^2	Area of each layer over which UAVs are deployed
P_T	Transmit power
R	Data rate over the transmission link
d_g	Radial distance
k	Number of bits
$\mathcal{E}_e$	Energy consumed by the electronic circuit for transmitting one bit
$\mathcal{E}_{fs}$	Free space power amplification energy
$\mathcal{E}_{mp}$	Power amplification energy in the multipath fading model
E_{Tr}	Energy consumed for transmitting k bits of data
E_d	Energy consumed for transmitting k bits of data over d_g distance
τ	Communication technology
r_τ	Range of the communication technology
λ	Learning rate for the proposed RL algorithms
γ	Reward parameter for the proposed RL algorithms
Δ	Discount parameter for the proposed RL algorithms
T_{Tr}	Transmission delay
T_{prop}	Propagation delay
A_τ	Area occupied by the communication technology
P_{pr}	Probability that a UAV present A_τ
W_1	Weighing parameter corresponding to E_{Tr}
W_2	Weighing parameter corresponding to E_d
$E_{N/W}$	Total energy consumed by the network
w	Weights of the policy neural network
$\hat{w}$	Weights of the target neural network
ζ	Interval after which target network weights are updated (in terms of number of steps)
ϕ	Denotes if terminal state is reached

A UAV-GS link is considered to select one of the four communication technologies from BLE, LTE, Wi-Fi, and LoRa. Once a communication link is established to a UAV-GS link, the data will be transmitted with a transmit power of P_T and a data rate of R. This will incur a delay of T_d seconds. Further, two path loss models

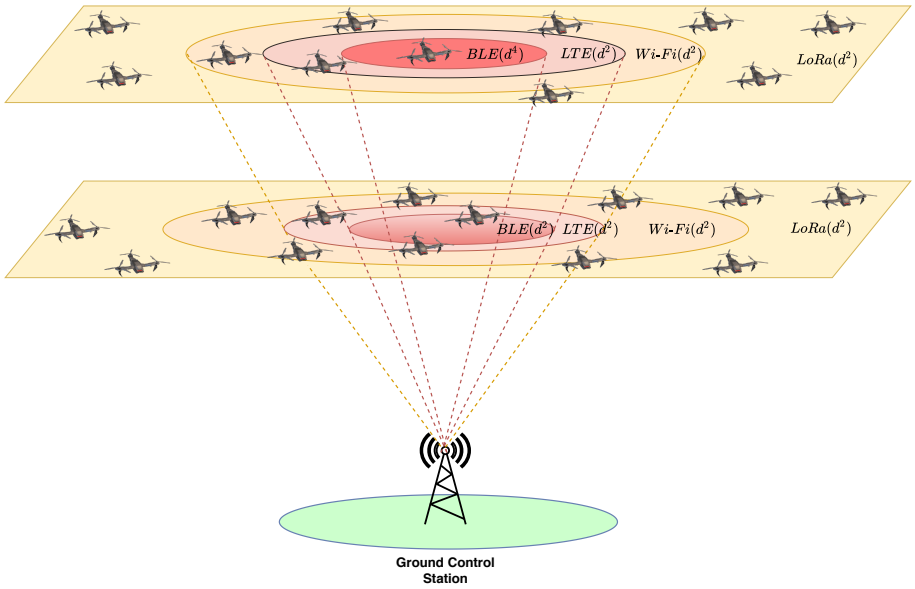


Figure G.1: System model for the hybrid BLE/LTE/Wi-Fi/LoRa scheme for two layers.

such as FS and FSMP are considered for inserting a communication technology for a UAV-GS link. In the FS model, a UAV-GS link can select one of the communication technologies, if the radial distance of the link is less than r_τ where τ represents one of the four communication technologies BLE, LTE, Wi-Fi, or LoRa. In the FS model, the energy consumption follows d^2 model. In the FSMP model, a UAV-GS link can select any of the communication links irrespective of its distance. However, the energy consumption follows d^2 model if the radial distance is less than r_τ, else, the energy consumption follows d^4 model.

G.2.1 Problem Formulation

In this work, our aim is to minimize the network energy consumption. Thus, we first obtain the overall network energy consumption. The total network energy consumption is obtained as the sum of the energy consumption at each UAV-GS link. The energy consumption at a UAV-GS link is the sum of the energy consumption for transmitting the data with a transmit power and the energy consumed for transmitting this data over a radial distance d_g. With P_T being the transmit power and R as the data rate, the energy consumed for transmitting k bits of data is obtained as [15]

$$E_{Tr} = \frac{kP_T}{R}. \tag{G.1}$$

As mentioned earlier, the energy consumption depends on the path loss model and the radial distance. Next, we derive the energy consumption for the FS model and FSMP model.

G.2.1.1 FS Model

In the FS model, a link can be connected to a communication technology in case the radial distance is less than the threshold distance, r_τ, of that communication technology. With d_g being the radial distance, the energy consumed for transmission of data over a distance d_g is obtained as [16]

$$E_{d,fs} = k\mathcal{E}_e + k\mathcal{E}_{fs}d_g^2, \tag{G.2}$$

where $\mathcal{E}_e$ is the energy consumed by the electronic circuit for transmitting one bit and $\mathcal{E}_{fs}$ denotes the free space power amplification energy.

G.2.1.2 FSMP Model

In the FSMP model, a UAV-GS link will be assigned with any of the communication technologies irrespective of its radial distance. The energy consumed for transmitting k bits of data over a radial distance of d_g is obtained as [16]

$$E_{d,fsmp} = \begin{cases} k\mathcal{E}_e + k\mathcal{E}_{fs}d_g^2, & \text{when } d_g < r_\tau\,; \\ k\mathcal{E}_e + k\mathcal{E}_{mp}d_g^4, & \text{when } d_g \geq r_\tau\,. \end{cases} \tag{G.3}$$

Here, $\mathcal{E}_{mp} = \mathcal{E}_{fs}/r_\tau^2$ denotes the power amplification energy in the multipath fading model. We consider $\mathcal{E}_e = 25 \times 10^{-9}$ J/bit and $\mathcal{E}_{fs} = 10 \times 10^{-12}$ J/bit/m^2 for the simulation evaluation [17].

Finally, the total cost of the UAV-assisted wireless network in terms of energy consumption is obtained as

$$C = \sum_{i=1}^{N}(E_{Tr,i} + E_{d,i}). \tag{G.4}$$

The aim of this work is to minimize the cost, defined in (G.4), of the network in terms of energy consumption.

G.3 Energy-Efficient Data Transmission over Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network Formation

The proposed scheme aims at the creation of a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network for energy-efficient data transmission. We consider two path loss models for generating the network which are described below.

G.3.1 FS model

In the FS model, a UAV-GS link can select a communication technology in case the radial distance is less than the threshold distance r_τ.

Algorithm 6: Algorithm to obtain the FS model-based Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network.

Input: N, d_g

Output: Network energy consumption and average network delay

if $d_g < r_{BLE}$ **then**
  Calculate the energy consumed over a UAV-GS link using (G.1) and (G.2) for all communication protocols;
  $E_{\min} = \min\{E_{BLE}, E_{LTE}, E_{Wi\text{-}Fi}, E_{LoRa}\}$;
  Choose the communication protocol with $E_{\min}$;
else if $d_g < r_{LTE}$ **then**
  Calculate the energy consumed over a UAV-GS link using (G.1) and (G.2) for LTE, Wi-Fi, and LoRa;
  $E_{\min} = \min\{E_{LTE}, E_{Wi\text{-}Fi}, E_{LoRa}\}$;
  Choose the communication protocol with $E_{\min}$;
else if $d_g < r_{Wi\text{-}Fi}$ **then**
  Calculate the energy consumed over a UAV-GS link using (G.1) and (G.2) for Wi-Fi and LoRa;
  $E_{\min} = \min\{E_{Wi\text{-}Fi}, E_{LoRa}\}$;
  Choose the communication protocol with $E_{\min}$;
else
  Choose the LoRa communication protocol;
end
Obtain network energy consumption and average network delay

Table G.2: Transmit power, data rate, and the path loss reference distance for all the communication protocols considered in this work.

Protocol	$\boldsymbol{P_T}$ **(W)**	$\boldsymbol{R}$ **(Mbps)**	$\boldsymbol{P_T/R}$	$\boldsymbol{r_\tau}$
BLE [18], [19]	0.01	1.36	0.007×10^{-6}	200
LTE [20], [21]	0.1	1	0.1×10^{-6}	400
Wi-Fi [22], [23]	2	10	0.2×10^{-6}	600
LoRa [23], [24], [25]	0.025	0.050	0.5×10^{-6}	1500

Algorithm 6 describes the steps involved in the selection of a communication protocol for a UAV-GS link. From Table G.2, it is observed that the threshold distance of BLE is less than that of LTE. This will be followed by Wi-Fi and LoRa. When the radial distance of a link is less than r_{BLE}, the link will be assigned to one of the four communication technologies that consume less energy. When the radial distance is less than r_{LTE}, the link will be assigned to one of the communication technologies from LTE, Wi-Fi, and LoRa based on minimum energy consumption. When the radial distance is less than $r_{Wi\text{-}Fi}$, the link will be assigned with one of the communication technologies from Wi-Fi and LoRa based on minimum energy consumption. Else, it will be connected to LoRa.

Algorithm 7: Algorithm to obtain the FSMP model-based Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network.

Input: N, d_g

Output: Network energy consumption and average network delay

1 Calculate the energy consumed over a UAV-GS link using (G.1) and (G.3) for all communication protocols

2 $E_{min} = \min\{E_{BLE}, E_{LTE}, E_{Wi\text{-}Fi}, E_{LoRa}\}$

3 Choose the communication protocol with $E_{\min}$

4 Obtain network energy consumption and average network delay

G.3.2 FSMP model

In the FSMP model, a link will be assigned with one of the communication links based on the energy consumption irrespective of its radial distance.

Algorithm 7 describes the procedure for the creation of hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network to evaluate the energy consumed over a link for different communication technologies. Assign the communication technology that consumes less energy for transmission.

G.4 Analytical model for Energy-Efficient Data Transmission over Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network

In this section, we derive the mathematical expressions for the network energy consumption for both FS and FSMP models.

G.4.1 FS Model

In the FS model, a link can be assigned to a communication technology in case the radial distance is less than the threshold distance. This implies that the energy consumption follows d^2 model. Thus, the connection to a communication protocol completely depends on the E_{Tr} as the E_d is the same for all the communication protocols. From Table G.2, it can be observed that the P_T/R ratio is less for BLE, followed by LTE, Wi-Fi, and LoRa. Thus, the number of UAV-GS links connected to any of the communication technologies depends on the presence of the UAV in its communication range. For example, if a UAV is present in the BLE range, it will be connected to BLE. In case the UAV is present in the LTE range and not in the BLE range, then it will be connected to LTE. The same applies for Wi-Fi. In case the UAV is not present in the Wi-Fi range, it will be connected to LoRa. Thus, each communication technology occupies a circular range for its connection. Thus, the circular area over which each communication technology occupies in a layer of

height h_i is obtained as

$$A_\tau = \begin{cases} \pi x_{BLE,fs}^2, & \text{for BLE} \\ \pi x_{LTE,fs}^2 - \pi r_{BLE,fs}^2, & \text{for LTE} \\ \pi x_{Wi\text{-}Fi,fs}^2 - \pi r_{LTE,fs}^2, & \text{for Wi-Fi} \\ B^2 - \pi x_{Wi\text{-}Fi,fs}^2, & \text{for LoRa} \end{cases} \tag{G.5}$$

where, τ represents the different communication technologies and x_τ is given as

$$x_\tau = \begin{cases} \sqrt{r_\tau^2 - h_i^2}, & \text{if } r_\tau > h_i; \\ 0, & \text{otherwise,} \end{cases} \tag{G.6}$$

Let P_{pr} represent the probability that a UAV is present in the area corresponding to a communication technology which is obtained as

$$P_{pr,\tau} = \frac{A_\tau}{B^2}. \tag{G.7}$$

Thus, the number of UAV-GS links connected to a communication technology is obtained as

$$n_\tau = P_{pr,\tau} N \tag{G.8}$$

The energy consumed by communication technology is obtained as

$$E_\tau = P_{pr,\tau} N (E_{Tr} + E_{d,fs,avg}) \tag{G.9}$$

Finally, the total energy consumed by the network is obtained as

$$E_{N/W} = \sum_\tau P_{pr,\tau} N (E_{Tr,\tau} + E_{d,fs,\tau,avg}) \tag{G.10}$$

Here, $E_{d,fs,\tau,avg}$ represents the average energy consumption incurred by the communication technology τ. The average energy $E_{d,fs,\tau,avg}$ over an area D can be obtained as [26]

$$E_{d,fs,\tau,avg} = \frac{1}{D} \iint_D f(r,\theta)\, r\, dr\, d\theta \tag{G.11}$$

Evaluating the above integral over the region D defined as the area between two concentric circles with radius r_a and r_b ($r_b > r_a$) gives

$$E_{d,fs,\tau,avg} = \frac{1}{\pi(r_b^2 - r_a^2)} \int_0^{2\pi} \int_{r_a}^{r_b} (k\mathcal{E}_e + k\mathcal{E}_{fs} r^2) r\, dr\, d\theta \tag{G.12}$$

Upon solving, we obtain the generalized closed form expression for $E_{d,fs,\tau,avg}$ as

$$E_{d,fs,\tau,avg} = k\mathcal{E}_e + \frac{k\mathcal{E}_{fs}}{2}(r_a^2 + r_b^2) \tag{G.13}$$

where r_a and r_b are defined as

$$r_a = \begin{cases} 0, & \text{for BLE} \\ r_{BLE,fs}, & \text{for LTE} \\ r_{LTE,fs}, & \text{for Wi-Fi} \\ r_{Wi\text{-}Fi,fs}, & \text{for LoRa} \end{cases} \tag{G.14}$$

and

$$r_b = \begin{cases} r_{BLE,fs}, & \text{for BLE} \\ r_{LTE,fs}, & \text{for LTE} \\ r_{Wi\text{-}Fi,fs}, & \text{for Wi-Fi} \\ r_{LoRa,fs}, & \text{for LoRa} \end{cases} \tag{G.15}$$

Similarly the average delay of the network $T_{avg,fs}$ can be derived as,

$$T_{avg,fs} = \frac{1}{\pi(r_b^2 - r_a^2)} \int_0^{2\pi} \int_{r_a}^{r_b} (T_{trans} + T_{prop}) r \, dr \, d\theta \tag{G.16}$$

where the expressions for T_{trans} and T_{prop} are detailed in Section G.6. On solving, we obtain the expression for the average network delay $T_{avg,fs}$ as,

$$T_{avg,fs} = \frac{k}{R} + \frac{2(r_a^2 + r_b^2 + r_a r_b)}{3c(r_a + r_b)} \tag{G.17}$$

where r_a and r_b are defined as above. The parameter c is the speed of light in vaccum which is equal to 3×10^8 m/s.

G.4.2 FSMP Model

In contrast to the FS model, a link can be assigned with any of the communication technology irrespective of its radial distance. Thus, the connection to a communication technology depends on both E_{Tr} and E_d. As described in the previous section, the link will be connected to a communication protocol if the UAV is present in the range of that communication protocol. Now we need to obtain the range beyond which a communication protocol extends for energy-efficient data transmission over hybrid UAV-assisted wireless network. For example, the BLE can be assigned to a link even if the radial distance is greater than the threshold distance. The requirement is that the total energy consumed with d^4 model of BLE should be less than the total energy consumed with d^2 model of LTE which is defined as

$$E_{Tr,BLE} + E_{d,fsmp,BLE} < E_{Tr,LTE} + E_{d,fs,LTE} \tag{G.18}$$

We need to derive the range of BLE, r_{BLE} by considering equality in (G.18).

$$\frac{kP_{T,BLE}}{R_{BLE}} + k\mathcal{E}_e + k\mathcal{E}_{mp} d_g^4 = \frac{kP_{T,LTE}}{R_{LTE}} + k\mathcal{E}_e + k\mathcal{E}_{fs} d_g^2 \tag{G.19}$$

Solving the above equation, we get the threshold distance corresponding to BLE which is defined as

$$r_{BLE} = \sqrt{\frac{\mathcal{E}_{fs} \pm \sqrt{\mathcal{E}_{fs}^2 - 4(\frac{P_{T,BLE}}{R_{BLE}} - \frac{P_{T,LTE}}{R_{LTE}})}}{2\mathcal{E}_{mp}}} \tag{G.20}$$

Similarly, we can obtain the range for LTE, Wi-Fi, and LoRa. With the new range, we can obtain the area over which BLE, LTE, Wi-Fi, and LoRa exist.

$$A_\tau = \begin{cases} \pi x_{BLE,fsmp}^2, & \text{for BLE} \\ \pi x_{LTE,fsmp}^2 - \pi r_{BLE,fsmp}^2, & \text{for LTE} \\ \pi x_{Wi\text{-}Fi,fsmp}^2 - \pi r_{LTE,fsmp}^2, & \text{for Wi-Fi} \\ B^2 - \pi x_{Wi\text{-}Fi,fsmp}^2, & \text{for LoRa} \end{cases} \tag{G.21}$$

where, τ represents the different communication technologies and x_τ is given as

$$x_\tau = \begin{cases} \sqrt{r_\tau^2 - h_i^2}, & \text{if } r_\tau > h_i; \\ 0, & \text{otherwise,} \end{cases} \tag{G.22}$$

Let P_{pr} represents the probability that a UAV is present in the area corresponding to a communication technology which is obtained as

$$P_{pr,\tau} = \frac{A_\tau}{B^2}. \tag{G.23}$$

Thus, the number of UAV-GS links connected to a communication technology is obtained as

$$n_\tau = P_{pr,\tau} N \tag{G.24}$$

The energy consumed by communication technology is obtained as

$$E_\tau = P_{pr,\tau} N (E_{Tr} + E_{d,fsmp,avg}) \tag{G.25}$$

Finally, the total energy consumed by the network is obtained as

$$E_{N/W} = \sum_\tau P_{pr,\tau} N (E_{Tr,\tau} + E_{d,fsmp,\tau,avg}) \tag{G.26}$$

Similar to (G.12) for the FS model, the average FSMP energy consumption $E_{d,fsmp,\tau,avg}$ for a communication technology τ is obtained as

$$E_{d,fsmp,\tau,avg} = \frac{1}{\pi(r_b^2 - r_a^2)} \int_0^{2\pi} \left[\int_{r_a}^{r_c} (k\mathcal{E}_e + k\mathcal{E}_{fs} r^2) r \, dr \, d\theta + \int_{r_c}^{r_b} (k\mathcal{E}_e + k\mathcal{E}_{mp} r^4) r \, dr \, d\theta \right] \tag{G.27}$$

Solving the above equation, the expression for the average energy consumption $E_{d,fsmp,\tau,avg}$ is given as

$$E_{d,fsmp,\tau,avg} = k\mathcal{E}_e + \frac{k}{(r_b^2 - r_a^2)}\left[\frac{k\mathcal{E}_{fs}}{2}(r_c^4 - r_a^4) + \frac{k\mathcal{E}_{mp}}{3}(r_b^6 - r_c^6)\right] \tag{G.28}$$

where r_a, r_b, and r_c are defined as

$$r_a = \begin{cases} 0, & \text{for BLE} \\ r_{BLE,fsmp}, & \text{for LTE} \\ r_{LTE,fsmp}, & \text{for Wi-Fi} \\ r_{Wi\text{-}Fi,fsmp}, & \text{for LoRa} \end{cases} \tag{G.29}$$

$$r_b = \begin{cases} r_{BLE,fsmp}, & \text{for BLE} \\ r_{LTE,fsmp}, & \text{for LTE} \\ r_{Wi\text{-}Fi,fsmp}, & \text{for Wi-Fi} \\ r_{LoRa,fsmp}, & \text{for LoRa} \end{cases} \tag{G.30}$$

and

$$r_c = \begin{cases} r_{BLE,fs}, & \text{for BLE} \\ r_{LTE,fs}, & \text{for LTE} \\ r_{Wi\text{-}Fi,fs}, & \text{for Wi-Fi} \\ r_{LoRa,fs}, & \text{for LoRa} \end{cases} \tag{G.31}$$

Similarly the average delay of the network $T_{avg,fsmp}$ can be derived as,

$$T_{avg,fsmp} = \frac{1}{\pi(r_b^2 - r_a^2)}\int_0^{2\pi}\int_{r_a}^{r_b}(T_{trans} + T_{prop})r\,dr\,d\theta \tag{G.32}$$

where the expressions for T_{trans} and T_{prop} are detailed in Section G.6. Solving the above equation, we obtain the expression for the average network delay $T_{avg,fsmp}$ as,

$$T_{avg,fsmp} = \frac{k}{R} + \frac{2(r_a^2 + r_b^2 + r_a r_b)}{3c(r_a + r_b)} \tag{G.33}$$

where r_a, r_b, and c are defined as above. It can be observed that the average network delay for both the hybrid FS and FSMP schemes is equal. This is expected as the delay in both cases is directly proportional to d_g. This is also verified through simulations in the subsequent sections.

G.5 RL-Based Energy-Efficient Data Transmission over Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network Formation

RL is a branch of machine learning in which an agent learns the optimal policy strategy through a set of actions. In a RL framework, the agent is the entity that

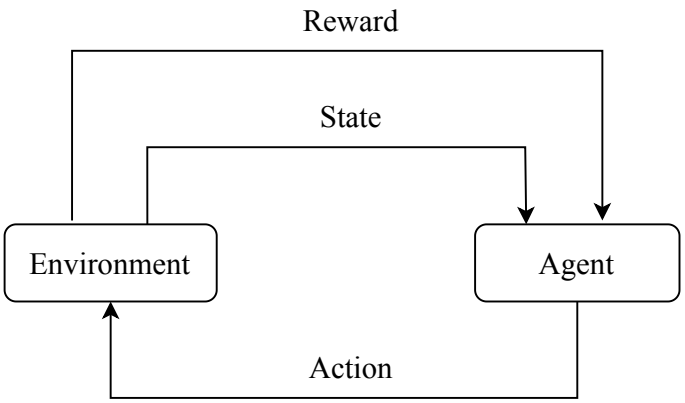


Figure G.2: An illustration of Q-learning.

Table G.3: Q-table corresponding to state-action pairs.

States	Actions			
	A_1	A_2	$\cdots$	A_M
S_1	$Q(S_1, A_1)$	$Q(S_1, A_2)$	$\cdots$	$Q(S_1, A_M)$
S_2	$Q(S_2, A_1)$	$Q(S_2, A_2)$	$\cdots$	$Q(S_2, A_M)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S_N	$Q(S_N, A_1)$	$Q(S_N, A_2)$	$\cdots$	$Q(S_N, A_M)$

tries to learn and model the environment that it interacts with. The agent interacts with the environment through a set of trial-and-error actions. Based on these actions the environment either rewards or penalizes the agent. The feedback thus obtained allows the agent to navigate the environment by transitioning to new states, thereby updating its optimal policy strategy. RL is generally used to solve problems involving Markov Decision Process (MDP) where the state transition probability of moving to the next state is dependent only on the present state and not the previous other states. In this paper, we consider two RL algorithms such as Q-learning and DQN which are described next.

G.5.1 Q-Learning

In this section, we describe the proposed Q-learning approach for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. Q-learning is a model-free reinforcement learning algorithm in which the agent directly interacts with the environment to achieve the optimal policy. The algorithm is value-based and finds the value function without any prior knowledge of the state transition probabilities [27]. Compared to model-based algorithms, Q-learning offers relatively less execution time and storage cost, provided the state and action space are not large. Since the algorithm operation is reward-dependent, proper design of the reward function is crucial for the algorithm's performance and conduct [28, 29]. In this work, the goal of the Q-learning algorithm is to transmit the data from UAV to GS with minimum energy consumption. In our hybrid system, the states, actions, and rewards as shown in Fig. G.2 are detailed as follows:

G.5.1.1 States

Let $S = \{s_n\}$ represent a set of N states for the Q-learning algorithm where $n = 1, 2, ..., N$. In this work, each state represents the communication link between the UAV and GS. Since there are N UAVs in the network, the number of UAV-GS communication links is N, and hence the number of states is N.

G.5.1.2 Action

Let each state $s_n \in S$ be associated with a set of M actions $A = \{a_m\}$, where each action represents the agent choosing a communication technology to assign to a UAV-GS link. In our hybrid system, there are four communication technologies and hence, there are four actions i.e., $M = 4$ in our Q-learning algorithm.

G.5.1.3 Reward/Penalty

Based on the action, the agent receives a reward/penalty from the environment. If γ is denoted as the reward/penalty, then $\gamma = \gamma(s_n, a_m)$ is the reward/penalty that an agent obtains when it is present in state s_n and performs an action a_m. In our hybrid system, the agent will receive the reward/penalty after selecting the communication technology for the UAV-GS link. Since the objective of the proposed work is energy-efficient data transmission over a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network, γ is considered in terms of the network energy consump-

tion. The expression for γ is defined as

$$\gamma = \begin{cases} \dfrac{10^{-}6}{E_{N/W}} & \text{(G.34)} \\ \dfrac{1}{E_{N/W}} & \text{(G.35)} \\ e^{-E_{N/W}} & \text{(G.36)} \\ \dfrac{W_1}{E_{1,N/W}} + \dfrac{W_2}{E_{2,N/W}}, & \text{(G.37)} \end{cases}$$

where, $E_{1,N/W}$ and $E_{2,N/W}$ denote the network energy consumed for transmitting k-bits of information and the network energy consumed for transmitting over a distance. Further, W_1 and W_2 are the weights assigned for $E_{1,N/W}$ and $E_{2,N/W}$, respectively. Since a higher reward translates to an optimized network, we consider γ as the inverse function of energy.

G.5.1.4 Updating the Q-value

The reward/penalty from the agent is used to update the Q-value corresponding to a state-action pair as given in Table G.3. When a UAV-GS link is assigned to a communication protocol, it results in a reward. The new Q-value is obtained as

$$Q(s_t, a_t) \leftarrow (1-\lambda)Q(s_t, a_t) + \lambda(\gamma + \Delta \max_a Q(s_{t+1}, a)), \tag{G.38}$$

where, λ and Δ are the learning rate and discount corresponding to the Q-learning approach, respectively. Here, $\max_a Q(s_{t+1}, a)$ takes the maximum of the future reward and applies it to the reward for the current state.

Algorithm 8 describes the Q-learning approach proposed in this work. In the proposed Q-learning algorithm, initialize the Q-matrix with the number of UAV-GS links as rows and the communication protocols as columns. The UAV-GS links are ordered as rows in decreasing order of their distance to the ground station. Now, starting from the first row of the Q-matrix, randomly select a communication protocol. With the selected communication protocol, calculate the network energy consumption. Calculate the reward using (G.34), (G.35), (G.36), and (G.37) and update the Q-value of the link using (G.38). Next, transition to the next state based on the decreasing distance to the ground station. The state transition rule employed in this work is partially motivated by [30], where the authors have used decreasing channel state information as a means to transition to the next state. Repeat the above procedure until the UAV-GS link closest to the ground station is updated. The above steps are repeated again for a large number of iterations such that all links and communication protocols are covered. Once the training is completed, connect the UAV-GS link with a communication protocol which results in the highest Q-value. Calculate the network energy consumption and average network delay for the energy-efficient data transmission over a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. With the increase in the number of state-action pairs, the

Algorithm 8: Q-learning approach to determine the optimized hybrid UAV-assisted communication network

Input: N UAVs, M communication technologies, Maximum number of episodes K_{max}, λ, Δ, exploration-exploitation factor ε

Output: Network energy consumption and average network delay

Initialize the rows and columns of the Q-matrix with zeros;
Assign the UAV-GS links to the rows of the Q-matrix in descending order of distance to the ground station;
Assign the communication technologies to the columns of the Q-matrix;
Training
for $i = 1$ **to** K_{max} **do**
    Select the first state s_0;
    for $t = 1$ **to** N **do**
        Generate a random number $u \in (0, 1)$;
        if $u > \varepsilon$ **then**
            Select action from Q-matrix which has maximum Q-value $a_t = \max_a Q(s_t, a_t)$;
        else
            Select a random communication technology as action a_t;
        end
        Calculate the network energy consumption using (G.1), (G.3), and (G.4);
        Obtain γ using (G.34), (G.35), (G.36), and (G.37);
        Update the Q-value using (G.38);
        if $s_t == s_N$ **then**
            break;
        else
            Update to next state $s_t = s_{t+1}$ in the order of decreasing distance to ground station;
        end
    end
end
Validation
In each row of the Q-matrix, select the indices with maximum Q-value;
Assign the communication technology with the highest Q-value to the UAV-GS link;
Calculate the network energy consumption and average network delay;

Q-matrix employed in the Q-learning algorithm would require additional memory to store the Q-values. This can often lead to increased memory overhead thereby affecting the performance of the algorithm [30]. The DQN algorithm is set to resolve these issues by using a neural network to approximate the Q-values [31]. By using a neural network, the DQN algorithm essentially preserves the relative significance

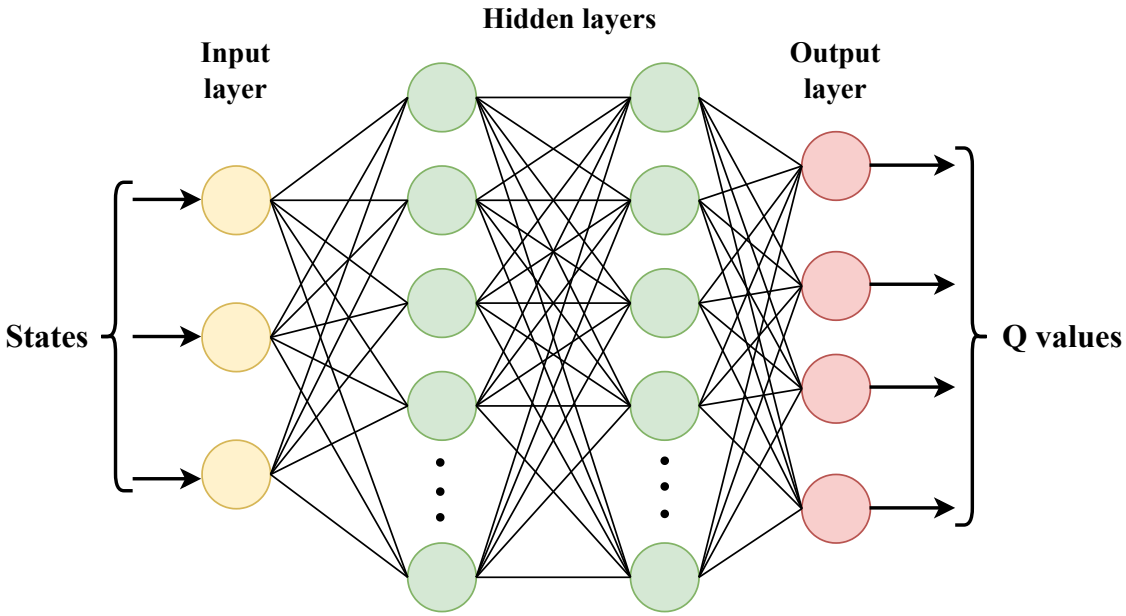


Figure G.3: Neural network model used in DQN.

between the Q-values and hence provides similar results as that of Q-learning [32]. Next, we discuss the details of the proposed DQN algorithm.

G.5.2 Deep Q-Network (DQN)

In this section, we present the DQN algorithm for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network.

The DQN algorithm when applied to this work follows a similar construct as the Q-learning algorithm. States and actions follow the same descriptions as defined for Q-learning. Hence, for a UAV-assisted wireless network with N UAVs, the number of UAV-GS communication links is N, and hence the number of states is N. As there are four communication technologies to choose from, there are four actions. For a state s_t, the DQN algorithm follows a ε-greedy policy to select an action a_t. This means that a random action is selected (exploration) with a probability ε and the action corresponding to maximum $Q(s_t, a_t)$ is selected with probability $(1 - \varepsilon)$ (exploitation). Thus an exploration-exploitation tradeoff is provided so that the DQN algorithm is able to train successfully without falling into a local optima [30]. At any step t, the reward/penalty γ_t for state s_t and action a_t is calculated based on (G.35).

The DQN network is primarily made up of two neural network models, a policy network and a target network. Each UAV-GS link with its corresponding length is given as input to the neural networks. As shown in Fig. G.3, a fully connected neural network with two hidden layers forms the policy network. The first hidden layer consists of 256 neurons while the second hidden layer contains 128 neurons. All the layers are activated using the ReLu activation function during training. The target network is a cloned replica of the policy network. It has the same architecture as the

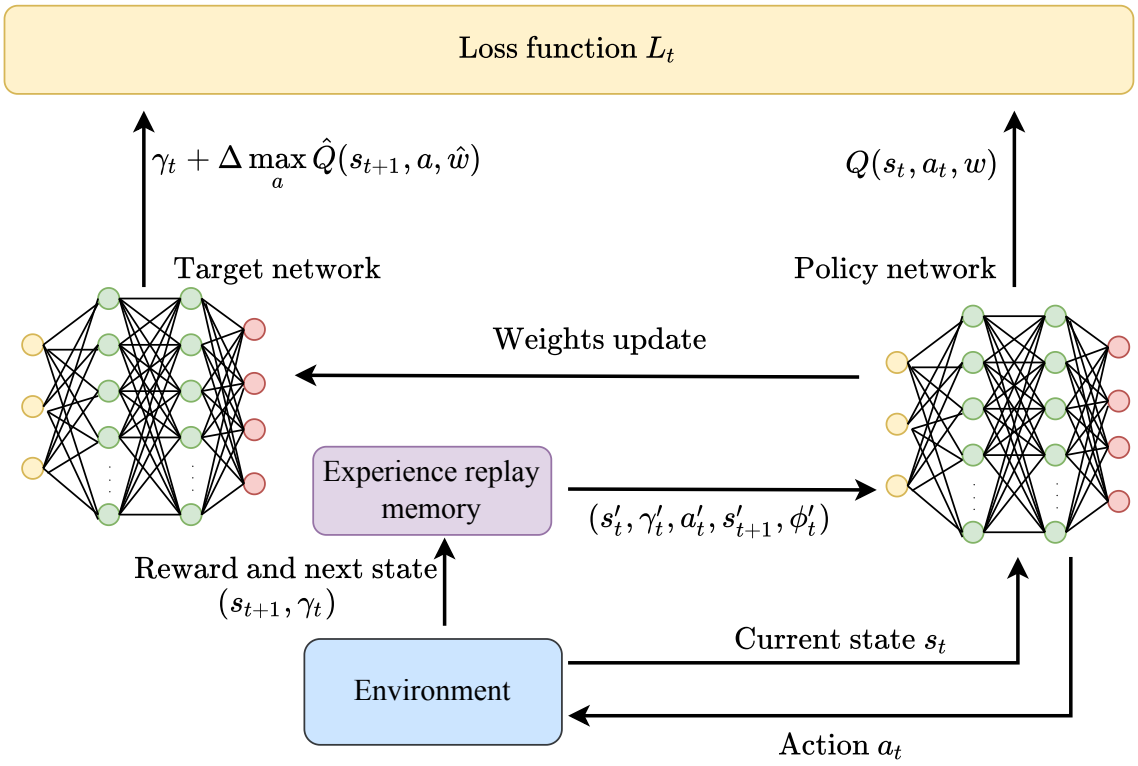


Figure G.4: Block diagram for the DQN algorithm.

policy network and serves the purpose of providing stability to the DQN algorithm while training. The output layer of the policy network consists of 4 outputs with a linear activation. The outputs correspond to the prediction of the state-action values $Q(s_t, a_t)$ at step t.

As shown in Fig. G.4, the action a_t for state s_t is selected based on ε-greedy policy. Choosing action a_t results in a state transition from state s_t to state s_{t+1} and a reward γ_t. The state transition follows the same distance-based rule used in Q-learning. The tuple $K = \{s_t, \gamma_t, a_t, s_{t+1}, \phi_t\}$ is then stored as a deque in the experience replay memory. Here, ϕ_t is a variable that indicates the final state. In the next step, a random mini-batch sample $(s'_t, \gamma'_t, a'_t, s'_{t+1}, \phi'_t)$ is selected which is used for training the policy network. The intention of selecting a random sample is to reduce the correlations between states and provide stability to the training process. The training is now carried out and the loss function is minimized. The loss function for step t is obtained as [30]

$$L_t = E\left[\left(\gamma_t + \Delta \max_a \hat{Q}(s_{t+1}, a, \hat{w}) - Q(s_t, a_t, w)\right)^2\right] \tag{G.39}$$

where Δ represents the discount factor ($\Delta \in [0, 1]$), $Q(s_t, a_t, w)$ represents the Q-value predicted using the policy network when trained with weights w, and $(\gamma_t + \Delta \max_a \hat{Q}(s_{t+1}, a, \hat{w})$ represents the output from the target network that is trained with weights $\hat{w}$. It is to be noted that the weights of the target network $\hat{w}$ are copied from the policy network every ζ step. Both the policy network and target network are trained using the Adam optimizer. The policy network is updated using the

below equation [30]

$$Q^*(s_t, a_t) \leftarrow Q(s_t, a_t) + \lambda(\gamma_t + \Delta \max_a \hat{Q}(s_{t+1}, a) - Q(s_t, a_t)) \tag{G.40}$$

where λ is the learning rate and $Q^*(s_t, a_t)$ is the new updated Q-value. The detailed steps of the proposed DQN algorithm are described in Algo. 9.

G.6 Evaluation Metrics

In this section, we derive the expressions for the performance metrics such as network energy consumption, and average network delay considered for the evaluation of the performance of the proposed method.

G.6.1 Average Network Delay

It is defined as the ratio of the sum of the delays incurred for transmitting the data in each UAV-GS link, to the total number of UAVs. The total delay of a UAV-GS link is the sum of the propagation delay and the transmission delay of the link.

G.6.1.1 Propagation delay

It is the delay incurred over a UAV-GS link for propagating the data over a radial distance d_g [33]. The expression for propagation delay is obtained as

$$T_{prop} = \frac{d_g}{c}, \tag{G.41}$$

where c is the velocity of the light which is equal to 3×10^8 m/s.

G.6.1.2 Transmission delay

It is the delay incurred for transmitting k bits of information over a UAV-GS link [33]. The expression for transmission delay is obtained as

$$T_{Tr} = \frac{k}{R}. \tag{G.42}$$

Thus, the average network delay is obtained as

$$T_{avg} = \frac{1}{N} \sum_{i=1}^{N} T_{total,i}, \tag{G.43}$$

where $T_{total} = T_{prop} + T_{Tr}$.

G.6.2 Network Energy Consumption

It is the total energy consumed for transmitting the data from each UAV to the GS. It is expressed as

$$E_{total} = \sum_{i=1}^{N} (E_{Tr,i} + E_{d,i}). \tag{G.44}$$

Algorithm 9: DQN approach to determine the optimized hybrid UAV-assisted wireless network

Input: N UAVs, M communication technologies, Number of episodes K_{max}, λ, Δ, ε, ζ

Output: Network energy consumption and average network delay

Initialize the experience replay memory, policy network weights w, and target network weights $\hat{w}$;
Arrange the UAV-GS links (states) in decreasing order of distance to the ground station;
Training
for $i = 1$ **to** K_{max} **do**
    Select the first state s_0 (first UAV-GS link) from the sorted list;
    $\phi_0 = 0$, $j = 0$;
    for $t = 1$ **to** N **do**
        Generate a random number $u \in (0, 1)$;
        if $u > \varepsilon$ **then**
            Select action $a_t = \max_a Q(s_t, a_t, w)$;
        else
            Select a random action a_t;
        end
        Calculate the network energy consumption using (G.1), (G.3), and (G.4);
        Obtain reward γ_t, using (G.35) ;
        Obtain next state s_{t+1} in order of decreasing distance to ground station;
        if $s_{t+1} == s_N$ **then**
            $\phi_t = 1$, $j = j + 1$;
        end
        Update experience replay memory with $K = (s_t, \gamma_t, a_t, s_{t+1}, \phi_t)$;
        if $\phi_t = 1$ **then**
            break;
        end
        When replay memory is full, randomly select a minibatch sample $(s'_t, \gamma'_t, a'_t, s'_{t+1}, \phi'_t)$;
        Predict $Q(s_t, a_t, w)$ from policy network;
        Predict $(\gamma_t + \Delta \max_a \hat{Q}(s_{t+1}, a, \hat{w})$ from target network;
        Compute loss using (G.39);
        Update weights w of policy network using (**??**);
        if $j == \zeta$ **then**
            Update the weights of target network $\hat{w}$;
        end
    end
    Decrease ε with a decay rate;
end

Validation
For each UAV, obtain the Q-values from the trained policy network;
Assign the communication technology with the highest Q-value to the UAV-GS link;
Calculate the network energy consumption and average network delay;

G.6.3 Energy Consumption for UAV Movement

The energy consumed by a UAV (in joules) to move a distance d is given by [34]

$$E_v = T \cdot d \tag{G.45}$$

where T represents the thrust force to move forward that is measured in Newtons ($kg\text{-}m/sec^2$). The thrust force is obtained as

$$T = \frac{m \times g}{r} \tag{G.46}$$

where m is the total weight of the aircraft (in kg), g is the acceleration due to gravity (9.8 m/sec^2), and r is a unitless parameter defined as the lift-to-drag ratio [34]. The lift-to-drag ratio essentially denotes the efficiency of aircraft design. A recommended lift-to-drag ratio is required to keep the aircraft airborne during steady flight. Commercial passenger aircraft have a lift-to-drag ratio between 10-20 whereas r for cruising helicopters is about 4. Typical lift-to-drag value for small and large scale UAVs is 3 [34].

G.7 Numerical Results

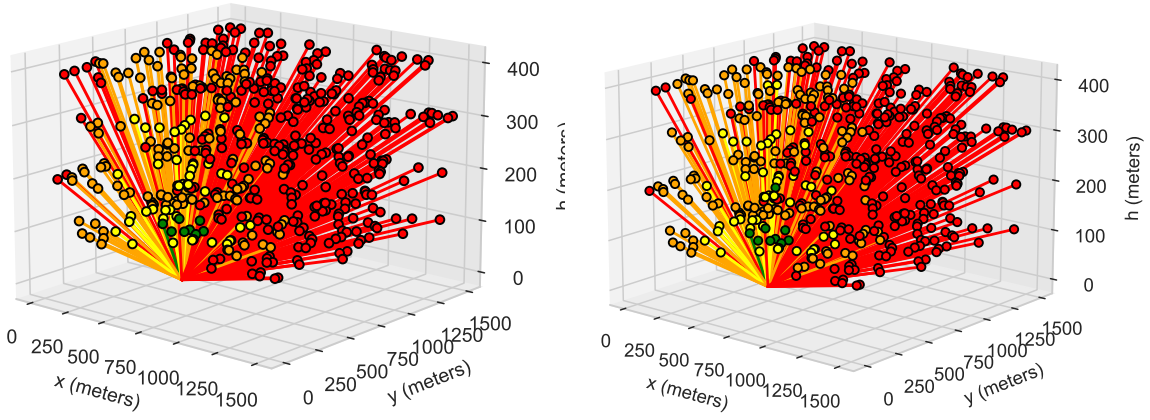
In this section, we first present the simulation setup considered for generating the simulation results. Then, we present the simulation results corresponding to both FS and FSMP models to verify the analytical derivations. We also present the simulation results to show the effect of the velocity of UAVs and packet size on the network energy consumption and average delay. Finally, the simulation results are presented to compare the performance of the proposed RL algorithms with other models.

G.7.1 Simulation Setup

The simulation setup is considered over a 1500 $\times$ 1500 m^2 area where the GS is located at $(500, 500, 0)$. A total of 500 UAVs are deployed over the area with varying heights of 100, 200, 300, and 400 meters. Each UAV is equipped with four different communication technologies such as BLE, LTE, Wi-Fi, and LoRa. The packet arrival rate follows Poisson distribution with randomly chosen mean taken from the set $\{1, 2, \cdots, 100\}$. Each packet contains 128 bits and hence the total number of bits received for each UAV follows a random distribution. The various simulation

Table G.4: The values of the parameters used for the performance evaluation and comparison of different methods.

Parameter	**Value**
Horizontal area	1500×1500
Hovering heights	$\{100, 200, 300, 400\}$
Number of UAVs, N	500
Location of GS	$(500, 500, 0)$
Velocity of UAV, ν	2
Packet length, k	128 bits
The parameter of Poisson distribution, λ	100
Speed of the transmission, c	3×10^8
$\mathcal{E}_e$	25 nJ/bit
$\mathcal{E}_{fs}$	10 pJ/bit/m^2
Initial energy at a UAV E	10 Joules



(a) (b)

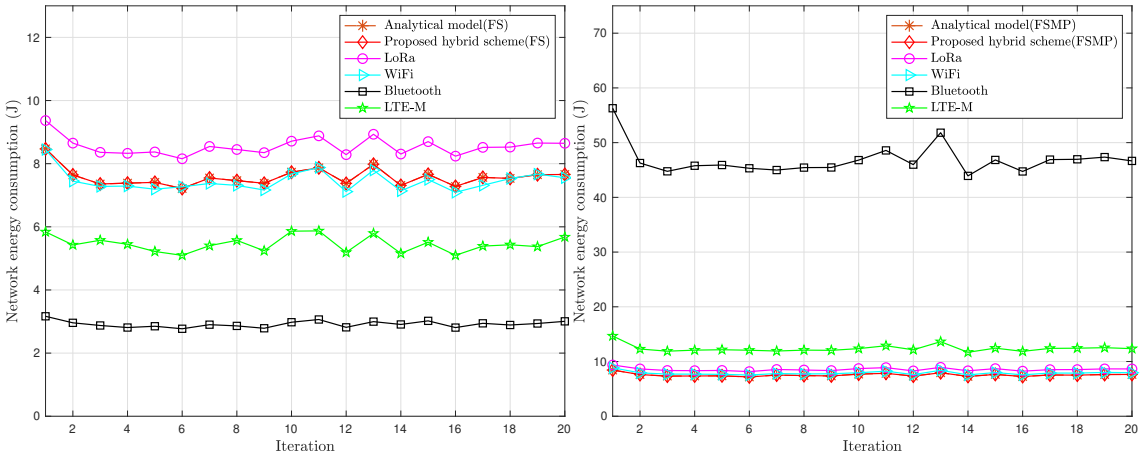
Figure G.5: An illustration of the UAV-assisted wireless network topology for (a) FS configuration and (b) FSMP configuration. Here, GS is considered to be located at $(500, 500, 0)$ which is the center of the terrestrial area. The colors green, yellow, orange, and red depict BLE, LTE, Wi-Fi, and LoRa communication protocols respectively.

parameters utilized for the different runs are listed in Table G.4. Additionally, the communication parameters for each specific technology are also provided in Table G.2.

G.7.2 Simulation Results (FS and FSMP)

Based on the above simulation parameters, we discuss the various numerical results obtained.

Fig. G.5a depicts the topology of the proposed hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network following the FS energy model. It can be observed from Fig. G.5a that the majority of the UAVs utilize the LoRa protocol (depicted

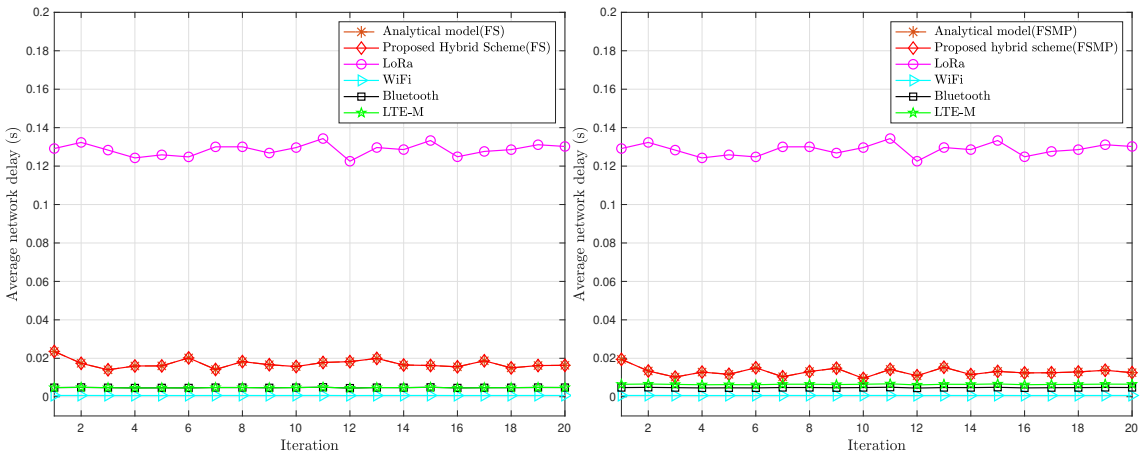


(a) (b)

Figure G.6: Variation of network energy consumption for (a) free space and (b) free space and multipath model for random UAV configurations. The analytical plots are obtained using equations (G.10) and (G.26), respectively. The ground station is located at $(500, 500, 0)$ for all the random configurations.

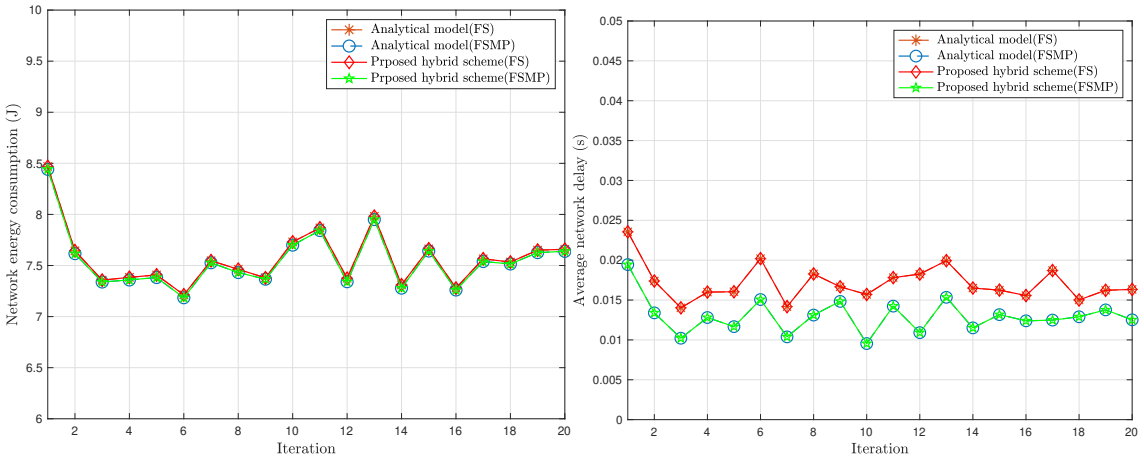
in red) to connect to the GS. The UAVs at medium range utilize Wi-Fi (yellow) and LTE (orange) to establish the UAV-GS link. The UAVs that are closer to the GS are connected via the BLE (green) communication protocol. When UAVs are closer to the GS, the BLE communication protocol consumes minimum energy as compared to other protocols. Hence, a small concentration of UAVs connected to BLE can be observed closer to the GS in Fig. G.5a. For UAVs situated farther than the path loss distance of BLE, the communication protocol with minimum energy consumption is connected. When the UAV-assisted wireless network follows the proposed hybrid scheme according to the FSMP energy model, the UAV-GS link connections are switched as shown in Fig. G.5b. It can be observed from Fig. G.5b that a small number of UAVs switch their communication protocol from LTE to BLE. This is due to the fact that the energy consumption of BLE while following the d^4 energy model is less than the energy consumption of LTE that is following the d^2 energy model. Similar changes can be observed between other communication protocols.

In Fig. G.6a, the network energy consumption for the proposed hybrid approach is compared with the conventional shortest path routing algorithm with individual communication technology while considering the FS path loss model. It can be observed from Fig. G.6a that the energy consumption exhibited by the proposed hybrid approach matches the energy consumption of the conventional approach for Wi-Fi protocol. However, the energy consumption is greater than that of BLE. This is due to the lower transmit power of BLE as compared to the average transmit power in the proposed hybrid scheme. The conventional approach for individual communication technology utilizes the hop-based shortest path algorithm to send packets from the UAV to GS. It is observed from Fig. G.6a that the energy consumption from the analytical model is equal to that of the simulation results, which is as expected. In Fig. G.6b, the proposed hybrid approach is compared to the



(a) (b)

Figure G.7: Variation of average network delay for (a) free space and (b) free space and multipath energy models for random UAV configurations. The analytical plots are obtained using equations (G.17) and (G.33), respectively. For all the configurations, the GS is located at $(500, 500, 0)$.



(a) (b)

Figure G.8: Variation of (a) network energy consumption and (b) average network delay for the proposed hybrid network. The plots corresponding to the analytical results are obtained using equations (G.10), (G.17), (G.26), and (G.33), respectively. In all these configurations, the GS is located at $(500, 500, 0)$.

conventional shortest path routing algorithm with individual communication technology in terms of network energy consumption. Here, we consider the FSMP path loss model. It can be observed from Fig. G.6b that the energy consumption with the proposed hybrid approach is lower than the conventional approach. It is also observed that the conventional approach utilizing BLE consumes more energy as compared to other approaches. This is due to the fact that r_τ is very low for BLE, resulting in a higher number of UAV-GS links following the d^4 model. Further, Fig. G.6b also illustrates the variation of the analytical results of the proposed approach.

It is observed from Fig. G.6b that the analytical results are matching with the simulation results verifying the analytical derivations.

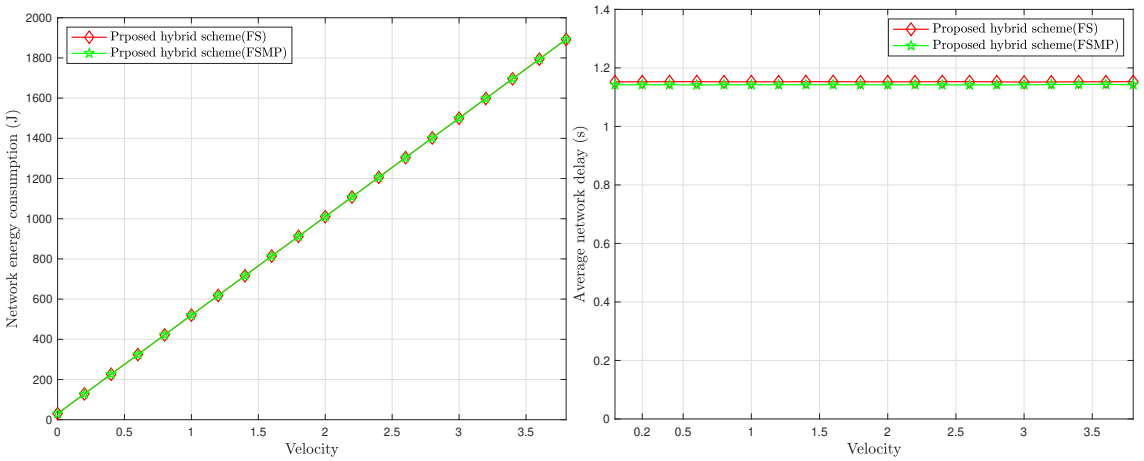
Figs. G.7a and G.7b provide the variation of average network delay for the proposed hybrid scheme with the conventional shortest path routing algorithm for individual communication technology. It can be observed from Figs. G.7a and G.7b that there is no significant variation in the delay for the proposed hybrid scheme that is following the FS and FSMP energy models. This can be attributed to the delay parameter that primarily depends on the transmission delay. The transmission delay depends on the data rate of the communication protocol. As LoRa has a low data rate compared to other communication protocols, LoRa has a higher delay. Thus, as observed from Figs. G.6b and G.7b, the proposed hybrid scheme following the FSMP energy model offers superior performance in terms of minimum energy consumption and low average network delay compared to the conventional shortest path routing algorithm with individual communication technology.

The variation of network energy consumption for the proposed hybrid scheme following the FS and FSMP energy models is depicted in Fig. G.8a. It is observed from Fig. G.8a, that the network energy consumption for the proposed hybrid scheme is similar across both the FS and FSMP energy models. Fig. G.8b shows the average network delay for the proposed hybrid scheme following the FS and FSMP energy models. It can be observed from Fig. G.8b that the proposed hybrid scheme with the FSMP energy model exhibits lower delay as compared to the proposed hybrid scheme following the FS energy model. This can be explained by observing the individual connections for each communication technology in both schemes. In the proposed hybrid scheme following the FS energy model, the distribution of UAVs connected to the different communication technologies are as follows: 14 connected to BLE, 153 connected to LTE, 279 connected to Wi-Fi, and 54 connected to LoRa. This distribution is altered when the proposed hybrid scheme with the FSMP energy model is used. The connections are switched with 21 UAVs connected to BLE and 162 UAVs connected to LTE. Moreover, the number of UAVs connected to LoRa reduces to 38. This change in the connection distribution decreases the delay in the proposed hybrid scheme utilizing the FSMP energy model. As BLE and LTE offer higher data rates as compared to the LoRa protocol, a lower delay is observed in the proposed hybrid scheme following the FSMP energy model. Thus the proposed hybrid scheme with FSMP energy model offers minimum energy consumption and reduced delay for the overall connected network.

G.7.2.1 Effect of UAV Velocity

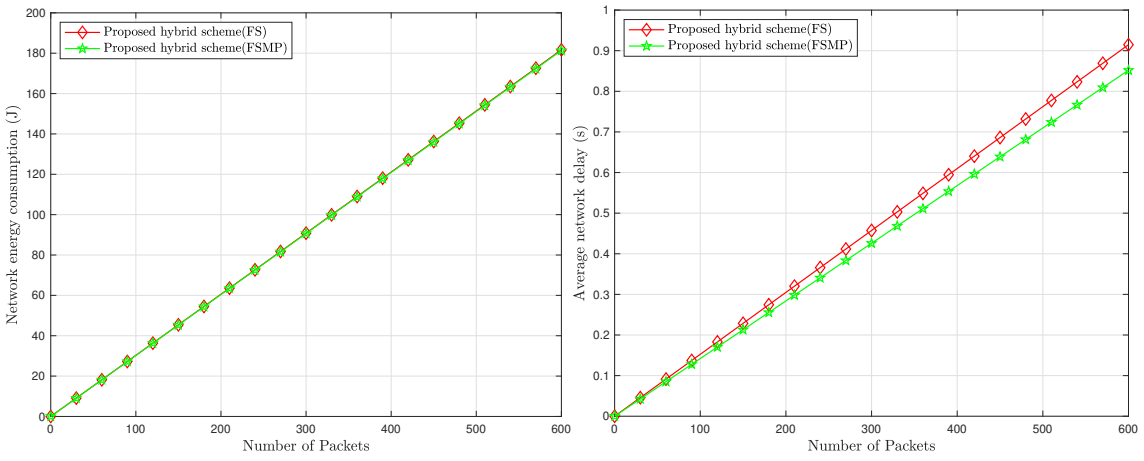
In this section, we show the effect of UAV velocity on the performance of the proposed hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. To obtain the variation of network energy consumption and average network delay, the velocity of each UAV is increased linearly from 0 to 3.8 m/s in steps of 0.2 m/s.

Figs. G.9a and G.9b show the variation of the network energy consumption and average delay with increasing velocity of the UAV. Here, the time of travel (moving



(a) (b)

Figure G.9: Effect of UAV velocity on (a) network energy consumption and (b) average network delay for both FS and FSMP model for random UAV configurations. GS is located at $(500, 500, 0)$ for all the random configurations.



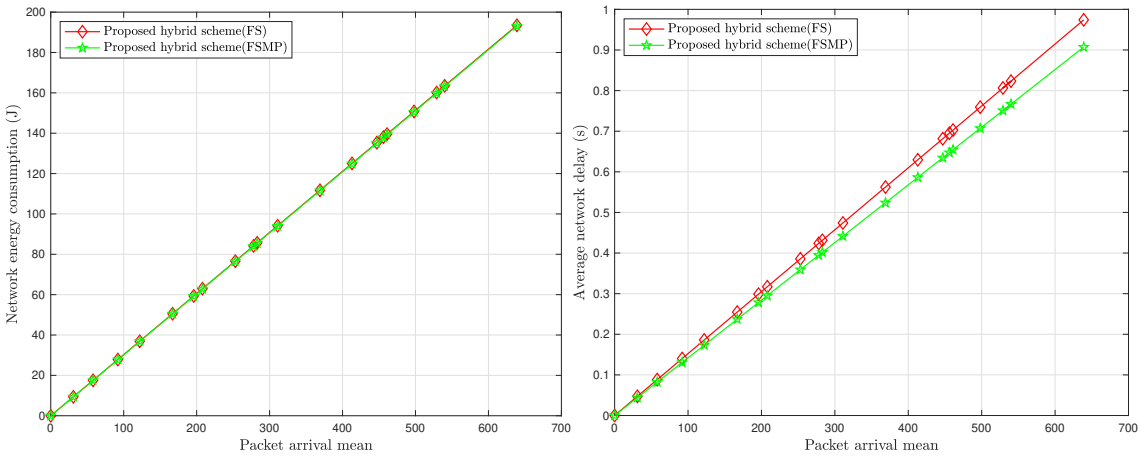
(a) (b)

Figure G.10: Effect of increasing the number of packets on (a) network energy consumption and (b) average network delay for both FS and FSMP model for random UAV configurations. GS is located at $(500, 500, 0)$ for all the random configurations.

time) for each UAV is considered to be 1 second. It is observed from Fig. G.9a that the energy consumption increases linearly with velocity. This is because, for a fixed time of travel, the distance traveled increases with increased velocity which consumes more energy. From Fig. G.9b, it can be observed that the velocity has a constant effect on the average delay as the time of travel is fixed for all velocities of UAVs.

G.7.2.2 Effect of Packet Size

In this section, we show the effect on the performance of the proposed hybrid system



(a) (b)

Figure G.11: Effect of increasing the mean of the packet arrival rate on (a) network energy consumption and (b) delay for both FS and FSMP model for random UAV configurations. GS is located at $(500, 500, 0)$ for all the random configurations.

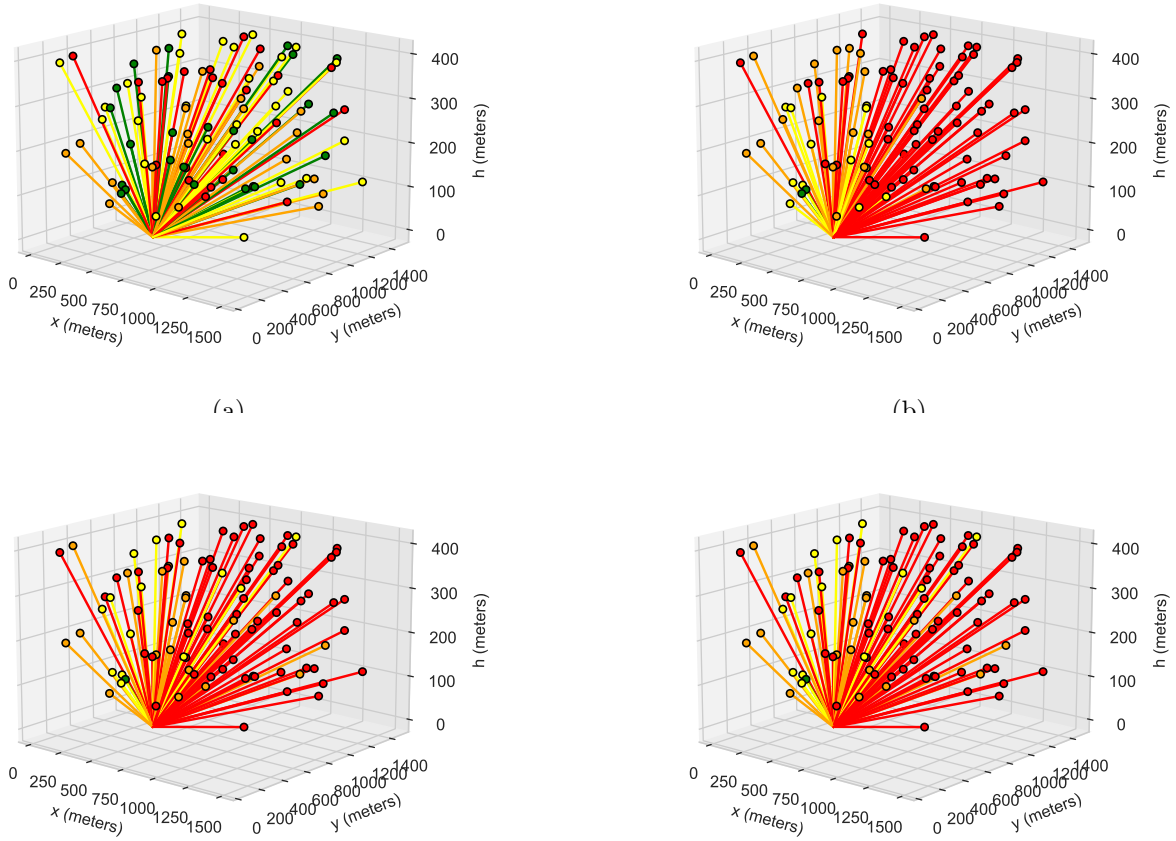
when the number of packets is increased. First, we consider a constant number of packets at each UAV which we increase from 1 to 600. Thereafter, we vary the mean of the Poisson distribution (packet arrival distribution) from 1 to 600.

Figs. G.10a and G.10b show the variation of the network energy consumption and average delay, respectively, for increasing values of the available packets at each UAV. Here, we varied the number of packets from 1 to 600. It is observed from Fig. G.10b that the energy consumption and average delay increase linearly with the number of packets. Further, Fig. G.10b also shows that the average delay is more for the FS model as compared to the FSMP model. This is due to the presence of multi-hop communication in the FS model which adds the queuing delay and processing delay at each hop.

Figs. G.11a and G.11b show the variation of the network energy consumption and average delay, respectively, for increasing values of the mean of the packet arrival rate. It is observed that an increase in the mean linearly increases the energy consumption and average delay as can be seen from Figs. G.11a and G.11b, respectively. Further, it is observed from Fig. G.11b that the average delay is more for the FS model as compared to the FSMP model. This is due to the presence of multi-hop communication in the FS model which adds the queuing delay and processing delay at each hop.

G.7.3 Performance Evaluation of the Proposed RL Algorithms

In this section, we present the numerical results comparing the performance of the proposed Q-learning (Algorithm 8) and DQN (Algorithm 9) algorithms with rule-based algorithm and hybrid random approach in which UAV-GS links are selected uniformly at random from BLE, LTE, Wi-Fi, and LoRa. Further, we show the



(a) (b)

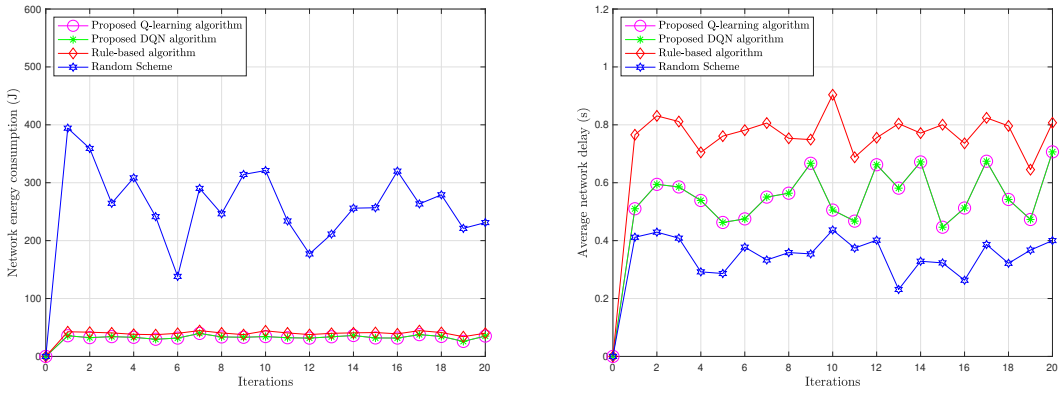
(c) (d)

Figure G.12: An illustration of hybrid UAV-assisted wireless network topology formed with (a) random hybrid scheme, (b) rule-based algorithm, (c) proposed Q-learning algorithm, and (d) proposed DQN algorithm. Here, GS is considered to be located at (500, 500, 0) which is the center of the terrestrial area. The colors green, yellow, orange, and red depict BLE, LTE, Wi-Fi, and LoRa communication protocols, respectively.

effectiveness of the proposed algorithms, by considering the PHY layer parameter, which denotes the number of transmissions required for a packet to get successfully delivered. The simulation setup is similar to the previous scenario wherein an area of $1500 \times 1500\ m^2$ is considered and the GS is placed at $(500, 500, 0)$. We consider a total of 100 UAVs deployed randomly over the considered area with varying heights of 100, 200, 300, and 400 meters. All other simulation parameters remain the same as defined in the previous section.

To study the behavior of Q-learning parameters we carried out extensive simulations using the learning rate and discount with different reward/penalty expressions defined in (G.34), (G.35), (G.36), and (G.37). When the reward is chosen from (G.34), (G.35), and (G.36), the Q-learning algorithm results in the optimized hybrid network at $\lambda = 0.7$ and $\Delta = 0.7$.

For the reward defined in (G.34), (G.35), and (G.36), we varied the Q-learning



(a) (b)

Figure G.13: Performance comparison of the proposed Q-learning algorithm, proposed DQN algorithm, rule-based algorithm, and random hybrid scheme in terms of (a) network energy consumption and (b) average delay for random UAV configurations. The ground station is located at $(500, 500, 0)$ for all the random configurations.

Table G.5: Variation of network energy consumption as a weighted sum of individual energy components $E_{1,N/W}$ and $E_{2,N/W}$. Here, W_1 and W_2 represent the weights.

$\boldsymbol{W_1}$	$\boldsymbol{W_2}$	$\boldsymbol{E_{total}}$ **(J)**
0.1	0.9	0.0204
0.2	0.8	0.0287
0.3	0.7	0.0392
0.4	0.6	0.0512
0.5	0.5	0.0650
0.6	0.4	0.0797
0.7	0.3	0.1002
0.8	0.2	0.1430
0.9	0.1	0.2589

parameters λ and Δ from 0.1 to 0.9 in intervals of 0.1. It is observed that the Q-learning algorithm started to attain the hybrid network with minimum energy consumption when λ is between 0.4 to 0.9 for all Δ. Table G.5 lists the network energy consumed for different weights when the reward is chosen as (G.37). Here, the weights are chosen from $\{0.1, 0.2, \cdots, 0.9\}$. We consider $\lambda = 0.7$ and $\Delta = 0.6$ which are best fit for the proposed Q-learning algorithm. As seen from Table G.5, $W_1 = 0.1$ and $W_2 = 0.9$ lowers the energy consumption. This means that $E_{2,N/W}$ has a higher impact than $E_{1,N/W}$ to attain the hybrid network with minimum energy consumption. This is due to the fact that for a UAV-GS link with k-bits of data, the energy consumed for transmission, $E_{1,N/W}$, is fixed, whereas, the energy consumed for the propagation, $E_{2,N/W}$ varies with the radial distance of the link. The DQN algorithm is trained under the same simulation setting as that of Q-learning. The

Table G.6: Parameters used for the DQN algorithm.

Parameter	Value
Number of hidden layers	2
Number of neurons in first layer	256
Number of neurons in first layer	128
Learning rate (neural network)	0.01
Discount factor	0.4
Batch size	64
Replay memory size	50000
Number of episodes	6400
Minibatch size	64
Epsilon	1 to 0.001

policy and target network are made up of fully-connected neural networks each with two hidden layers that contain 256 and 128 neurons, respectively. The exploration-exploitation factor ε is set to vary from 1 to 0.001. Other parameters related to the DQN algorithm are provided in Table G.6. We achieved the best performance when the learning rate and discount factor were set to 0.01 and 0.4, respectively. The DQN algorithm converged after 6400 episodes resulting in the formation of the hybrid UAV-assisted wireless network with minimum network energy consumption.

Fig. G.12a depicts the topology of the hybrid random network. Here, BLE is represented with green lines, LTE with yellow lines, Wi-Fi with orange lines, and LoRa with red lines. It can be observed from the figure that the connections are evenly distributed among the four protocols irrespective of any criteria due to the random distribution. Fig. G.12b shows the UAV-assisted wireless network obtained with the rule-based algorithm. From Fig. G.12b, it can be noticed that UAVs further away from the GS with high radial distance are connected to the LoRa protocol while the ones closer to the GS are connected to BLE. Figs. G.12c and G.12d show the hybrid network obtained with the proposed Q-learning and DQN algorithms, respectively. From Fig. G.12c and G.12d, it can be noticed that a UAV-GS link is connected to a communication technology that consumes minimum energy with the number of transmissions required.

We compare the performance of the proposed algorithms with the rule-based algorithm [35] and random hybrid scheme in terms of network energy consumption and average network delay as shown in Figs. G.13a and G.13b, respectively. The Q-learning algorithm utilizes the reward defined in (G.35) for generating these results. It is observed that the Q-learning and DQN algorithms outperform the rule-based algorithm in terms of both energy consumption and average delay as shown in Figs. G.13a and G.13b, respectively. This is primarily due to the learning characteristic of RL algorithms. In RL algorithms, the environment characteristics are learned during the training or exploration phase by which an accurate estimation of the network energy consumption is obtained. However, the rule-based algorithm assigns the communication technology based on the distance irrespective of the PHY layer

characteristics. The unknown PHY layer characteristics increase the number of retransmissions required for successful transmission of bits which in turn will further increase the energy consumption and delay. The proposed RL algorithms thus exhibit an edge over other schemes in learning and incorporating additional parameters that can affect the communication link in a seamless manner. It is also observed that the random hybrid scheme consumes a large amount of energy when compared to the proposed and rule-based algorithms. This is due to the fact that the number of UAVs connected with BLE for hybrid random network configuration is higher than the RL-based hybrid network as can be observed from Figs. G.12a and G.12c. This increases energy consumption as most of these links' geographical distance is greater than r which leads to d^4 energy consumption. Since the BLE offers a higher data rate than other communication technologies, the transmission delay is much less, which lowers the average network delay as can be observed from Fig. G.13b. To provide additional clarity on the performance of all the schemes, we provide explicit values from the simulation. In the 10-th iteration, the hybrid random network exhibits a network energy consumption of 1160 Joules and the rule-based algorithm expends 190 Joules of energy. The proposed algorithms however outperform all the schemes with the minimum network energy consumption of about 151 Joules as can be seen from Fig. G.13a. In terms of average network delay, the rule-based algorithm offers an average delay of about 0.71 seconds, the random hybrid scheme offers an average delay of about 0.344 seconds, and the proposed RL algorithms exhibit 0.5 seconds at 10-th iteration as can be seen from Fig. G.13b. From Figs. G.13a and G.13b, it can be observed that the DQN algorithm displays similar performance as that of the Q-learning algorithm. This is expected as DQN essentially follows the same mathematical principles as that of Q-learning. Compared to Q-learning, the DQN algorithm utilizes a neural network to learn features from the input and provide an approximate Q-value. If the amount of data increases, storing and searching for Q-values using Q-learning can lead to performance degradation. In this scenario, DQN outperforms Q-learning in terms of reduced memory consumption and computational efficiency. However, the DQN algorithm requires significantly large training time and fine-tuning as compared to Q-learning. For example, the simulation time required to train Q-learning is 20 minutes whereas DQN requires approximately 6 hours.

G.8 Conclusion and Future Work

In this work, we have proposed two RL algorithms such as Q-learning and DQN for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network. The proposed RL algorithms take any random network as an input and learn it. Based on the learning, a hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network is formed by assigning the best communication technology to a link based on learning. We have also proposed the theoretical framework for energy-efficient data transmission over hybrid BLE/LTE/Wi-Fi/LoRa UAV-assisted wireless network for both free space and free space multipath path loss models. Further, we have derived the analytical expressions for the network energy consumption and

average network delay. Through extensive results, we verified the analytical expressions. We have also analyzed the effect of the velocity of UAVs and the number of packets on the performance of the proposed framework. Finally, it has been shown that the proposed RL algorithms result in better performance in terms of network delay and energy consumption when compared to rule-based algorithm and random hybrid scheme. In the future, we plan to incorporate the dataset obtained from physical layer parameters to evaluate the performance of the proposed algorithms.

References

[1] Dimosthenis C Tsouros, Stamatia Bibi, and Panagiotis G Sarigiannidis. A review on uav-based applications for precision agriculture. *Information*, 10(11):349, Nov. 2019.

[2] Manjula Sharma et al. Survey on unmanned aerial vehicle for Mars exploration: deployment use case. *Drones*, 6(1):4, Dec. 2021.

[3] Aakanksha Chowdhery and Kyle Jamieson. Aerial channel prediction and user scheduling in mobile drone hotspots. *IEEE/ACM Transactions on Networking*, 26(6):2679–2692, Nov. 2018.

[4] Arvind Merwaday and Ismail Guvenc. UAV assisted heterogeneous networks for public safety communications. In *Proc. IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 329–334, New Orleans, LA, USA, 2015.

[5] Yuntao Wang et al. Task offloading for post-disaster rescue in unmanned aerial vehicles networks. *IEEE/ACM Transactions on Networking*, pages 1–15, Jan. 2022.

[6] Sky Magic - Drone Light Shows. [Online]. Available: https://skymagic.show.

[7] Eyuel D. Ayele, Kallol Das, Nirvana Meratnia, and Paul J.M. Havinga. Leveraging BLE and LoRa in IoT network for wildlife monitoring system (WMS). In *Proc. IEEE World Forum on Internet of Things (WF-IoT)*, pages 342–348, Singapore, 2018.

[8] Yeduri Sreenivasa Reddy et al. Optimisation of indoor hybrid PLC/VLC/RF communication systems. *IET Communications*, 14(1):117–126, Jan. 2020.

[9] Mohamed Kashef, Mohamed Abdallah, and Naofal Al-Dhahir. Transmit power optimization for a hybrid PLC/VLC/RF communication system. *IEEE Transactions on Green Communications and Networking*, 2(1):234–245, Nov. 2018.

[10] Mengjie Yi et al. Deep reinforcement learning for fresh data collection in UAV-assisted IoT networks. In *Proc. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 716–721, Toronto, ON, Canada, 2020.

[11] Shu Fu et al. Towards energy-efficient UAV-assisted wireless networks using an artificial intelligence approach. *IEEE Wireless Communications*, pages 1–11, May 2022.

[12] Sabarish Krishna Moorthy, Maxwell Mcmanus, and Zhangyu Guan. ESN reinforcement learning for spectrum and flight control in THz-enabled drone networks. *IEEE/ACM Transactions on Networking*, 30(2):782–795, Nov. 2022.

[13] Yirga Yayeh Munaye, Rong-Terng Juang, Hsin-Piao Lin, and Getaneh Berie Tarekegn. Resource allocation for multi-UAV assisted IoT networks: A deep reinforcement learning approach. In *Proc. International Conference on Pervasive Artificial Intelligence (ICPAI)*, pages 15–22, Taipei, Taiwan, 2020.

[14] Xijian Zhong, Yan Guo, Ning Li, and Yancheng Chen. Joint optimization of relay deployment, channel allocation, and relay assignment for UAVs-aided D2D networks. *IEEE/ACM Transactions on Networking*, 28(2):804–817, Feb. 2020.

[15] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[16] S. M. Mahdi H. Daneshvar, Pardis Alikhah Ahari Mohajer, and Sayyed Majid Mazinani. Energy-efficient routing in WSN: A centralized cluster-based approach via grey wolf optimizer. *IEEE Access*, 7:170019–170031, Nov. 2019.

[17] Om Jee Pandey and Rajesh M. Hegde. Low-latency and energy-balanced data transmission over cognitive small world WSN. *IEEE Transactions on Vehicular Technology*, 67(8):7719–7733, May 2018.

[18] Bluetooth Wireless Technology. [Online]. Available: https://www.bluetooth.com/learn-about-bluetooth/tech-overview/.

[19] A. N. Wilson, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Embedded sensors, communication technologies, computing platforms and machine learning for UAVs: A review. *IEEE Sensors Journal*, 22(3):1807–1826, Feb. 2022.

[20] 3gpp Standards for the IoT. [Online]. Available: https://www.3gpp.org/news-events/3gpp-news/1805-iot_r14.

[21] 3gpp Release 13 Overview. [Online]. Available: https://www.3gpp.org/news-events/3gpp-news/1628-rel13.

[22] Evgeny Khorov, Anton Kiryanov, Andrey Lyakhov, and Giuseppe Bianchi. A tutorial on IEEE 802.11ax high efficiency WLANs. *IEEE Communications Surveys & Tutorials*, 21(1):197–216, Sep. 2019.

[23] Pycom OEM module. [Online]. Available: https://pycom.io/product/l04-oem-module/.

[24] Semtech SX1276 - Datasheet. [Online]. Available: https://www.semtech.com/products/wireless-rf/lora-transceivers/sx1276.

[25] Xianjin Xia, Yuanqing Zheng, and Tao Gu. LiteNap: Downclocking LoRa reception. *IEEE/ACM Transactions on Networking*, 29(6):2632–2645, Jul. 2021.

[26] Tom M Apostol. *Calculus, Volume 1*. John Wiley & Sons, 1991.

[27] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*. MIT press Cambridge, 1998.

[28] Tom M Mitchell and Tom M Mitchell. *Machine learning*, volume 1. McGraw-hill New York, 1997.

[29] Kai-Chu Tsai, Zirui Zhuang, Ricardo Lent, Jingyu Wang, Qi Qi, Li-Chun Wang, and Zhu Han. Tensor-based reinforcement learning for network routing. *IEEE Journal of Selected Topics in Signal Processing*, 15(3):617–629, 2021.

[30] Fan Jiang, Zesheng Gu, Changyin Sun, and Rongxin Ma. Dynamic User Pairing and Power Allocation for NOMA with Deep Reinforcement Learning. In *IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2021.

[31] Yejian Zhao, Yanhong Wang, Yuanyuan Tan, Jun Zhang, and Hongxia Yu. Dynamic Jobshop Scheduling Algorithm Based on Deep Q Network. *IEEE Access*, 9:122995–123011, 2021.

[32] Sifat Rezwan and Wooyeol Choi. Priority-Based Joint Resource Allocation With Deep Q-Learning for Heterogeneous NOMA Systems. *IEEE Access*, 9:41468–41481, 2021.

[33] Wei Ye, J. Heidemann, and D. Estrin. Medium access control with coordinated adaptive sleeping for wireless sensor networks. *IEEE/ACM Transactions on Networking*, 12(3):493–506, Jun. 2004.

[34] Juan Zhang, James F. Campbell, Donald C. Sweeney II, and Andrea C. Hupman. Energy consumption models for delivery drones: A comparison and assessment. *Transportation Research Part D: Transport and Environment*, 90:102668, 2021.

[35] H Chtourou and Mohamed Haouari. A two-stage-priority-rule-based algorithm for robust resource-constrained project scheduling. *Computers & industrial engineering*, 55(1):183–194, 2008.