Q-learning-based Optimization of Smart Home's Wireless Sensors Network Lifetime

Ismael Jrhilifa[1‡], Hamid Ouadi[1], Abdelilah Jilbab[2]

[1]ERERA, ENSAM Rabat, Mohammed V University in Rabat, Rabat, Morocco

[2]E2SN, ENSAM Rabat, Mohammed V University in Rabat, Rabat, Morocco

(Ismael.jrhilifa@gmail.com, hamidouadi3@yahoo.fr, a_jilbab@yahoo.fr)

[‡]Corresponding Author; Ismael.jrhilifa@gmail.com

Received: 26.11.2022 Accepted: 05.01.2023

Abstract- Wireless sensor networks (WSN) have witnessed increased utilization in recent years, particularly with the internet of things (IoT) trend in numerous sectors such as health, agriculture, marine, and intelligent buildings. The main challenge of these networks is energy efficiency; typically, sensor nodes are powered by tiny batteries with limited capacity. This research suggests a Q-learning-based routing algorithm for optimizing WSN lifespan and sending vast amounts of data in smart home applications. The suggested routing approach takes advantage of the benefits of Q-learning to discover the ideal routing path to transmit data with the least amount of energy dissipation. To simulate the WSN in a smart home environment, the simulation is implemented in 3D. The presented routing method is evaluated in comparison to three different protocols: QLRP, EQL, and Dijkstra. The findings show that the created routing approach surpasses the other protocols in terms of prolonging WSN lifetime, total transferred data, and energetic network cost.

Keywords WSN, smart home, lifetime optimization, reinforcement learning, Q-Learning, routing protocol energy efficiency

1. Introduction

In the last twenty years, WSN could be used and found in many areas counting intelligent homes and buildings [1,2]. WSN is a network composed of numerous devices called sensor nodes, with sensing aptitudes ranging from a few nodes to hundreds or even thousands. Those nodes are equipped with different types of sensors such as weather, pressure, electricity, motion, image, chemical, acoustic, etc [3]. Due to these sensors' diversity, WSN exploitations are on a high level, starting from agriculture, defence, medicine, and industry to smart homes and buildings.

Usually, a node of sensors is composed of four units, namely, a power unit, a sensing unit, a processing unit, and a communication unit [4]. Once WSN is installed, the power module, which consists of a small battery, should continuously supply power to the remaining three modules for months or years without any intervention [5]. The WSN use case in this paper is to control and monitor an intelligent home in a smart building. It should offer numerous measurements and information, such as the temperature, humidity, pressure, luminosity, occupants' information, electrical power consumption, security, safety, etc. The energy consumption of the sensors node is crucially important to optimize because of its small battery with short autonomy [6]. The idea behind utilizing WSN powered by batteries in a smart building context is to avoid the electrical installation complication and implement it easily in edifices already in service.

In this work, we concentrate on the routing protocol method to enhance the smart home's WSN lifetime. Routing is a procedure of choosing the path for delivering information from the source (sensor node) to the destination (sink/ base station (BS)). WSN routing could be in different forms; Direct routing to the sink means each node sends data directly to the sink; this technique is called star topology. The routing also could be hierarchical; in this case, the network is constituted of many clusters, and each cluster has a Clusterhead in charge of receiving the data sent by the nodes and transmitting them to the central node. Another strategy is called the multi-hop routing protocol; the data packet crosses several nodes to arrive at the sink. The third is the considered approach in this paper.

The availability of sensors allows for realizing many tasks in smart buildings and, when it is combined with the use of actuators, offers advantages in terms of increased occupant comfort [7]. More meaningfully, the use of actuators and sensors in a smart edifice supports environmentally friendly operation: the usage of technology permits persons to regulate energy losses, for example, by adjusting temperature or lighting without human intervention [8]. According to the network architecture, WSNs routing protocols can be classified into flat and hierarchical [9]. This research paper proposes a Q-learning (QL) based transmitting protocol to enhance a flat WSN longevity implemented in a 3D space environment. QL is a model-free reinforcement-learning algorithm approach for learning the value action in a specific state. QL produces an optimum policy that maximizes the anticipated value of the total reward across subsequent actions starting from the initial state, without the requirement for previous knowledge of the environment or the capacity to cope with random transitions and rewards without the need for adaptations. [10]. Reinforcement learning problems are considered a Markov decision process.

The rest of this paper is organized as follows: Section two will go through some similar works in the literature. The smart home's WSN concept will be presented in the third section. After that, fourthly the sensors node energy model will be introduced. Section five will be reserved to discuss the proposed routing algorithm based on QL. The sixth section is kept to exhibit and discuss the simulation results. Finally, this research item will be terminated with a conclusion.

2. Related Works

To best address the challenges of energy efficiency and WSN lifetime, numerous techniques have been explored in the literature. The following lines discuss the pertinent works:

The QL-based WSN lifetime optimization developed in [11] considered every node as an agent attempting to find the optimal way to deliver the data packet by selecting the best neighbor node from the neighbors' table. The forwarder node is chosen based on the remaining energy of the neighbor node, its hop count, and the Euclidean distance between the actual and its neighbors. However, when a node's neighbors' energy is depleted, the node becomes isolated and cannot find a way to deliver the data packet. The authors of [12] proposed another QL-based optimization. They regarded the sensors node as the state, and the data packet as an agent acting on the left, right, up, and down to reach the sink node. The drawback of this strategy is that the optimization is based solely on the count steps of the data packet, which is insufficient for optimal optimization. In [13], a QL-based energy balancing technique for WSNs was introduced. It computes the optimal pathways utilizing feedback received as a reward from routing decisions and modifies the routing table for a better selection of the next hop going forward. The residual energy and hop count to the sink are the only inputs used to compute the reward function. As a result, this technique has a flaw since, despite the node's low hop count, the distance between the node and the forwarder may be too big. In [14], the authors proposed an optimization approach based on QL with the aim of enhancing the WSN lifetime. They divided the WSN into clusters, each having a cluster head (CH). The authors supposed that the cluster heads constitute the environment and the agent is a mobile sink (MB) in action. The shortcoming of this approach is that the reward function is based only on one optimization parameter, which is the Euclidean distance between two cluster heads.

This work differs from our previous work in terms of the QL algorithm will not choose in the neighbor nodes' table the best node as a next forwarder, but it will choose the best path from a node's routing table. Additionally, it differs from the preceding listed articles in terms of the parameters or components that were selected for the optimization function: residual energy, hop count, and transmission distance. This paper's contribution appears in:

- Proposing a new reward function, which a cost function to optimize by QL with the aim of maximizing the WSN lifetime.
- Many routing paths are established using Spanning tree protocol (STP).
- The WSN contains two parts: a control unit and the QL Algorithm, which work together to recognize the optimum routing in real-time applications and detecting the death nodes, and the network part.
- Permitting both sending more packets while consuming less energy, as revealed by the residual energy using Q-learning algorithm.
- The WSN deployment is in a 3D space as well, making this study the first to address a WSN lifetime optimization problem in a 3D environment.
- A comparative study between the proposed approach and three other routing protocols.

For information, those parameters are inspired from [15]; the parameters are selected to be used in the QL algorithm, which is considered the most essential reinforcement-learning procedure to learn the less energetic data packet transmitting continuously according to the previous parameters. This technique permits the optimization of the WSN lifespan, ameliorating the energy efficacy, and reducing the network energetic bills.

3. The WSN Idea in Smart Homes

WSNs and actuators have grown in relevance in an era where technology surrounds us from all sides, including agriculture, industry, medicine, defense, and academia. Home automation is one of the technology's principal application fields. The SHWSN (Smart House Wireless Network) allows observing and controlling uses for building occupant comfort and effective home management [16]. Intelligent buildings in general, and intelligent houses in particular, contain hundreds to thousands of sensors. The sensor nodes comprise a variety of tightly limited embedded devices, some of which are battery-powered and have low-power radio frequency (RF) transceivers. Because wired solutions require conduits or cable, RF transmission permits fluid addition and exclusion of devices from the network while also lowering installation costs since wired solutions require conduits or cable trays. However, radio propagation dynamics, resource constraints, and some devices' mobility pose challenges to the design of SHWSNs.

The proposed WSN is intended to be used to remotely manage and monitor a smart dwelling that is part of a smart building. This vision provides many services, such as energy monitoring, management, and adjusting environmental conditions. Figure 1 depicts the notion of smart buildings as well as the role of sensor nodes in this technological sector.

Each sensor node can include one or more sensors, and each sensor has a specific task to realize.

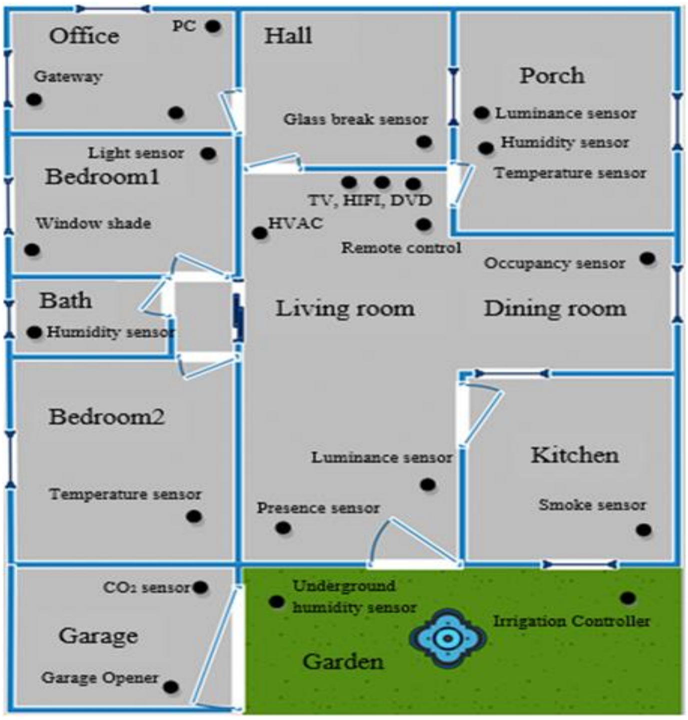

Fig. 1. Concept of SHWSN.

The following is a non-exhaustive collection of use case examples:

Light control: the lights can be controlled wirelessly, through remote control, or via a smartphone, and they can be turned on automatically when the presence and brightness sensors identify persons in the area.

Remote control: (WNs) can be exploited to operate electrical items like as air conditioners, washing machines, water heaters, and other household appliances.

Measuring of environmental measures: extra use of WN is the measurement of humidity, temperature, sound level, pressure, and so on.

Power measurement: WNs may similarly be used to measure electrical voltage and current, active and reactive power, and so on. They can also warn of electrical problems such as surges, overcurrents, etc.

4. Energy Model

We use the suggested model in [17], known as the first-order model, to characterize the energy model of WSNs. It is used to compute energy depletion in both conveying and receiving modes. When a sensor node transmits data, the energy usage in diffusion mode is as follows:

$$E_{Tx}(k,d)=E_{elec}k+\varepsilon_{amp}kd^{m} \qquad (1)$$

The energy usage in receiving mode for a sensor node when it gets a packet is determined by (2):

$$E_{Rx}(k)=E_{elec}k \qquad (2)$$

Where k stands for a packet's length, d for the distance of transmission, and $E_{Tx}(k,d)$ and $E_{Rx}(k)$ stand for the energy used to send and receive a packet with a length of k bits across a distance of d, respectively. m, E_{elec}, and ε_{amp} are three constant parameters. E_{elec} stands for the energy required for the transmitter or receiver circuitry to send or receive one unit of data, ε_{amp} refers to the energy consumed by the transmitter amplifier to transmit unit data to unit distance, and m is a propagation attenuation exponent.

The following are the parameters' values:

$$E_{elec}=50nJ/_{bit},\ \varepsilon_{amp}=100\ pJ/_{bit/_{m^2}},\ and\ m=2\ or\ 4$$

5. Proposed Approach

The suggested technique is detailed in depth in this section. The network under consideration is made up of two parts: a control unit and the QL Algorithm, which work together to identify the optimum routing in real-time applications, and the network part. The control plane is isolated from the data plane in this method. The control plane houses the controller, establishing routing patterns for nodes to follow and collecting data packets.

Initially, the control unit uses the Spanning Tree Protocol (STP) [18] to scan for all available routing pathways or the routing table (RT), preventing the network from looping. The controller then allocates each routing table to its matching node. To choose the optimal path from RT, *Q*L collaborates with the controller; it assists the controller in learning the optimum path in real-time. *Q*L, one of the most significant reinforcement-learning algorithms, is an off-policy temporal difference approach that assigns a *Q*-value to each conceivable action, indicating the quality of that action. During the learning process, the agent executes an action based on *Q*-value. The agent advances to the following state *S'* and earns a Reward *R* from the environment for each action *A* from a state, *S*. An agent's purpose is to maximize its rewards in order to encourage it to learn the optimum policy.

$$Q(s_{t+1},a_{t+1})\leftarrow(1-\alpha)Q(s_t,a_t)+\alpha[r_{t+1}+\theta_{max}Q(s_{t+1},a_{t+1})] \qquad (3)$$

The state-action function, sometimes referred to as the *Q*-function $Q(s_t,a_t)$, is introduced in Eq. 3 and is used to estimate the entire reward that an agent expects to get by doing a particular action a_t a given state s_t. At each instant $t+1$, the *Q*-values are updated, taking into consideration the previously stored value and adding the most recent reward, $r\ (t+1)$. Each agent keeps a *Q*-table with records that are |S| |A|. The optimum policy may be created when *Q*-values are learned in a static environment by selecting the action with the highest *Q*-value in each state s_t [19].

In (3), $0 \le \alpha \le 1$ denotes the learning degree. In height values of α accelerate learning and are typically reliant on the environment's dynamism. When is α= 1, the agent just utilizes the updated reward value. Mostly, α = 0.5 is used for all t. The discount factor, $0 \le \theta \le 1$, enables the agent to change how he prefers long-term rewards. When θ = 0, The agent merely takes into account immediate rewards, and when θ = 1, The weight of the instantaneous and discounted rewards is equal. [20].

A routing problem is choosing a data delivery path, which may be thought of as a Markov Decision Process (MDP). An MDP is defined as a Q-learning problem as follows:

• E: symbolizes the environment, which is WSN in this case.

• S: stands for the state variables, which include the distance between the transmitter node and the sink, the hope count *h(m)*, and the residual energy of the path *E(m)*.

• A: denotes the set of available actions for choosing a routing path.

• R: represents the reward gained from the environment as a result of doing an action.

Agent and environment are the two main parts of the QL algorithm. An agent regards the environment's current state and acts in accordance with the existing policy. The environment will reward the agent once it takes a certain action

In the proposed method, the control unit is viewed as an agent, and in the event that a node generates or receives a data packet, the controller must select the optimal routing path from the routing table and broadcast it to the node by determining the Q-value of each path in the node's routing table using equation (4).

$$Q(p)=(1-\alpha)Q(p)+\alpha\left(\sum_{k=1}^{X} R(n_k,n_{k+1})+Q(n_{k+1})\right) \quad (4)$$

Where p is the path number, α denotes the learning rate, $Q(p)$ is the expected path quality from the current node to the sink through a path. $\sum_{k=1}^{X} R(n_k,n_{k+1})$ symbolizes the expected reward from taking a path instead of another. The reward function is formulated as follows in equation (5):

$$\sum_{k=1}^{X} R(n_k,n_{k+1})=\sum_{k=1}^{X} \frac{E(n_{k+1})}{d^{\beta}(n_k,n_{k+1})h(,n_{k+1})} \quad (5)$$

X indicates the total node in the path, k is the node number, and $E\ (n_{k+1})$ is the residual energy of the next node in the path, $h(,n_{k+1})$ is the next node hop count.

$d(n_k,n_{k+1})$ denotes the 3D Euclidean distance between the consecutive nodes in a path, which is formulated as next in equation (6).

$$d(n_k,n_{k+1})=\sqrt{\Delta_x + \Delta_y + \Delta_z} \quad (6)$$

Where $\Delta_x=(x(n_{k+1})-x(n_k))^2$, $\Delta_y=(y(n_{k+1})-y(n_k))^2$, and $\Delta_z=(z(n_{k+1})-z(n_k))^2$

Where x, y, and z are the location coordinates of the nodes that belong to a path.

In addition, β is as next:

$$\beta = \left\{ \begin{array}{ll} 2 & \text{if} \quad d \le d_0 \\ 4 & \text{otherwise} \end{array} \right\}$$

$Q(n_{k+1})$ represents the path quality from a routing path's nodes n_k to the sink. It 's computed by (7).

$$Q(n_{k+1})=\sum_{k=1}^{X} \frac{E(n_{k+1})}{h(,n_{k+1})} \quad (7)$$

The controller then determines the optimum way by selecting the one with the highest Q-value, estimating the energy used to convey the data packet, and updating the remaining energy of each node along the chosen path. The controller uses the STP or Minimum Spanning Tree function to remove a node from the routing routes if its energy runs out. It repeats the procedure until the last node's energy is depleted. Lastly, the sensor nodes receive the controller's planned routing path. The same route is used to transport the data packet to the sink. The node gets delivery from a neighbor node, validates the packet's address, and passes it to the destination address by placing it in a transmission queue. Fig.2 depicts the organizational chart of the proposed approach including the two units (control unit and network side). The chart presents an operational sample on one node.

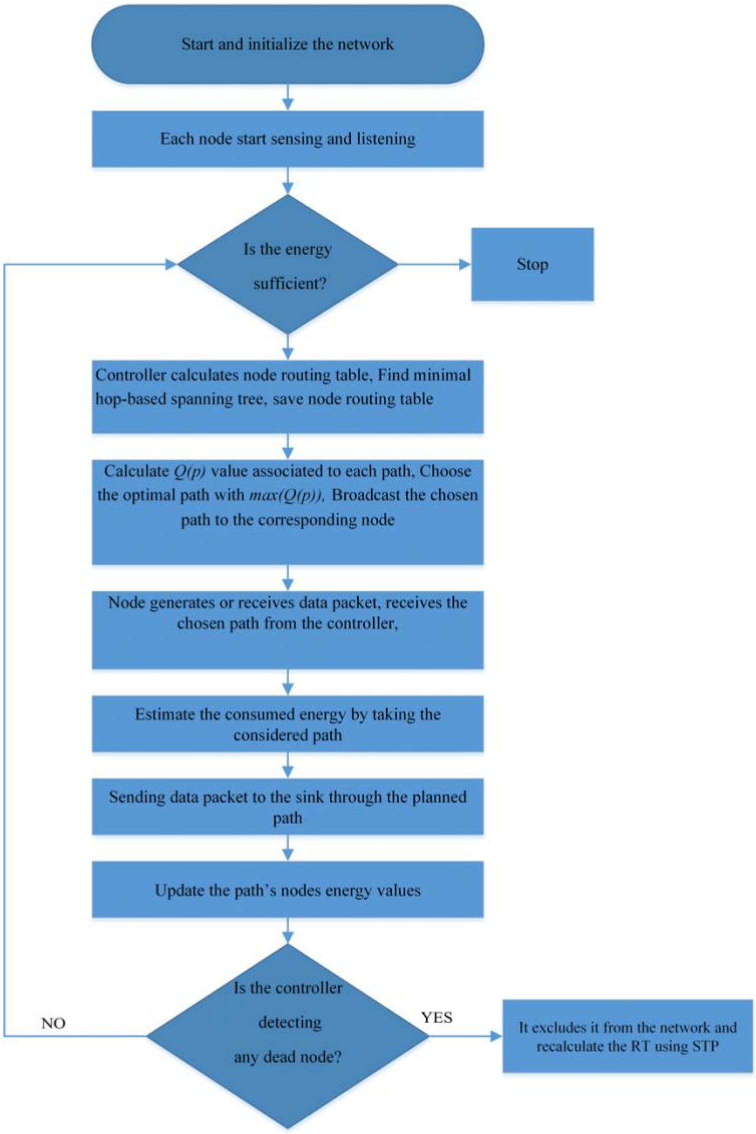

Fig. 2. The organizational chart of the proposed approach**.**

6. Simulations & Outcomes

6.1 Simulation Parameters

We exhibit the simulation results of the proposed routing protocol in this section to assess its performance in terms of network lifetime and total conveyed data. As it's noticed in figure 3 many blue circle points deployed in 3D space; each one of them symbolizes a sensors node containing a microcontroller (ESP32, Raspberry Pi, Arduino, STM, etc.…), and different sensors depending on the quantity to sense. The red triangle in the middle refers to the sink node or the base station. It's charged to collect and aggregate all the sensed and transmitted data, store them on a database such MySQL, MongoDB, etc., and display them on a monitoring application like a Web application, mobile application, and desktop software to offer visibility and the access to the homeowner. The nodes also have the option to receive commands for controlling actuators such as turning on or off appliances. The simulation is organized in several steps described in the following subsections individually.

The proposed technique's performances are compared with two other QL-based routing protocols and the Dijkstra algorithm. The first one is the proposed approach named QLRP developed in [7]. The second one is the developed QL-based routing technique called EQ proposed in [8]. The SHWSN under consideration is made up of 45 nodes including the sink node as an illustrative example. These nodes are spread out over a 3-Dimensional space of 16*13*2.5 meters (Fig.4). Each node is fixed at predetermined coordinates (x, y, and z). The starting energy of E0 is 0.5 J, and the data packet size is 512 bits because the residential data (temperature, humidity, presence, etc.) does not need a large coding size. The scales of all the figures have been standardized to facilitate comparisons. As can be shown, the suggested routing strategy outperforms the other three in terms of network lifetime optimization and total transmitted data.

The simulation is developed using Python programming language and Networkx Library, this last one is open access library developed for the creation, manipulation, and of the structure, dynamics, and functions of complex networks. Table 1 below resumes all the simulation settings and their values.

Table 1. Simulation parameters

Simulation parameters	values
Amount of nodes	44
Simulated area dimensions	16*13*2.5 m^3
The sink node	1
starting energy	0.5 J
Data amount	512 bits

6.2 Routing Protocols Behaviours in 2D Space

Firstly, the four routing protocols are developed in 2D space in order to analyze and compare their performance under identical conditions. Figure 3 depicts a comparison of four network lifetimes using three different routing techniques. The suggested technique completed almost 6000 cycles with a total network residual energy of 5 J, whereas the QLRP completed 5430 rounds, however, the network energy was fully depleted. The total number of rounds completed by the network by using the EQL protocol is 4000, with the network's residual energy equal to zero. The Dijkstra algorithm completed 3457 transmission cycles.

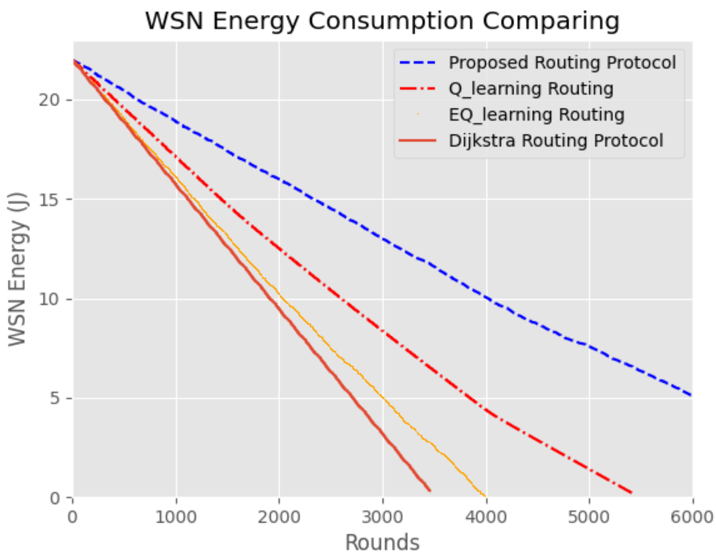

Fig. 3. Remaining energy analysis of each protocol in 2D space

In table 2 for the QRLP, we notice that the total amount of communicated data through 5430 rounds is 2780.16 Kbits, concerning the EQ method permits to transmit of about 2048 Kbits. The Dijkstra algorithm enables to send of 1769,98 kbits of data during 3457 transmission cycles. For the proposed approach, the total communicated data of 4864 Kbits can be conveyed by this protocol. As a result, the proposed approach permits to both sending more packets while consuming less energy, as revealed by the residual energy.

Table 2. Total conveyed data for every protocol in 2D space.

Routing algorithm	WSN lifetime (Rounds)	Total conveyed data amount (Kbits)
QLRP	5430	2780,16
EQ protocol	4000	2048
Proposed approach	**9500**	**4864**
Dijkstra protocol	3457	1769,98

6.3 Routing Protocols Behaviours in 3D Space

Now moving to the 3D space, to see how the four routing protocols behaved and to evaluate their performances under such conditions. According to figure 4., each routing protocol has selected its own optimal routing path to convey the data packet from node 0 to the sink. The dashed path is the selected path by the proposed routing protocol. The dash-dot line refers to the chosen path by QLRP. The dotted path denotes the preferred path by the EQ method. The solid line is the data path designated by the Dijkstra algorithm. The consumed energy by the first path is 0.0386 joule, by the second path is 0.0293 joule, by the third one is 0.0307 joule, and by the fourth one is 0.0345 joule.

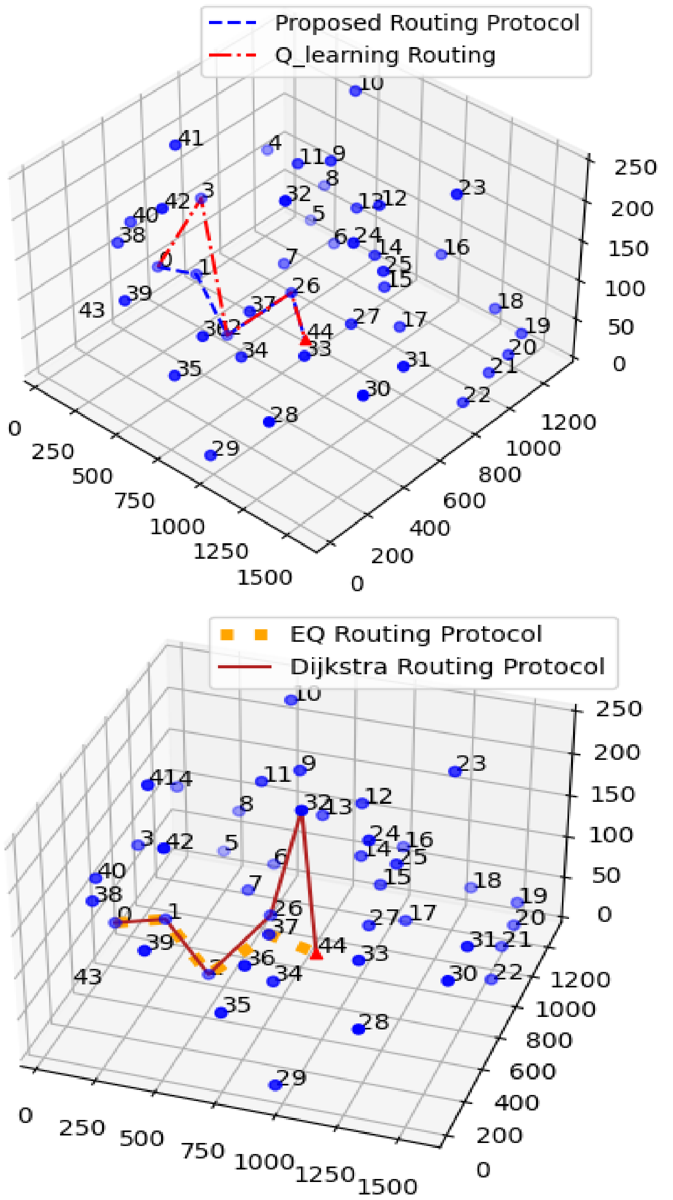

Fig. 4. The selected path to deliver the packet from node 0 to the sink using a): using Proposed Approach and QLRP. b): using EQ protocol, and Dijkstra.

As we notice, in figure 5, the total rounds achieved by SHWSN using the proposed routing technique is more than 6000 rounds and the network residual energy is 2.75 J. The QLRP network executed 4700 rounds, however, the network energy is exhausted. The network reached 3000 rounds for the protocol EQ, consuming all network energy. The Dijkstra algorithm completed 3737 transmission cycles. As a deduction, the proposed routing protocol has ameliorated the SHWSN lifetime more efficiently and permitted traffic of too much data than the three protocols.

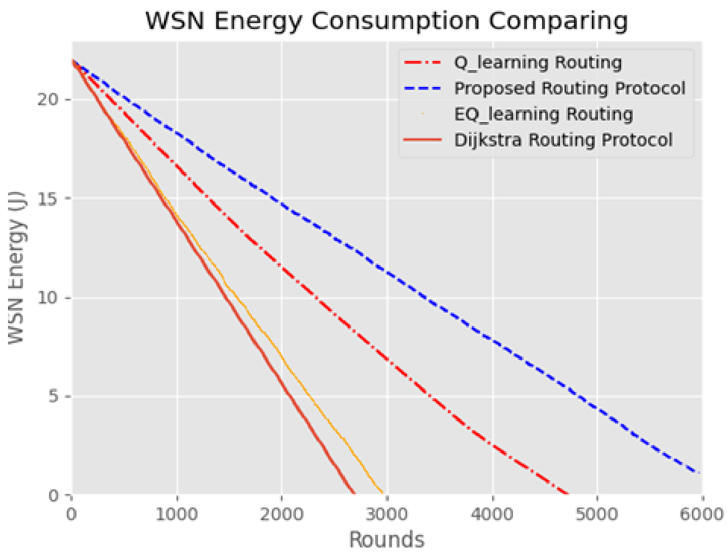

Fig. 5. Residual energy analysis of each protocol in 3D space

Table 3 resumes the total rounds accomplished by each routing strategy and the total communicated data to the sink node. the proposed technique permits sending a total data quantity of 3205.12 Kbits during 6260 rounds. And by using QLRP, SHWSN was able to send a total data quantity of 2406.4 Kbits through 4700 transmission cycles. By utilizing the EQ protocol, the network was capable to diffuse 1536 Kbit during 3000 rounds. The Dijkstra algorithm enables to send of 1401,34 Kbits of data during 2737 transmission cycles. As a result, the proposed approach permits to both sending more packets while consuming less energy, as revealed by the residual energy.

Table 3. Total conveyed data for every protocol in 3D space

Routing algorithm	**WSN lifetime(Rounds)**	**Total conveyed data amount (Kbits)**
QLRP	4700	2406.4
EQ protocol	3000	1536
Proposed approach	**6260**	**3205.12**
Dijkstra protocol	2737	1401,34

6.4 Increasing the Node Numbers in 3D Space

In order to examine more the energetic efficiency of the developed protocol, the number of sensor nodes is increased to become 54 nodes in total. To see how the network acts on the energy consumption and how its lifetime will be. Its results are compared with the QL-based routing protocols. The consumed energy for transmitting one data packet by using the proposed routing strategy is 0,0273 Joule, by QLRP is 0,0281 Joule, by EQ is 0,0292 Joule, and by Dijkstra is 0,0307 Joule.

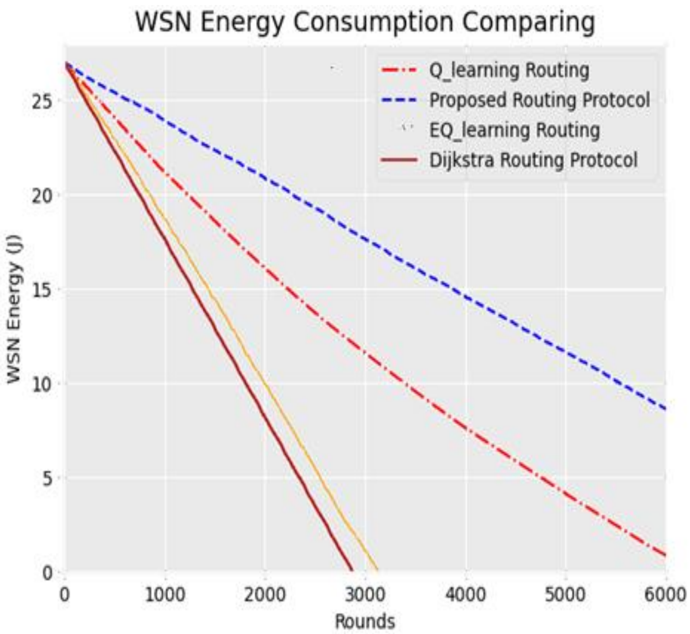

Fig.6. Residual energy analysis of each protocol in 3D space after increasing the size of the network.

As we are remarking, in figure 6, the total rounds achieved by SHWSN using the proposed routing technique is more than 6000 rounds and the network residual energy is 8J. For the QLRP, the network also accomplished 6000 rounds but the network energy is close to 1 J. For the EQ protocol, the network attained 3200 rounds consuming all the network energy. The Dijkstra algorithm completed 2877 transmission cycles. As results, the proposed routing protocol has ameliorated the SHWSN lifetime more efficiently than the other methods.

Table 4 lists the total rounds accomplished by each routing algorithm, and the total communicated data to the sink node. the proposed technique permits sending of a total data quantity of 6144 Kbits during 12000 rounds. And by using QLRP, SHWSN was able to send a total data quantity of 3220.48Kbits through 6290 transmission cycles. By utilizing the EQL protocol, the network was capable to diffuse 1638.4 Kbit during 3200 rounds. The Dijkstra algorithm enables to send of 1473 Kbits of data during 2877 transmission cycles.

Table 4. Total transmitted data for each protocol in 3D space after increasing the network size.

Routing algorithm	**WSN lifetime(Rounds)**	**Total conveyed data amount (Kbits)**
QLRP	6290	3220.48
EQL protocol	3200	1638.4
Proposed approach	**12000**	**6144**
Dijkstra protocol	2877	1473

Histograms on Fig.7 and Fig.8 describe a comparison between the four routing protocols in the three simulation experiments in terms of the total rounds accomplished by each network, and in terms of the total transmitted data to the sink. We can see clearly the impact of the proposed on the network lifetime, when the network size is increased; the lifetime was doubled into twofold comparing to the network with 45 nodes, also the total transmitted data are increased to the double. Another observation that can be drawn from the two histograms is the effect of implementation space; it is extremely evident that 3D space affects network lifetime more than 2D space (see the comparison between 45 nodes 2D and 45 nodes 3D).

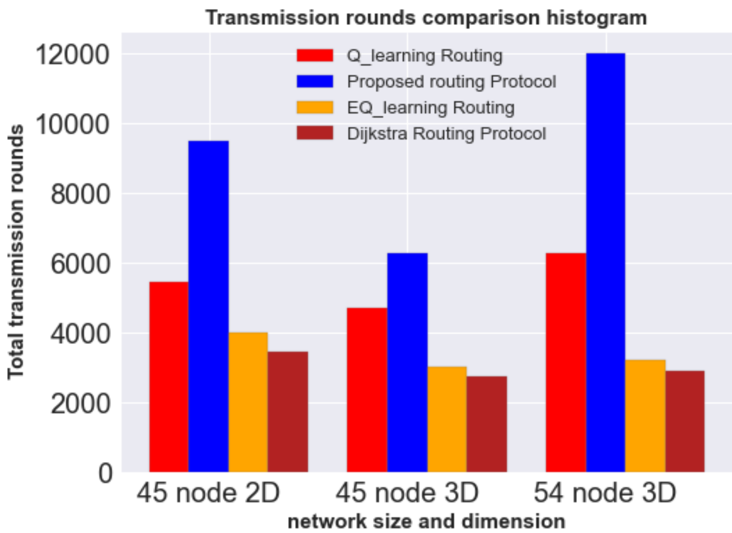

Fig.7. Networks lifetime comparison histogram

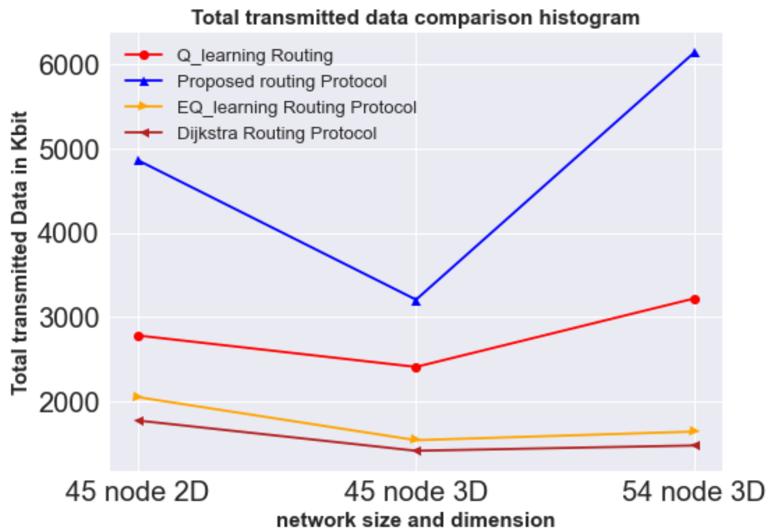

Fig. 8. Total transmitted data comparison histogram

6.5 Flexibility and Robustness Test

To inspect the flexibility and robustness of the proposed routing protocol, we have put all three protocols under a test of flexibility and robustness to see how they will perform. We canceled the energy of five random sensor nodes in five different random rounds. The results of this test are presented in the three curves in following Fig. 9. The network ran 10500 rounds of the flexibility and robustness test using the developed routing strategy. That means the WSN lifetime fell by 1500 rounds, or 14.28%. With the scale normalization, the network's remaining energy is 4,5 joule at the round 6000.The total rounds accomplished by the network by using the QLRP protocol is 5000 rounds under the same test; the network lifetime decreased with 1290 rounds regarding the normal condition which means a percentage of 20.5%. The network lifetime by using the EQ method decreased by 400 data transmission cycle, it performed a 2800, which means a percentage of 12.5 %. The network ran 2510 rounds of the same test using the Dijkstra algorithm. That means the WSN lifetime fell by 367 rounds, or 12.75%. According to this statistics, The proposed QL based routing approach reacted rapidly to such modifications and was capable to remain routing traffic competently.

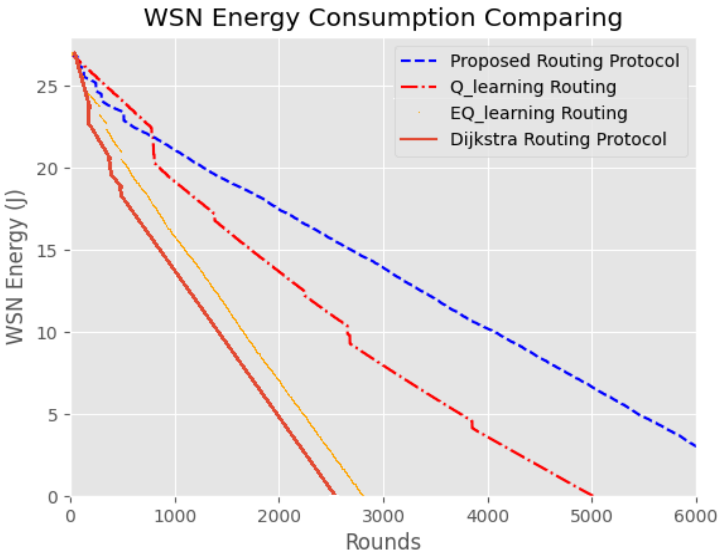

Fig. 9. Remaining energy analysis of each protocol under the test of flexibility and robustness.

7. Conclusion

Many automated operations need the usage of WSN for smart home applications. WSN lifetime is a critical performance for this. This research introduced an advanced technique, Q learning, as one of the most widely used reinforcement learning algorithms, to maximize the WSN lifetime in a smart home application. We exhibited the smart house prototype in 2D and 3D views, with a variety of network sizes that included the sensors node and the sink. We calculated the energy usage in transmission and receiving activities using the first-order Radio model. The simulation results shown that in a 2D network with 45 nodes, the suggested routing technique improved network lifespan by 42.8% when compared to the QLRP protocol, and by 57.8% when compared to the EQ algorithm. The lifetime of the WSN was increased by 63,61% when compared to the Dijkstra algorithm. For the same network in 3D perspective the suggested routing technique improved network lifespan by 24,9% when compared to the QLRP protocol, and by 52.07% when compared to the EQ algorithm. The lifetime of the WSN was increased by 56,27% when compared to the Dijkstra algorithm. Furthermore, the 3D area was protected, and the network size was increased to 54 nodes to observe how it would affect the network longevity. The proposed routing strategy increased network longevity by 47.8% when compared to the 45-sized WSN, 47.5% when compared to the QLRP protocol, and 73.3% when compared to the EQ algorithm. When compared to the Dijkstra method, the WSN's lifespan was extended by 76%. We examined the suggested algorithm for robustness and flexibility and discovered that the proposed QL-based routing strategy responded immediately to unforeseen events and therefore could continue routing traffic satisfactorily. As an outcome, the WSN energy cost will be lowered, which is considered an economic benefit. In future study, we hope to adopt another energetic model that takes into account the obstacles to energy attenuation in order to get closer to the real-world operating of WSN. Furthermore, the acquired data will be used in other research such as physical quantities prediction, smart building diagnostics, and energy consumption predictions in buildings.

References

[1] F .Viola, G .Ala, M.T. Nguyen, H.T. Tran, C.V. Nguyen, & I. Colak "Hybrid Solar-RF Energy Harvesting Mechanisms for Remote Sensing Devices." International Journal of Renewable Energy Research, 2022, doi: 10.20508/ijrer.v12i1.12807.g8403.

[2] Z. Q. Mohammed Ali and S. T. Hasson, "Simulating the Wireless Sensor Networks Coverage area," 2022 10th International Conference on Smart Grid (icSmartGrid), 2022, pp. 387-390, doi: 10.1109/icSmartGrid55722.2022.9848616.

[3] E. Irmak, A. Köse and G. Göçmen, "Simulation and ZigBee based wireless monitoring of the amount of consumed energy at smart homes," 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), 2016, pp. 1019-1023, doi: 10.1109/ICRERA.2016.7884488.

[4] R. Gill and A. N. Hasan Albaadani, "Developing a Low Cost Sensor Node using IoT Technology with Energy Harvester for Precision Agriculture," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021, pp. 187-194, doi: 10.1109/ICOSEC51865.2021.9591849.

[5] S.H. Yang, "Hardware Design for WSNs." Signals and Communication Technology, pp. 49-72, 2013, doi: 10.1007/978-1-4471-5505-8_3.

[6] M. U. Younus, M. K. Khan, M. R. Anjum, S. Afridi, Z. A. Arain, and A. A. Jamali, "Optimizing the Lifetime of Software Defined Wireless Sensor Network via Reinforcement Learning." IEEE Access, vol. 9, pp. 259-272, 2021, doi: 10.1109/access.2020.3046693.

[7] J. Milanezi, J. P. C. L. Da Costa, and E. P. De Freitas, "Improved radiofrequency energy harvesting based on a rectenna array system and its feasibility evaluation in urban environments." 2014 International Conference on Renewable Energy Research and Application (ICRERA), 2014, doi: 10.1109/icrera.2014.7016447.

[8] N. Havard, S. McGrath, C. Flanagan and C. MacNamee, "Smart Building Based on Internet of Things Technology," 2018 12th International Conference on Sensing Technology (ICST), 2018, pp. 278-281, doi: 10.1109/ICSensT.2018.8603575.

[9] A. Mehmood, J. Lloret, M. Noman, H. Song, (2015). Improvement of the Wireless Sensor Network Lifetime using LEACH with Vice-Cluster Head. Adhoc and Sensor Wireless Networks. 28(1-2):1-17. http://hdl.handle.net/10251/65889

[10] B. Jang, M. Kim, G. Harerimana, and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications." IEEE Access, vol. 7, pp. 133653-133667, 2019, doi: 10.1109/access.2019.2941229.

[11] I. Jrhilifa, H. Ouadi and A. Jilbab, "Smart home's wireless sensor networks lifetime optimizing using Q-learning," IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society, 2021, pp. 1-6, doi: 10.1109/IECON48115.2021.9589460.

[12] A. K. Yadav, P. Sharma, and R. K. Yadav, "A novel algorithm for wireless sensor network routing protocols based on reinforcement learning." International Journal of System Assurance Engineering and Management, vol. 13, no. 3, pp. 1198-1204, 2021, doi: 10.1007/s13198-021-01414-2.

[13] V. K. Mutombo, S. Y. Shin, and J. Hong, "EBR-RL." Proceedings of the 36th Annual ACM Symposium on Applied Computing, 2021, doi: 10.1145/3412841.3442063.

[14] M. Krishnan and Y. Lim, "Reinforcement learning-based dynamic routing using mobile sink for data collection in WSNs and IoT applications." Journal of Network and Computer Applications, vol. 194, p. 103223, 2021, doi: 10.1016/j.jnca.2021.103223.

[15] W. Guo, C. Yan, and T. Lu, "Optimizing the lifetime of wireless sensor networks via reinforcement-learning-based routing." International Journal of Distributed Sensor Networks, vol. 15, no. 2, p. 155014771983354, 2019, doi: 10.1177/1550147719833541.

[16] G. M. Toschi, L. B. Campos, and C. E. Cugnasca, "Home automation networks: A survey." Computer Standards & Interfaces, vol. 50, pp. 42-54, 2017, doi: 10.1016/j.csi.2016.08.008.

[17] B. Abidi, A. Jilbab, and E. H. Mohamed, "Wireless body area network for health monitoring." Journal of Medical Engineering & Technology, vol. 43, no. 2, pp. 124-132, 2019, doi: 10.1080/03091902.2019.1620354.

[18] E. Longo, A. E. Redondi, M. Cesana, A. Arcia-Moret, and P. Manzoni, "MQTT-ST: a Spanning Tree Protocol for Distributed MQTT Brokers." ICC 2020 - 2020 IEEE International Conference on Communications (ICC), 2020, doi: 10.1109/icc40277.2020.9149046.

[19] H. Tiansi and F. Yunsi, "QELAR: A Machine-Learning-Based Adaptive Routing Protocol for Energy-Efficient and Lifetime-Extended Underwater Sensor Networks." IEEE Transactions on Mobile Computing, vol. 9, no. 6, pp. 796-809, 2010, doi: 10.1109/tmc.2010.28.

[20] G. Künzel, G. P. Cainelli, I. Müller, and C. E. Pereira, "Weight Adjustments in a Routing Algorithm for Wireless Sensor and Actuator Networks Using Q-Learning ⋆ ⋆The authors thank the Federal Institute of Education, Science and Technology of Rio Grande do Sul - IFRS for financial support." IFAC-PapersOnLine, vol. 51, no. 10, pp. 58-63, 2018, doi: 10.1016/j.ifacol.2018.06.236.