applied sciences

MDPI

Article

An Intelligent Clustering-Based Routing Protocol (CRP-GR) for 5G-Based Smart Healthcare Using Game Theory and Reinforcement Learning

Abdul Ahad *, Mohammad Tahir *, Muhammad Aman Sheikh, Kazi Istiaque Ahmed and Amna Mughees

Department of Computing and Information Systems, Sunway University, Petaling Jaya 47500, Malaysia; amans@sunway.edu.my (M.A.S.); 19030576@imail.sunway.edu.my (K.I.A.); 19099621@imail.sunway.edu.my (A.M.)

* Correspondence: abdul.a22@imail.sunway.edu.my (A.A.); tahir@sunway.edu.my (M.T.)

Abstract: With advantages such as short and long transmission ranges, D2D communication, low latency, and high node density, the 5G communication standard is a strong contender for smart healthcare. Smart healthcare networks based on 5G are expected to have heterogeneous energy and mobility, requiring them to adapt to the connected environment. As a result, in 5G-based smart healthcare, building a routing protocol that optimizes energy consumption, reduces transmission delay, and extends network lifetime remains a challenge. This paper presents a clustering-based routing protocol to improve the Quality of services (QoS) and energy optimization in 5G-based smart healthcare. QoS and energy optimization are achieved by selecting an energy-efficient clustering head (CH) with the help of game theory (GT) and best multipath route selection with reinforcement learning (RL). The cluster head selection is modeled as a clustering game with a mixed strategy considering various attributes to find equilibrium conditions. The parameters such as distance between nodes, the distance between nodes and base station, the remaining energy and speed of mobility of the nodes were used for cluster head (CH) selection probability. An energy-efficient multipath routing based on reinforcement learning (RL) having (Q-learning) is proposed. The simulation result shows that our proposed clustering-based routing approach improves the QoS and energy optimization compared to existing approaches. The average performances of the proposed schemes CRP-GR and CRP-G are 78% and 71%, respectively, while the existing schemes, such as FBCFP, TEEN and LEACH have average performances of 63%, 48% and 35% accordingly.

Keywords: clustering; routing; game theory; reinforcement learning; smart healthcare; 5G and IoT

check for updates

Citation: Ahad, A.; Tahir, M.; Sheikh, M.A.; Ahmed, K.I.; Mughees, A. An Intelligent Clustering-Based Routing Protocol (CRP-GR) for 5G-Based Smart Healthcare Using Game Theory and Reinforcement Learning. *Appl. Sci.* **2021**, *11*, 9993. https://doi.org/10.3390/app11219993

Academic Editor: Paolo Renna

Received: 8 September 2021
Accepted: 10 October 2021
Published: 26 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Internet of Things (IoT) and 5G have increasingly been integrated into various facets of daily life, from smart cities to smart agriculture and from traditional to smart healthcare applications [1]. IoT and 5G-based systems enable the development of more accurate diagnostic tools, more effective treatment, and devices that improve quality of life. When IoT is used in a medical setting, it is referred to as IoMT (Internet of Medical Things). IoMT has changed the medical field by enabling remote healthcare concerning social benefits and discernment by diagnosing diseases and patients monitoring with resource-efficient methods [2,3]. By utilizing pervasive computing methods based on the IoT, it is possible to monitor and control various things of importance in the medical domain, including medical devices, physician instruction, medication, drugs, and individuals. By integrating IoT and machine learning into remote healthcare monitoring, additional efficient medical-care methods can be discovered [4–6].

Advanced healthcare empowers telemedicine, telehealth, telesurgery and telerehabilitation, which permit remote monitoring and intensive care of subjects at hospitals/home [7–10]. The modern healthcare industry requires developing a network that

integrates the human body and medical devices to form a body sensor network. Furthermore, the medical data are exchanged with the help of the IoMT framework to the medical cloud [11]. Figure 1 shows the generic architecture of smart healthcare based on 5G and IoMT. There are three major components in smart healthcare based on 5G and IoT: (a) cloud data center, (b) gateways (Base-stations) and (c) body sensor networks. IoT and 5G will play an important role in providing healthcare services to distant individuals (e.g., patients, physicians, and insurance companies). The information generated by the medical devices related to the human body can be provided to relatives and medical staff to check up on the patient anywhere, at any time.

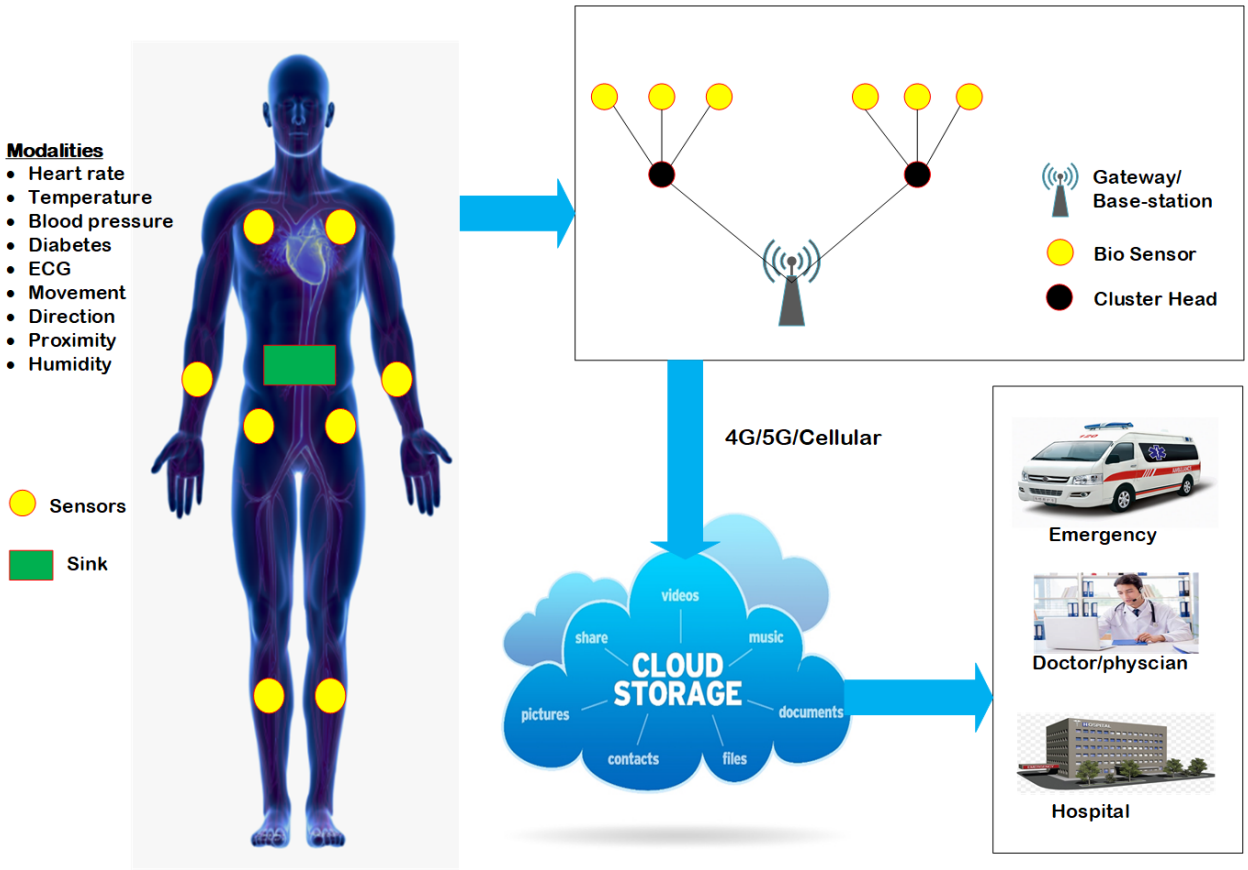


Figure 1. A general architecture of smart healthcare of IoMT.

Furthermore, in 5G and IoT-based smart healthcare, gateways are utilized as central hubs between medical devices (i.e., sensors nodes) and a cloud data center. The gateway acts as a hub to collect data and perform computation in a network for health monitoring. In addition, the gateway connects the nodes present in the network to clinic sites. These features can be successfully utilized by equipping gateways with networking, processing, and appropriate intelligence to develop smart gateways for remote healthcare monitoring.

In 5G and IoT-based smart healthcare applications, the wearable/ambient nodes are constrained in resources, including battery, processing power, and memory. Therefore, designing a framework with energy-efficient communication for medical devices to improve the network lifetime is necessary. Clustering-based routing may be a viable technique that can perform energy-efficient communication in WSN by managing various devices in the network properly [12]. It plays a crucial part to decrease the requirements of numerous nodes, which takes part in the transmission [13]. The medical nodes are organized in groups with the help of a clustering mechanism. Each group must have at least one central coordinator node designated as the Cluster Head (CH), and the remaining nodes are designated as cluster nodes (CN). All cluster nodes in the group transmit the medical information through CH. If the cluster head (CH) selection is not optimal, addi-

tional communication may occur, resulting in increased energy consumption in a smart healthcare network.

Our Contribution

Several researchers have presented a clustering-based routing scheme. However, all the existing approaches have different issues when considering 5G-based smart healthcare applications. The optimal cluster head selection and best routing path for data transmission are challenging tasks due to the high density and high heterogeneous network with different parameters such as different energy levels, mobility speeds, and low transmission ranges. The main contributions of this research are:

1. Clustering-based routing protocol using game theory and reinforcement learning to reduce energy consumption and increase network lifetime for smart healthcare scenarios.
2. An algorithm to select the best optimal cluster head (CH) from the available cluster heads (CHs) to avoid the situation of more cluster head (CH) occurrence.
3. Reinforcement learning-based route selection algorithm for data transmission.
4. Comparison of existing approaches with the proposed method.

The rest of the paper is organized as follows. In Section 2 the brief overview of existing protocols is presented. In Section 3 network model, energy model, clustering game and Theoretical analysis are presented. In Section 4 the detail of our clustering algorithm is given. In Section 5 the detail of our routing algorithm is presented. In Section 6 the time complexity of our proposed algorithms is presented. In Section 7 results analysis are discussed. Finally, the conclusion is presented in Section 8.

2. Literature Review

In recent years, IoT has been used for a wide range of applications, including smart healthcare, smart homes, and many others. Due to the limited capabilities of sensor nodes, the utilization of available resources efficiently is the main challenge for IoT-based systems. Therefore, several researchers are motivated to design energy-efficient approaches for IoT-based WSN. In [14], a joint reliable and energy-efficient technique is presented, where a game-theoretical approach is used to provide secure communication in a wireless sensor network. The overhead associated with the trust-based technique is mitigated by employing a game-theoretical approach. The results show that the presented technique is suitable for IoT-based applications in terms of security and energy efficiency. In [15], an under-water channel model is presented. The proposed method does not consider channel-aware energy saving (i.e., duty cycle) for IoT-based smart healthcare. In [16], a joint product life-cycle and IoT management algorithms are presented to save power and extend battery life. The proposed approach is based on duty-cycle management and power transmission control to optimize battery charge consumption. The researchers did not consider energy parameters, wireless channels, and clustering techniques at physical and network layers. In [17], a delay balancing approach is presented for transmission of data and energy utilization for wireless sensor networks. Besides, a work-load management policy is also considered for the network. In [18], a sensing layer-based technique is presented to analyze the power dissipation of numerous nodes in the network. Additionally, a novel architecture is presented based on base-station, sensing, and control layers. The study did not entertain duty-cycle and channel-aware energy saving. In [19], a novel framework is presented for IoT-based applications. The study focuses mainly on three layers, e.g., sensing, information processing, control, and presentation. However, they did not consider energy-saving strategies based on clustering and the use of numerous energy-optimizing parameters across the network.

In [20], a Low Energy Adaptive Clustering Hierarchy (LEACH) is proposed, known as a node selection and data transmission protocol of wireless sensor networks. In this protocol, a random technique is used to save CH energy in the network. CH selection is based on standard rules that how many times a node can be a CH. In [21], a k-mean clustering approach is presented for wireless sensor networks. In this approach, the

main issue is discovering the centroid vector, which leads to nodes group separation, guaranteeing the disconnection of connectivity. However, the benefits of these approaches give an overhead of cluster designing and selection of CH [22–24]. The overhead issues can be addressed with the help of medium access control (MAC). MAC introduces the concept of sleeping nodes when they have no data for transmission. MAC is defined by two categories: contention-free and contention-oriented [25]. These techniques initiate a part of collision when trying to use the concurrent channel. Additionally, packet loss is also increased when the network and sensor density are increased. Subsequently, the protocol that has conflict in route selection is not suitable for a dense network. Therefore, time-division multiple access (TDMA) is proposed to solve this issue [26,27]. TDMA uses a time slot schedule that assigns an individual slot to each cluster member, resulting in increased network energy efficiency. Although the clustering algorithms described in [28,29] are based on TDM, they do not take into account the potential of data failure in the network. In addition, deploying clustering involves a planning overhead, which has an impact on network resources.

In [30], an interference aware self-optimizing (IASO) is proposed, which minimizes the interference of the network. The technique has the ability of multichannel sensing and the capability of control gain. In [31], a greedy model small world (GMSW) is proposed for applications based on IoT to improve the robustness of the topological structure.

To our knowledge, no one has combined QoS and energy consumption challenges in a clustering-based routing model for 5G and IoT-based smart healthcare. Because of the variety of the network and the limited energy available to nodes, balancing QoS and energy consumption is a difficult challenge. Table 1 shows the summary of different clustering-based routing schemes.

Table 1. Review of existing clustering-based routing algorithms.

Algorithms	Latency	Energy	Reliability	Simulators	Benchmarks
LEACH	NO	YES	NO	Matlab	Direct, MTE, Static
LEACH-C	YES	YES	NO	NS-2	LEACH, Static
HEED	NO	YES	NO	Matlab	LEACH
BEEM	YES	YES	NO	Matlab	LEACH, HEED
TEEN	NO	YES	NO	NS-2	LEACH, LEACH-C
RINtraR	NO	YES	YES	Matlab	LEACH
LEACH-TL	NO	YES	NO	NS-2	LEACH
DWEHC	NO	YES	NO	NS-2	HEED
AWARE	NO	YES	YES	TOSSIM	Unaware LEACH
EECS	NO	YES	NO	Matlab	LEACH
ADAPTIVE	NO	YES	YES	Matlab	PARNET, Clustering TDMA
EEHC	NO	YES	NO	Matlab	MAX-MIN-D
HGMR	YES	YES	YES	NS-2	HRPM, GMR
MOCA	NO	YES	NO	Matlab	No Comparison
UCS	NO	YES	NO	Matlab	Equal Cluster Size
CCS	NO	YES	NO	Matlab	PEGASIS
BCDCP	NO	YES	NO	Matlab	LEACH, LEACH-C, PEGASIS
FLOC	NO	NO	NO	Matlab	No Comparison
APTEEN	NO	YES	NO	NS-2	LEACH, LEACH-C, TEEN
EEUC	NO	YES	NO	Matlab	LEACH, HEED
PEACH	NO	YES	NO	Matlab	LEACH, HEED, EEUC, PEGASIS
ACE	NO	YES	NO	Matlab	HCP
S-WEB	NO	YES	NO	Matlab	Short, Direct
PANEL	NO	YES	NO	TOSSIM	HEED
TTDD	YES	YES	YES	NS-2	No Comparison
PEGASIS	NO	YES	NO	Matlab	LEACH, Direct
MPRUC	NO	YES	NO	Matlab	HEED

3. System Model

3.1. Network Model

The proposed network model consists of base stations, sensor nodes and cluster head (CH). The end-user is connected to the base station through the internet. The sensor nodes in the network gather all the information needed and forward it to the base station through the cluster head (CH). Figure 2 shows the framework of clustering. The considered system model makes the following assumptions:

- All the sensors nodes in the network are mobile after the deployment.
- All three different categories of nodes have different energy levels. We consider pico-cells as advanced cells, femtocells as intermediate nodes, and other nodes as normal.
- All sensor nodes in the network have different mobility speeds and energy levels.
- The battery in each node in the network has a different initial energy level, and it is not rechargeable or replaceable.
- All sensor nodes in the network have their unique ID.

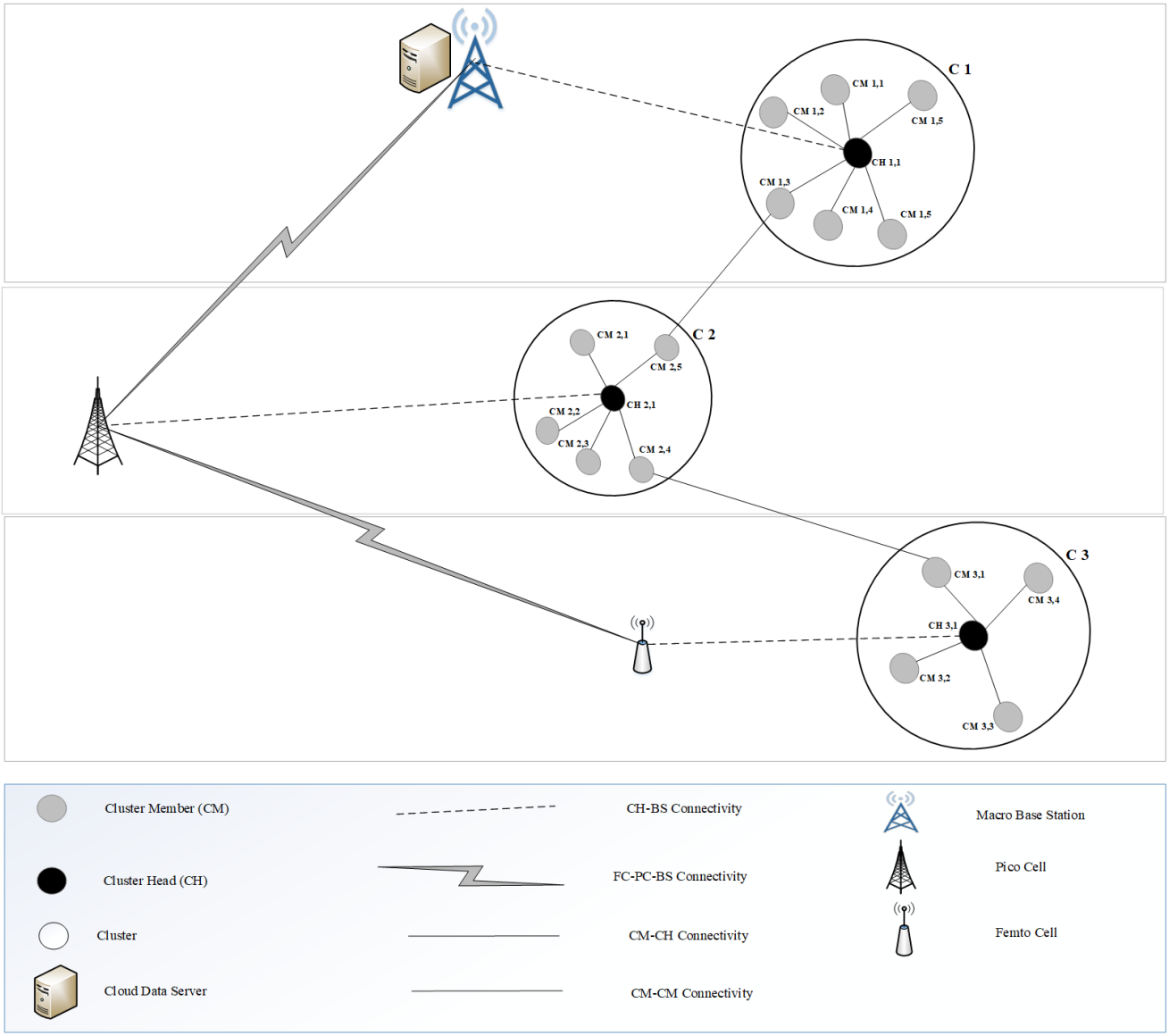


Figure 2. Cluster architecture of a 5G-based network.

3.2. Energy Model

Our proposed algorithm follows first order radio model [32–35] to handle energy dissipation as shown in the Figure 3. The distance between receiving and transmitting "*l*" bits of data is presented in the Equations (1)–(3) respectively.

$$E_{i,CH} = lE_{elec}^{Ta} + l\varepsilon_{fs}d^2, d < d_0, \quad (1)$$

$$E_{CH,sink} = lE_{elec}^{Ta} + l\varepsilon_{mp}d^4, d \geq d_0, \quad (2)$$

$$E_{Ra}(l) = lE_{elec}^{Ra}. \quad (3)$$

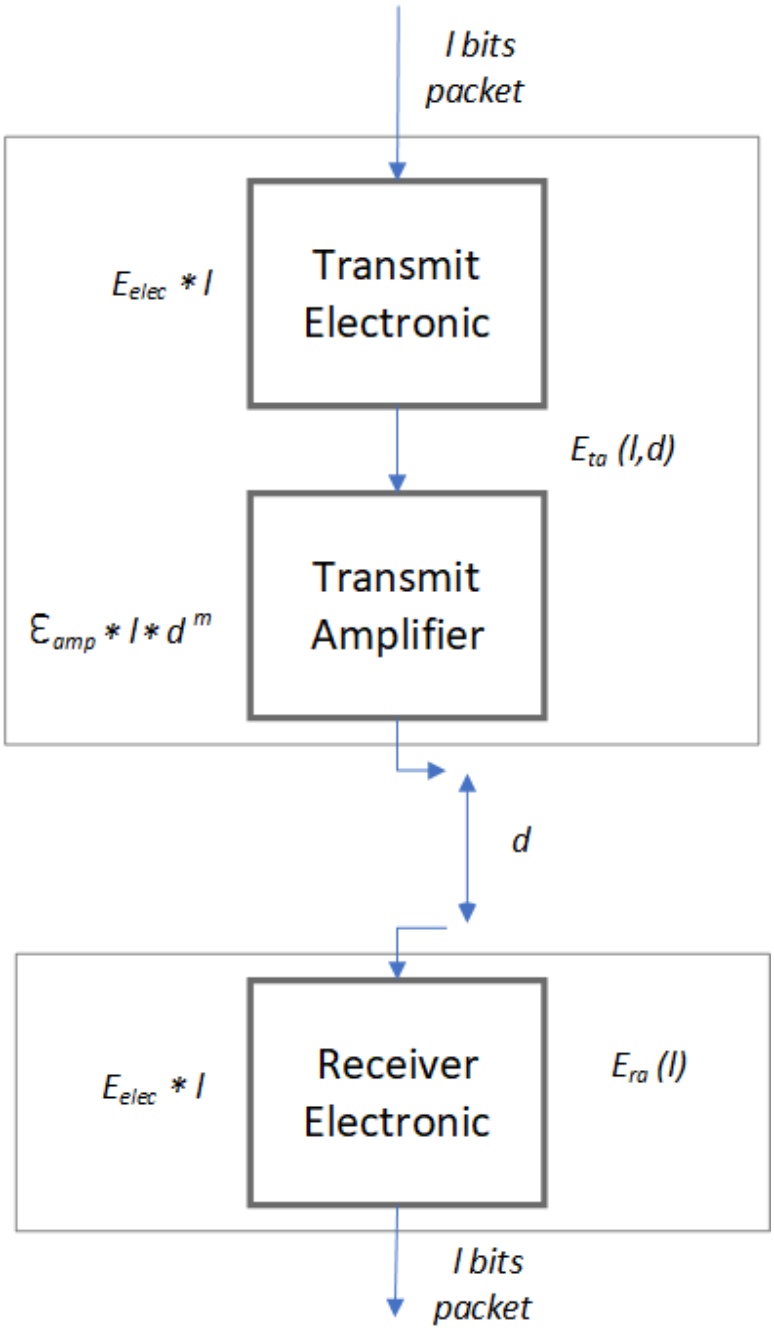


Figure 3. Dissipation model of radio energy.

E_{elec}^{Ra} and E_{elec}^{Ta} are the consumed energy via receiver and transmitter circuits per bit. Similarly, ε_{fs} and ε_{mp} are multipath and free space radio modes amplification factor. Also, threshold distance d_0 can be obtained as:

$$d_0 = (\frac{\varepsilon_{fs}}{\varepsilon_{mp}})_2^1 \tag{4}$$

Next, for the data bits "l", the aggregate energy is E_A is calculated as

$$E_A = NlE_1, \tag{5}$$

where "N" is the nodes in the cluster, "l" represent bits number and E_1 represent the aggregate energy of one bit.

The total energy consumed by a node E_c is calculated as follows:

$$E_c = E_{elec}^{Ra} + (E_{i,CH}) + E_m + (E_{CH,sink}) + E_A, \tag{6}$$

where $(E_{i,CH})$ and $(E_{CH,sink})$ represent the required transmitting energy from "i" members of the cluster to CH and from CH to sink (Base-station). E_A is the aggregate energy of a data, and E_m is mobility energy.

The distance between two nodes in the network is given by

$$d_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}, \tag{7}$$

where x_i and y_i are node "i" coordinates, while x_{i+1} and y_{i+1} are sink or neighbour node coordinates. The parameters considered for CH election are mobility speeds of nodes, the distance between nodes and the base station, and the nodes' remaining energy.

3.3. Cluster Game Modelling

The clustering game (CG) is a non-cooperative game that is used is used by the nodes to select the cluster head in the network. Every cluster in the network has a cluster head as a result of the clustering game. The cluster head collects data from the available nodes and sends it to the base station.

In the scenario where the nodes differ in any attributes or have found any difference among each other, then synchronization cannot be achieved. This leads to a cluster size of one, where each node in the network declares itself as cluster head (CH). In such an equilibrium, the expected payoff of each node of being cluster head (CH) is equal to the expected payoff of not being cluster head (CH) [36].

The non-cooperative clustering game (CG), is defined as $CG = \{N, A, U\}$, where "N" number of nodes, "A" is action set and "U" is a utility function. The proposed game for cluster head selection can be modeled with a mixed strategy game with the following elements:

- Player: N number of nodes
- Action: Cluster head (CH) and non-cluster head (NCH) are sets of actions for each player.
- Utility: Utilities of each player are denoted by the expected payoff function value, which is "0" which means no node declares as cluster head (CH).

Regarding payoffs, if none of the players (i.e., node) in the network declares itself as cluster head CH, then the payoff is zero and the player will be unable to send the data to the base station. If at least one player declares itself as cluster head CH, then the player's payoff is z (i.e., the successful delivery of data to base station). Finally, if the player declares themselves as the cluster head CH, then the payoff z is reduced by subtracting the cost c, being cluster head CH from the payoff z (i.e., $z - c$).

For the analysis of possible equilibria for two nodes case, the expected payoffs of these two nodes (2×2) are presented in Table 2. The payoff shows that the game is symmetrical and the payoff only depends on the node's strategies. The strategies $(z - c_j, z - c_j)$ (i.e.,

both the nodes declares itself as cluster head (CH)) is not a Nash equilibrium. This is because the node can get a better payoff if they change their strategies to Non-cluster head NCH (i.e., $z > z - c$). Similarly, the strategies (0, 0) do not follow a Nash equilibrium either because any node will prefer to deviate and declare itself as cluster head CH, leading to a positive payoff. The remaining conditions such as $(z - c_j, z)$ and $(z, z - c_j)$ are Nash equilibria (i.e., one of the nodes declares itself as cluster head CH and another declares itself as non-cluster head NCH such that none of the nodes have any incentive to change its strategy). The utility of function of this game is an optimal selection of energy-efficient cluster head (CH), which is given as U_{CG} for node "i" as follows:

$$U_{CG} = \begin{cases} 0, & if \quad S_{i,j} = NCH \\ z, & if \quad S_j = CH \quad and \quad \exists \quad i \in N, \quad S_i = NCH \\ z - c_j, & if \quad S_i = CH \quad s.t, z > c > 0 \end{cases} \tag{8}$$

Table 2. The (2 × 2) players payoff.

	Player 2	
Player 1	**Cluster Head (CH)**	**Non-Cluster Head (NCH)**
CH	$z - c_j, z - c_j$	$z, z - c_j$
NCH	$z - c_j, z$	*0, 0*

3.4. Expected Payoff

To reach the equilibrium, the nodes play mixed strategies. This means any node has permission to declare itself as cluster head (CH) with probability p and non-cluster head (NCH) with probability $p = 1 - q$.

Theorem 1. *The mixed strategies Nash equilibrium exists for the symmetrical clustering game, and the probability p for the player as cluster head (CH) in the equilibrium is given as*

$$p = 1 - (\frac{c}{z})^{\frac{1}{N-1}}.$$

Proof. To find out the Nash equilibrium in mixed strategies, which corresponds to the probability p of a node as cluster head (CH) by using the methodology as presented in the [36]. First of all, we need to find out the expected payoff for every available choice. When the node act as a cluster head (CH), the expected payoff is $U_{CH} = z - c$, which is independent of other node strategies. When the node is playing non-cluster head (NCH) then the expected payoff is

$$U_{NCH} = Pr\{no\ player\ declare\ as\ CH\} \cdot 0 + Pr\{at\ least\ one\ declare\ as\ CH\} \cdot z$$

$$= z \cdot (1 - q^{N-1})$$

$$= z \cdot (1 - (1 - p)^{N-1})$$

□

The payoffs are equal in the equilibrium. Therefore, no player has the incentive to change its strategy. Thus,

$$z - c = z \cdot (1 - (1 - p)^{N-1}).$$

We can compute the probability p by solving the above equation, which corresponds to the equilibrium.

$$p = 1 - (\frac{c}{z})^{\frac{1}{N-1}}.$$

Let us denote $\omega = \frac{c}{z} < 1$. Figure 4 shows the probability values as the nodes increases for different ω parameter values (i.e., 0.05, 0.1, 0.3, 0.5, 0.7 and 0.9). As the number of players increases, the probability "p" increases. When the attributes of the nodes (i.e., mobility speed, energy level and distance) are similar in a network, then the equilibrium condition is given by the probability as:

$$For N = 1, p = 1 - (\frac{c}{z})^{\frac{1}{N-1}} = 1. \quad (9)$$

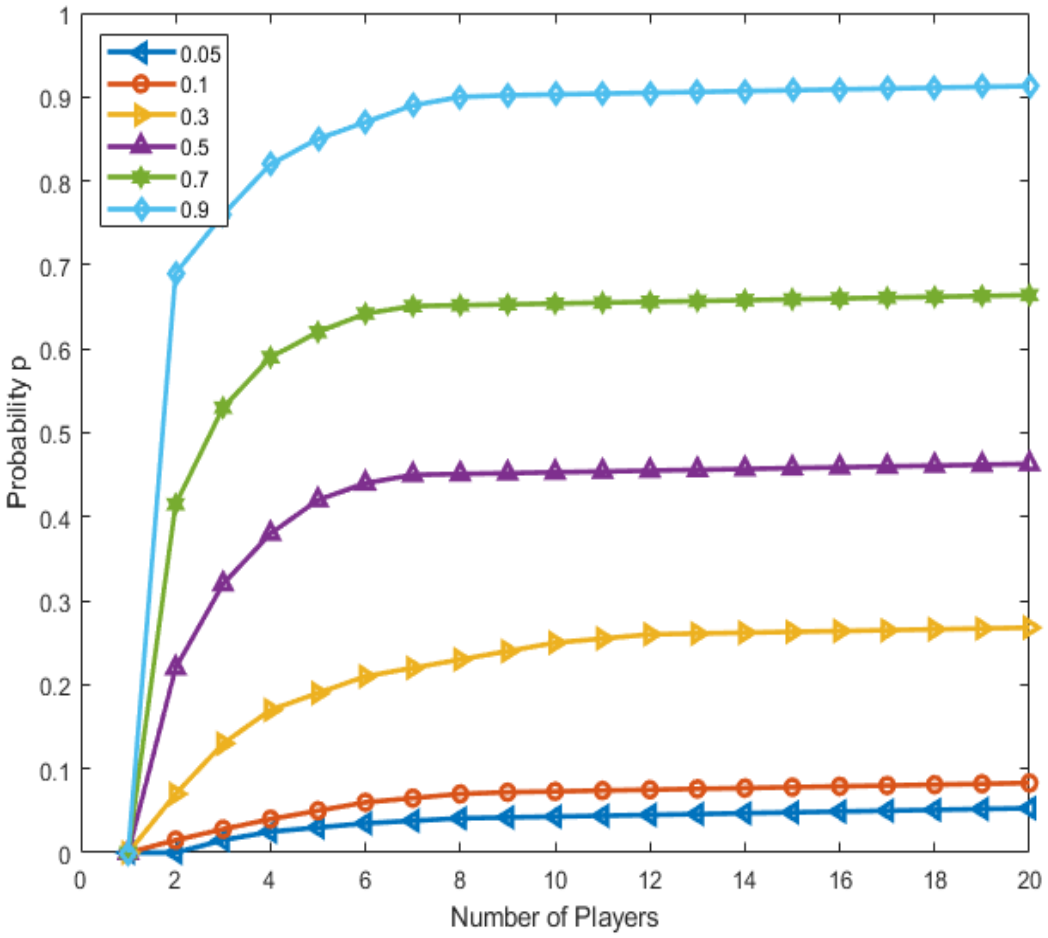


Figure 4. Probability p of a player for being a cluster head (CH) versus number of players for different 'ω' values.

The probability of being cluster head (CH) is always 1 if $N = 1$, i.e., that at least one node must play as cluster head (CH).

$$For' N' nodes \Rightarrow p = 1 - (\frac{c}{z})^{\frac{N}{N-1}}. \quad (10)$$

The average payoff of an arbitrary node "i" is specified as

$$P = (z - c) \cdot Pr\{S_i = CH\} + z \cdot Pr\Big\{S_i = NCH \bigcap \exists j s.t S_j = CH, j \neq i\Big\}$$

$$= (z-c)\cdot Pr\{S_i = CH\} + z\cdot Pr\{S_i = NCH\}\cdot Pr\{\exists js.tS_j = CH, j \neq i\}$$

$$= (z-c)\cdot p + z\cdot(1-p)\cdot(1-Pr\{S_j = CH, j \ni N, j \neq i\})$$

$$= (z-c)p + z(1-p)(1-(1-p)^{N-1})$$

$$P = z - c_j p - z(1-p)^N. \tag{11}$$

For the equilibrium strategy, the average payoff P_{NE} is given as

$$P_{NE} = z - c_j p - z(1-p)^N$$

$$= z - c\left(1-\left(\frac{c}{z}\right)^{\frac{1}{N-1}}\right) - z\left(1-\left(1-\left(\frac{c}{z}\right)^{\frac{1}{N-1}}\right)\right)^N$$

$$= z - c + c\left(\frac{c}{z}\right)^{\frac{1}{N-1}} - z\left(\frac{c}{z}\right)^{\frac{N}{N-1}}$$

$$P_{NE} = z - c_j. \tag{12}$$

To understand the scenario with nodes consisting of different attributes (i.e., mobility speed, energy level and distance), consider two players' games as shown in Table 2. The probability of a node being a cluster head (CH) for two nodes (n_1, n_2) with mixed strategy equilibrium is calculated as follows: let us suppose P_i is the probability of every node in the network with the cost c_i, for $i = 1, 2, 3, 4, \ldots n$. The probability of n_1 node being a CH with cost c_1 is derived as follows on the basis of Equation (11) :

$$z - c(p) = z\cdot\left(1-(1-p)^{N-1}\right)$$

$$c(p) = z(1-p)^{N-1}$$

$$\left(\frac{c_2}{p}\right) + (c_1) = z(1-p)^{N-1}$$

$$\left(\frac{c_2}{z}\right) + \left(\frac{c_1}{z}\right)p = p(1-p)^{N-1}$$

$$P_1 = 1 - \left(\frac{c_1}{z} + \frac{c_2}{z}\right)^{\frac{1}{N-1}}, \tag{13}$$

where c_1 and c_2 are the costs to be a cluster head (CH). The probability for "N" nodes is specified as

$$\sum_{i=1}^{i=n} P_i = 1 - \left(\frac{c_i}{z} + \frac{c_{i\pm1}}{z}\right)^{\frac{N}{N-1}}. \tag{14}$$

From the above equations, it can be seen that if nodes with differing attributes are found in the network, then every optimal probability depends on the costs of the others, i.e., being a cluster head (CH) costs more.

Being a cluster head (CH) in Nash equilibrium always depends on the neighbor node "j" cost. If "N" increases, probability becomes less, which means one node at least declares itself as cluster head (CH). Likewise, when "N" tends to 1, then "p" tends to 1, and then it constantly declares itself as a cluster head (CH).

$$\lim_{N\to1} p_n = 1. \tag{15}$$

The cost of being a cluster head (CH) is specified as:

$$c_n = (D + M_s)/E_{pr} \tag{16}$$

$$E_{pr} = E_{int}(i) - E_c(i).$$

$E_{int}(i)$ is the initial energy of node "i" and E_c is the energy consumed by node "i" in the data transmission to base station. "D" is distance between nodes and M_s is the mobility speed of node.

From the analysis of the clustering game, it is found that if the benefit and cost are not dependent on "N" value. As the probability decreases, at least one node must be declared as cluster head (CH) in equilibrium. Therefore, our proposed algorithm claims one cluster head (CH) per cluster, avoiding additional competition for cluster head (CH) selection.

4. Clustering Algorithm

In this section, we introduce a clustering algorithm based on game theory for 5G-based smart healthcare.

4.1. Initialization

Our proposed protocol consists of two phases, i.e., the setup phase and the steady-state phase. Cluster head (CH) selection and cluster formation are performed in the setup phase, while data transmission is performed in the steady-state phase. Initially, the number of members and an optimal number of clusters is obtained for the given "N" value. Then, each node in the network broadcasts a message to the neighbors for the nomination of the cluster head (CH). Finally, all the information is collected by the base station and saved.

4.2. Setup Phase

Probability P_k of node "i" in the "k" cluster is calculated as in the Equation (14). The probability of the node is compared with other cluster nodes, and the node with the highest probability is selected as cluster head (CH). Once the election is complete, the cluster head (CH) broadcasts the "CH message" in the network along with node ID and "Join-Request" set to "0". In the response, nodes send back the "Join-Response" field set to "1" to cluster head (CH) along with the information <node position, remaining energy, speed, Node ID> and declare itself as cluster member of that cluster. Although the nodes may join the nearby cluster as a cluster member and withdraw the nomination when they receive the "CH message". Hence, a cluster in the network is formed.

4.3. Steady-State Phase

When a node identifies an event in the network, the node transmits data to the cluster head (CH), and the cluster head (CH) transmits the data to the base station. In our proposed algorithm, reinforcement learning helps the cluster head (CH) and nodes to find the energy-efficient route for data transmission to the base station. It prevents the early death of cluster heads (CHs) and reduces the traffic in the network. Due to nodes, mobility in the network would change the topology, leading to the deletion and addition of member nodes, i.e., to hand over the control of the mobile node to another cluster head (CH) immediately. Whenever the node member goes out from the cluster region, the sink (base-station) estimates the new position of the node and is assigned as a member node to a new cluster. The proposed algorithm pseudo-code is given in Algorithm 1.

Algorithm 1: Clustering

Require: N: Number of nodes.
n_1, n_2 : Nodes among N.
U_i:Node utility function.
$p(n_1)$: n_1 node probability to become a cluster head(CH).
Count : Total number of times as to be a cluster head (CH).
C_k: Number of optimal clusters.
Ensure: CH(n_1): Cluster Head selection of the node n_1.
Functions:
Broadcast (Distance, Data);
Send (Data, Destinations);
Probability (c_n, z);
% *Initialization*
$U_i(n_1) \leftarrow$ n(CH, NCH).
is cluster head (n_1)= false;
r $\leftarrow$ 0.
MAIN:
For each round of clustering.
if *(count*==$\frac{C_k}{N}$*)* **then**
    is cluster head (n_1) $\leftarrow$false;
    p $\leftarrow$ (0,1);
else
    count $\leftarrow \frac{N}{C_k}$;
end
p $\leftarrow$ probability that is (c_n, z);
% Defined by Equation (14).
(node, p) broadcast i.e., message of candidate.
On the delivery of message candidate from CH candidates.
for *the every* $n \in N$ **do**
    if $p_(n_1) > p_(n_2)$ **then**
        CH $n_1 \leftarrow U_i\ (n_1)$
        is cluster head $(n_1) \leftarrow$ true;
        Broadcast (Stop Election-Message, N)
    else
        is cluster head $(n_1) \leftarrow$ false;
        CH $(n_2) \leftarrow U_i\ (n_2)$
    end
    if *cluster head* n_1 ==*true* **then**
        Broadcast (Cluster Head Message, N)
        Send Join Request Message;
    else
        On Getting CH-Message:
        Node Send Join Response Message to close Cluster Head (CH).
    end
    Count=count+1;
end
if *mobility of node* == *true* **then**
    forward Dis-Join Request Message to CH;
    forward Join Request Message to close CH;
else
    mobility of CH== true;
    forward Join Request Message to close CH;
    Re-election.
end

5. Routing Algorithm

In this section, we introduce a reinforcement learning-based routing algorithm for 5G-based smart healthcare. Figure 5 shows the data processing in the network.

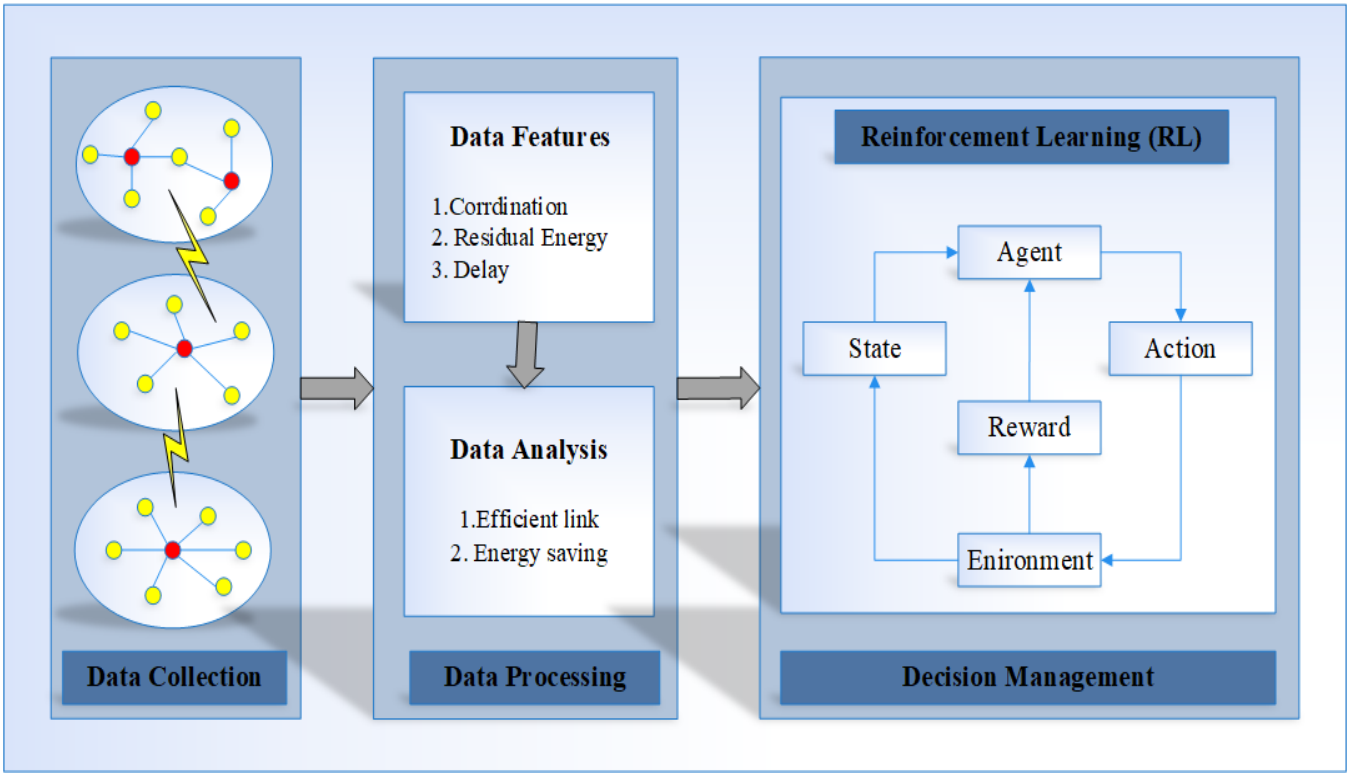


Figure 5. Data processing framework in 5G-based smart healthcare.

5.1. Q-Learning

Q-learning is a well-known reinforcement learning technique that selects the optimal action based on the current state and receives the delayed reward with the highest value without using a specific environment model [37]. Based on inputs, three primary factors are measured within the Q-learning.

s: State: i.e., energy level and node position.

v: Action: i.e., accessibility of the available next hope node.

R: Reward: i.e., successful data transmission, which is calculated by function reward.

Whenever an agent performs an action a, it is awarded with a reward, R, instantly. Hence, we can utilize a set of factors to illustrate the working procedure: $(s_0; v_0; R_1; s_1; a_1; R_2; s_2; \ldots\ldots)$, which shows that agent get a reward R_1 and changing from state s_0 to state s_1 with the action v_0; when an agent reached to state s_2 with the action v_2, it get the reward R_2 and so on. Furthermore, $Q(s_t; v_t)$ is called Q-value which is the actual value for the state-action pairs. To show the relationship between Q-values and reward, the following formula can be use.

$$Q : s \times v \rightarrow R. \tag{17}$$

Additionally, Q-values are updated by the core algorithm by calculating old and new information with by using Equation (18), given below.

$$Q_{t+1}(s_t, v_t) = (1 - \alpha) \cdot Q_t(s_t, v_t) + \alpha \cdot [\gamma \cdot maxQ_t(s_{t+1}, v_t) + R_{t+1}]. \tag{18}$$

In Equation (18), α is the learning rate which is $(0 \leq \alpha \leq 1)$. (s_{t+1}) is available next hopping nodes, γ is a discount factor which is $(0 \leq \gamma \leq 1)$, $maxQ_t(s_{t+1}, v_t)$ shows the most extreme approximation reward possible. Whenever the value is set to 0, new information

is irrelevant, and previous information is only considered. While, when α is closed to 1, only the new information is considered and past information is discarded. Furthermore, the γ factor is vital for upcoming rewards. Whenever the value is set to 0, then the agent considers short-term rewards. When the value is set to 1, then the agent is interested in long-term rewards.

$$E_{hop} = \frac{E_t}{E_{(t+1)} + 1/E_{pr}} \tag{19}$$

$$D_{hop} = \frac{D_t}{D_{(t+1)} + d_t}, \tag{20}$$

where D_t defines the shortest best path, E_t defines the highest energy best path, E_{pr} shows the energy remaining in the N_t node and d_t is the distance between nodes.

5.2. Our Proposed Routing Algorithm

Different routing algorithms have been proposed in the literature to achieve Quality of Services (QoS), energy consumption and link heterogeneity in WSN. Some of the famous algorithms are E-TORA (energy-aware TORA), EBCRP (energy balanced chain-cluster routing protocol), HGMR (hierarchical geographic multicast routing), EADAT (energy-aware data aggregation protocol) and many more, as mentioned above in Table 1. Due to different attributes, such as dynamic topology, distributed nature, high density and resource constraints of the 5G-based smart healthcare network, the above-mentioned cluster-based routing protocols are not suitable for this network. The earlier algorithms do not consider the delay in transmission, energy consumption, and unbalanced energy dissipation. To address these issues, we use a Q-learning-based algorithm. The network is divided into three grids (i.e., pico, femto and macro). In the learning phase, the grids periodically exchange information (i.e., the distance between nodes and the nodes' remaining energy). The value of each node state and topological relation of nodes with each other are stored in a Q-table. The agent makes a decision based on the Q-table for the next-hop selection. In this paper, we classified them based on Energy hop (E_{hop}) and Distance hop (D_{hop}), which improve the Quality of Services (QoS) and reduce energy consumption by maintaining energy balance between nodes. Furthermore, the values can vary considerably between the various range. Therefore, it is critical to scale them in a specific range. For the Y feature, we can scale it between range [0,1] as follows:

$$f(Y) = \frac{Y - min(Y)}{max(Y) - min(Y)}. \tag{21}$$

In Q-learning based routing protocol, Q_t and R_{t+1} at the hop can be calculated with Equations (23) and (22) which are associated with E_{hop} and D_{hop}. At this point, when the agent changes the state from s_t to s_{t+1}, which means data are transmitted from N_t to N_{t+1}.

$$Q_t(s_t, a_t) = f(E_{hop}) \cdot (1 - \beta) \tag{22}$$

$$R_{t+1} = f(D_{hop}) \cdot \beta. \tag{23}$$

Additionally, $\beta (0 \leq \beta \leq 1)$ is an administrative factor that imitates the significance of delay and energy consumption. When the β factor is set to value 1, the algorithm emphasizes decreasing transmission delay. While, when the β value is set to 0, the algorithm tries to balance the remaining energy. Otherwise, both transmission delay and remaining energy are considered instantaneously. In this way, we can alter the value of β according to QoS requirements.

By combining Equations (18), (22) and (23), the algorithm for routing is updated as follows:

$$Q_{t+1}(s_t, v_t) = f(E_{hop}) \cdot (1 - \beta) \cdot (1 - \alpha) + \alpha \cdot [\gamma \cdot maxQ_t(s_{t+1}, v_t) + \beta \cdot f(D_{hop})]. \tag{24}$$

The sink is the end node of the network, which is the information collection unit within the network. Every sink has the next hopping node, known as BS. Thus, $maxQ_t(s_{t+1}, v_t) = 0$. In this situation, the formula is updated as follows:

$$Q_{t+1}(s_t, v_t) = (1 - \beta) \cdot (1 - \alpha) \cdot f(E_{hop}) + \alpha \cdot f(D_{hop}) \cdot \beta \quad (25)$$

Algorithm 2 presents a detailed explanation of the proposed routing algorithm. Each node keeps information of the network and a Q-value. Before information gathering, the base station allows tasks with broadcasting. All nodes in the network update their Q-value according to Equation (24). Based on the updated table of Q-values, the path with the highest reward in Algorithm 3 is selected. The node determines the next available hopping node that has the highest Q-value from the available subsequent hopping nodes if there are subsequent hop nodes available for a node, implying that the node is in the vicinity of the base station. Otherwise, the node will select a node with the highest energy as the next hopping node from all nodes available in the reachable range. Figure 6 shows a Q-value fluctuation with several rounds. The Q-value fluctuates in the initial couple of rounds, but it converges after a certain number of rounds. Figure 7 shows that link disconnection is lower at learning rate of α = 1.0. It is due to higher network density because more nodes in the network increase the number of hops.

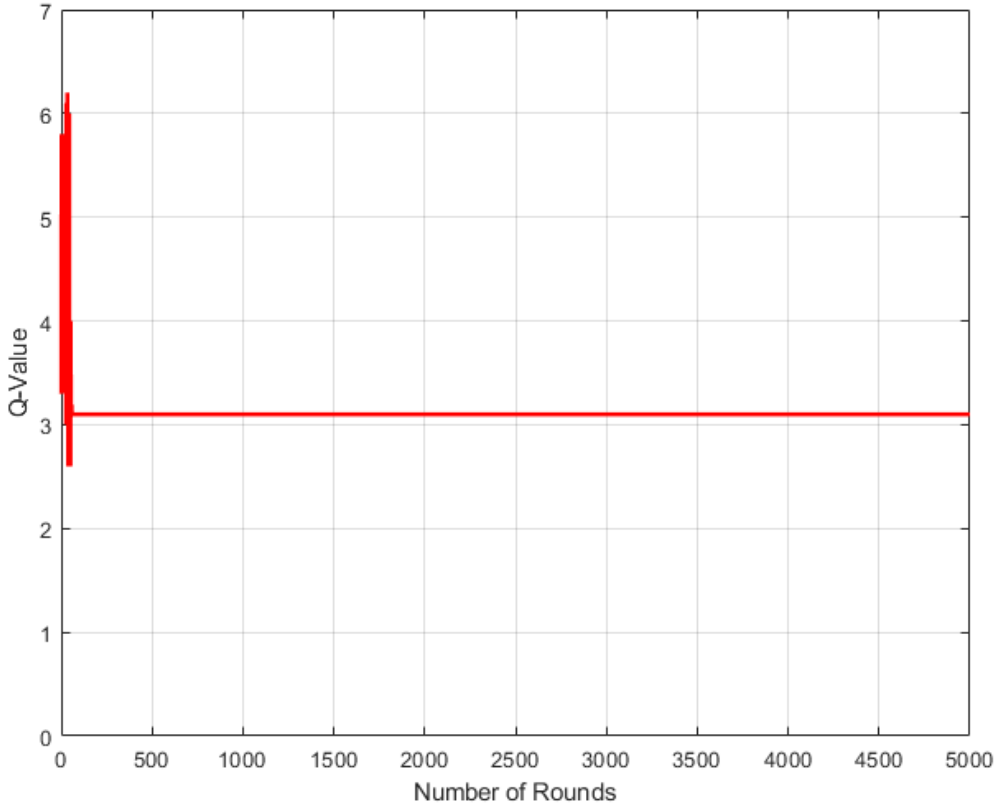


Figure 6. Fluctuation of the Q-value and coverage with respect to the number of rounds with a learning rate (α = 0.9).

Algorithm 2: QL-Algorithm

Require: Information of nodes
Ensure: *Q-table*
% *Initialization*;
Zeros-matrix $\longrightarrow$ *Q-table*;
Distance between nodes $\longrightarrow d_t$;
Maximum communication distance $\longrightarrow C_{max}$;
Energy remaining of nodes $\longrightarrow E_{pr}$;
% Run clustering algorithm;
while *CHs sets are not empty* **do**
  for *every* SN_i **do**
    % *computation of next hopping node;*
    Its coordinate $\longrightarrow y_t(i)$;
    for *every* SN_i **do**
      Its coordinate $\longrightarrow y_{t+1}(j)$;
      if $d_t(i,j) < C_{max}$ *and* $y_t(i) < y_{t+1}(j)$ **then**
        The available next hopping node ($SN_i \longrightarrow SN_j$);
      end
    end
    while SN_i *ensure the available next hopping node* **do**
      % *calculate the rewards and Q-table;*
      compute E_{hop} and D_{hop} using Equations (19) and (20);
      $Q_t(s_t, v_t) = f(E_{hop}) \cdot (1-\beta)$;
      $R_{t+1} = f(D_{hop}) \cdot \beta$;
    end
  end
  % *Q-table update accordingly ;*
  for *every round of number* **do**
    select a state s_t;
    if SN_i=*0* **then**
      $Q_{t+1}(s_t, v_t) = (1-\alpha) \cdot Q_t(s_t, v_t) + \alpha \cdot [\gamma \cdot maxQ_t(s_{t+1}, v_t) + R_{t+1}]$
    end
    $s_t \leftarrow s_{t+1}$;
  end
end

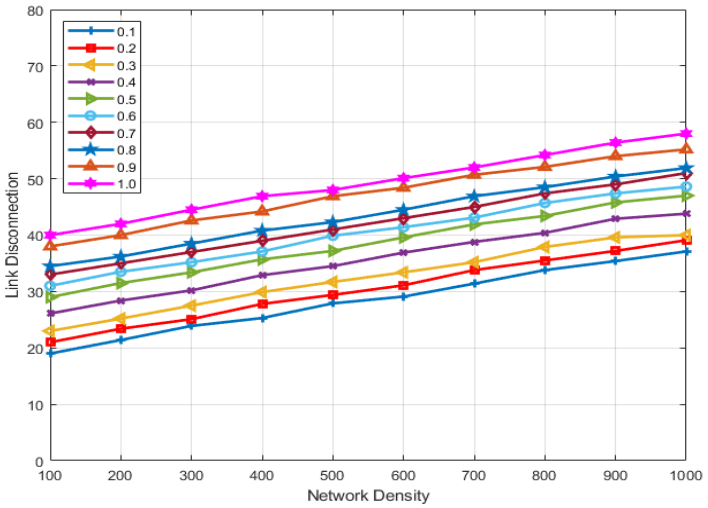


Figure 7. Link disconnection percentage increase with the increase of network density. The learning rate α = 1.0 shows the low number of link disconnection.

Algorithm 3: Best Path Selection Scheme

Require: Information of node
Ensure: Best path selection
% *Initialization*
All nodes update BS about their location and remaining energy.
On the basis of algorithm 1 Base station form clusters in the network and calculate the Q-table.
% *Data Sending*
for *every source node* **do**
    % Select route for data forwarding
    Chose state s_t as current node;
    while *if the data not reached to Base station* **do**
        if *Next hopping nodes are available* **then**
            Chose next state s_{t+1} with highest Q-value;
        else
            Chose the nearest node having the highest energy remaining as next state s_{t+1};
        end
        Forward node states from s_t to s_{t+1} along with collected information.
    end
    $s_t \leftarrow s_{t+1}$;
    % *calculate energy consumption*
    By using Equation (6) update the remaining energy;
end

6. Time Complexity of Proposed Algorithms

In this section, an analysis of the time complexity of the proposed algorithms is presented.

6.1. Time Complexity Clustering Algorithm.

The symbols used for calculation is given below:
N: Number of nodes.
n_1, n_2 : Nodes among N.
U_i: Node utility function.
$p(n_1)$: n_1 node probability to become a cluster head (CH).
$Count$: Total number of times as to be a cluster head (CH).
C_k: Number of optimal clusters.

The time complexity can be calculated step-by-step and then combined to show the algorithm's overall complexity.

Solution:

- Use the if-else statement from lines 18–23 to check count = $\frac{C_k}{N}$ and count $\leftarrow \frac{N}{C_k}$, taking a constant time due independent of data length. We denote this constant with C_1.
- $p \leftarrow (c_n, z)$ taking a constant time and denoting with C_2. It executes once by receiving a pre-calculated value.
- The statement if-else from lines 28–43 repeats for N number of nodes. This means that the probability p(n_1) will repeat n times and probability p(n_2) will also repeat n times. The time taken will be $n \cdot n$ times.
- Count = count + 1 will take n times to execute.
- The if-else statement from lines 46–53 takes a constant time to execute. We denote it as C_4 and C_5.

From the above time complexity at each step, we can then calculate

$$C_1 + C_2 + C_3 + n \cdot n + n + C_4 + C_5.$$

By dropping constants and n, we are left with n^2. The conclusion shows that the time complexity of the clustering algorithm is $O(n^2)$

6.2. Time Complexity of the Q-Learning Algorithm

The symbols used for calculation are given below:
Zeros-matrix $\longrightarrow$ *Q-table*
Distance between nodes $\longrightarrow d_t$
Maximum communication distance $\longrightarrow R_{com}$
Energy remaining of nodes $\longrightarrow E_{pr}$
The time complexity of an algorithm can be determined incrementally and then added together to determine the algorithm's overall complexity.
Solution:

- The for-loop statement from lines 10–13 takes n seconds to compute the next hopping node, and lines 14–20 take n seconds to check for available next hopping nodes. Therefore, it will take $n \cdot n$ times due to the inner loop.
- For calculating the reward and Q-table, the while-loop form lines 21–27 takes n time.
- For updating the Q-table, the for-loop form lines 29–33 takes n time.
- Changing state $s_t \leftarrow s_{t+1}$ takes a constant time.

From the above discussion, the time complexity can be calculated as:

$$n + n \cdot n + n + C_1.$$

By dropping constants and n times, we have n^2. The conclusion shows that the time complexity of the QL- Algorithm is $O(n^2)$.

6.3. Time Complexity of the Best Path Selection Algorithm

The symbols used for calculation are given below:
Zeros-matrix $\longrightarrow$ *Q-table*
Energy remaining of nodes $\longrightarrow E_{pr}$
Distance between nodes $\longrightarrow d_t$
The time complexity can be calculated step-by-step and then combined to show the algorithm's overall complexity.
Solution:

- The for-loop from lines 7–9 will take n time.
- The while-loop from lines 10–16 will take n time.
- DThe inner loop form lines 7–21 is takes $n \cdot n$ times.

$$n + n + n \cdot n$$

By dropping n times, we have n^2 left. The conclusion shows that the time complexity of the best path selection algorithm is $O(n^2)$.

7. Results and Discussion

We used MATLAB to test the performance of the proposed algorithm and obtain simulation results.

7.1. Evaluation Metrics

We evaluated the proposed algorithm in terms of throughput, residual energy, packet delivery ratio, end-to-end delay and network lifetime, concerning different mobility speeds of the network and compare it with the existing protocol. We presented two schemes: CRP-GR (i.e., based on game theory and reinforcement learning) and CRP-G (i.e., based on game theory). The results were generated at a learning rate $\alpha = 0.9$ and regulatory factor $\beta = 0.9$. Table 3 shows the simulation parameters.

Table 3. Simulation parameters.

Parameters	Values
Area of sensor deployment	100 × 100 m
Location of base-station	(50, 50) m
Total number of nodes	100
Number of normal nodes	50
Femtocells (Intermediate nodes)	30
Picocells (Advance nodes)	20
Initial energy of normal nodes	0.5 J
Initial energy of intermediate nodes	0.7 J
Initial energy of advance nodes	1 J
Mobility of normal nodes	1–5 m/s
E_{Ta}	50×10^{-10}
E_{Ra}	50×10^{-10}
Maximum rounds (rmax)	3000

7.1.1. Throughput

The throughput is the sum of the $Pkt_{(received)}$ by the base station in the $T_{(period)}$, where the packet sizes $Pkt_{(size)}$ are in bits. Equation (26) shows the numerical form of the throughput.

$$T_{(r)} = \frac{Pkt_{(received)} \times Pkt_{(size)}}{T_{(period)}}. \tag{26}$$

7.1.2. Residual Energy

Residual energy is the average energy consumed, $ENG_{(cons)}$, at each round by a node divided by the available total energy $ENG_{(available)}$. Equation (27) shows a numerical form of the residual energy.

$$RSD_{(ENG)} = \frac{ENG_{(cons)}}{ENG_{(available)}} \times 100. \tag{27}$$

7.1.3. Packet Delivery Ratio

The packet delivery ratio is the proportion of received data packets, $Pkt_{(received)}$, by a base station to the sum of sent data packets $Pkt_{(send)}$. Equation (28) shows a numerical form of the packet delivery ratio.

$$Pkt_{(DR)} = \frac{Pkt_{(received)}}{Pkt_{(send)}} \times 100. \tag{28}$$

7.1.4. Average End-to-End Delay

End-to-end delay can be defined as the average delay between the receipt and packet sources at corresponding delivery. Equation (29) shows the average end-to-end delay of data packets per round.

$$D_{average} = \sum_{i=0}^{Pkt_{rcd}} \frac{T_{received}(i) - T_{send}(i)}{Pkt_{(received)}}. \tag{29}$$

7.1.5. Network Lifetime

Network lifetime is the number of alive nodes in the network during simulation time or after a specified scenario comes to an end.

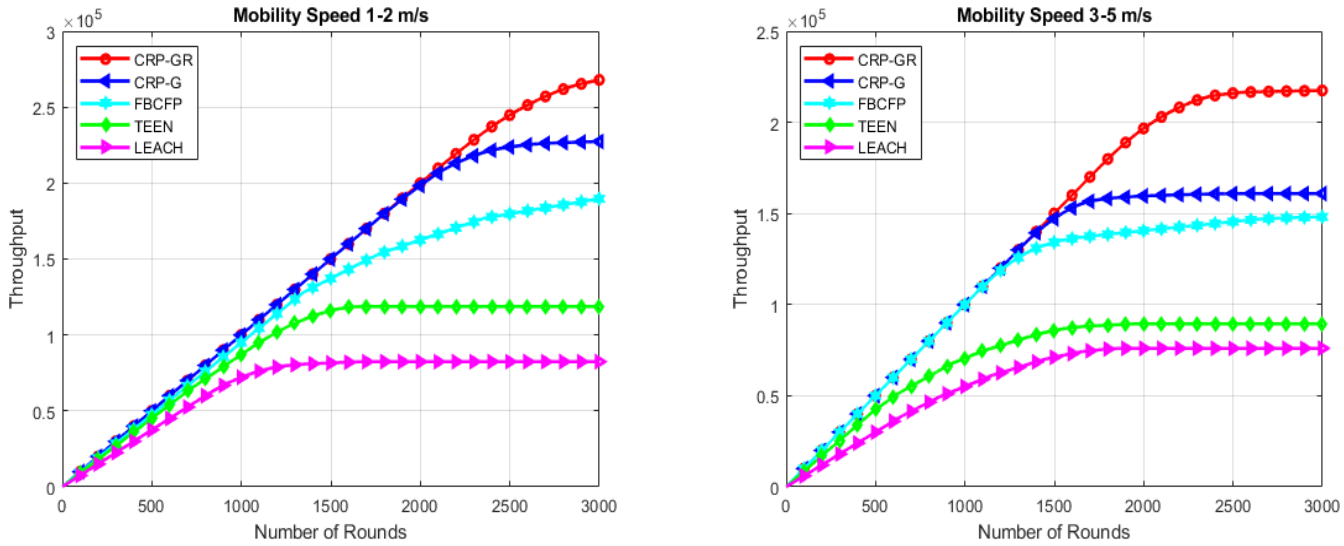
7.2. Results

The proposed schemes are evaluated and compared with FBCFP, TEEN and LEACH in terms of throughput, packet delivery ratio, residual energy, average end-to-end delay and network lifetime at different mobility speeds. The TEEN and LEACH protocols are

selected for comparison because these are the standard routing protocols and many more protocols are created based on these protocols, such as FBCFP [38].

7.2.1. Throughput

Every cluster group has a cluster head in the proposed scheme based on various factors (i.e., node mobility, the distance between node and base station and residual energy). Figure 8a,b shows throughput versus the number of rounds with different mobility speeds of nodes. It can be observed from the graphs that the throughput decrease with the increasing mobility speed of the nodes. It is due to the nodes dropping out of the cluster, which leads to frequent link disconnection. Furthermore, our proposed schemes CRP-GR and CRP-G provide better throughput as compared to FBCFP, TEEN and LEACH. The throughput of our proposed scheme is better due to optimum cluster head selection and best next hopping node selection with reinforcement learning. Similarly, decreasing packet loss is also improving throughput. The average performance of the proposed schemes CRP-GR and CRP-G with respect to throughput is 78% and 71%, while the existing schemes, such as FBCFP, TEEN and LEACH are 63%, 48% and 35% accordingly.

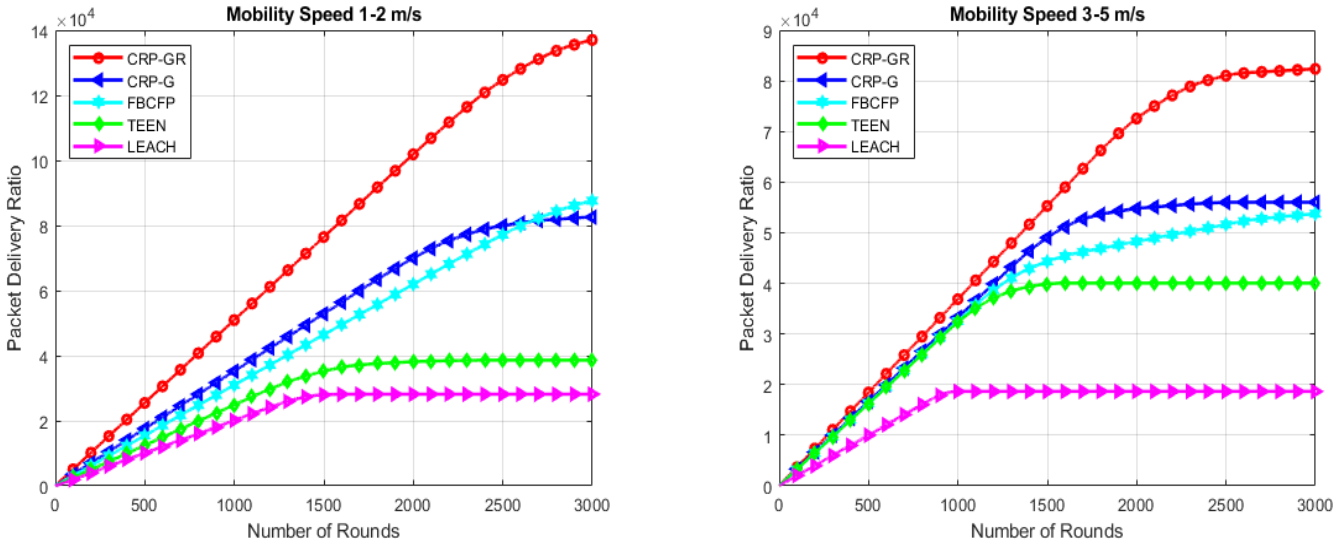


(**a**) (**b**)

Figure 8. Throughput vs number of rounds.

7.2.2. Packet Delivery Ratio

Our proposed schemes CRP-GR and CRP-G have the best delivery packet delivery ratio compared to FBCFP, TEEN and LEACH. Figure 9a,b show packet delivery ratio versus the number of rounds with different mobility speeds of the nodes. The packet delivery ratio decreases due to the available fixed bandwidth of the network by increasing the number of nodes. The FBCFP, TEEN and LEACH delivery ratio is less due to selecting repeated routes for data transmission from source to destination. Our proposed schemes perform better than other schemes due to selecting the optimal cluster head (CH) for every cluster and better node distribution, making fluent data transmission between cluster head (CH) and base station. The average performances of the proposed schemes, CRP-GR and CRP-G, with respect to the packet delivery ratio were 67% and 49%. The existing schemes, such as FBCFP, TEEN and LEACH, had average performances of 42%, 25% and 19% accordingly.

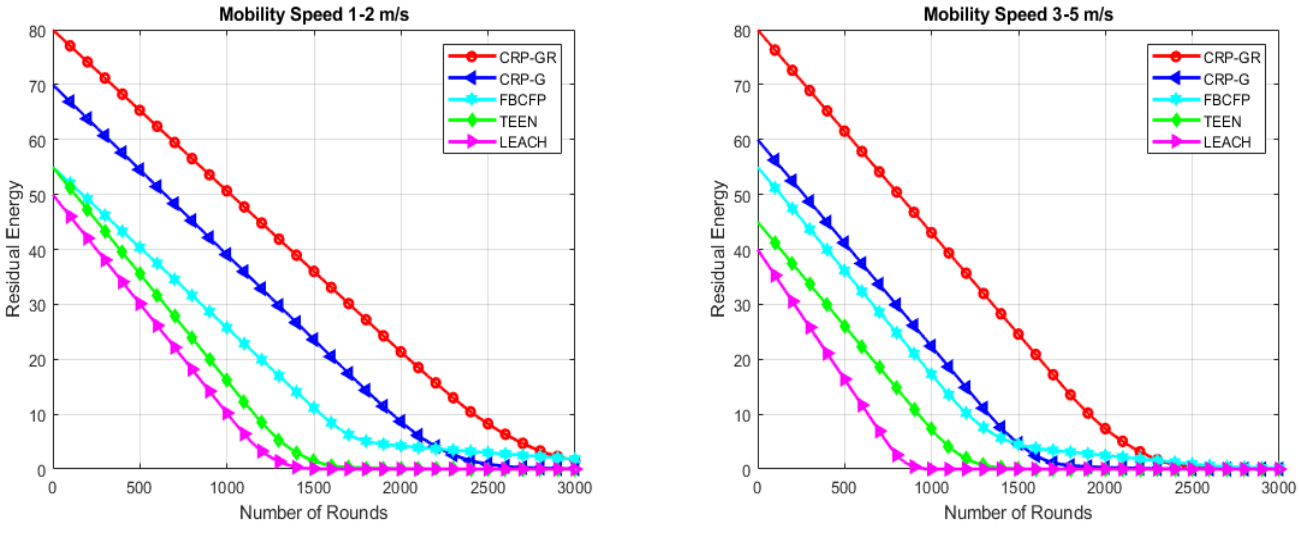


(**a**) (**b**)

Figure 9. Packet delivery ratio vs number of rounds.

7.2.3. Residual Energy

Our proposed scheme avoids the random selection of cluster heads (CHs) and uneven distribution in the network. Randomly selecting and distributing cluster heads (CHs) in a distributed and centralized network can lead to long-distance transmission and more energy consumption. Therefore, our schemes decrease energy consumption compared to FBCFP, TEEN and LEACH, as shown in Figure 10a,b with different mobility speeds of the nodes. In our proposed schemes, the cluster head (CH) was based on game theory by considering various factors (i.e., minimum distance, mobility and energy of nodes). Due to these, energy dissipation is reduced during communication, and nodes in the network consume less energy. The FBCFP, TEEN and LEACH spent more of their energy on cluster heads (CHs) selection at each round, data transmission and monitoring its cluster members. The average performances of the proposed schemes, CRP-GR and CRP-G, with respect to residual energy were 69% and 52%. The existing schemes, such as FBCFP, TEEN and LEACH had average performances of 37%, 24% and 19%, accordingly.

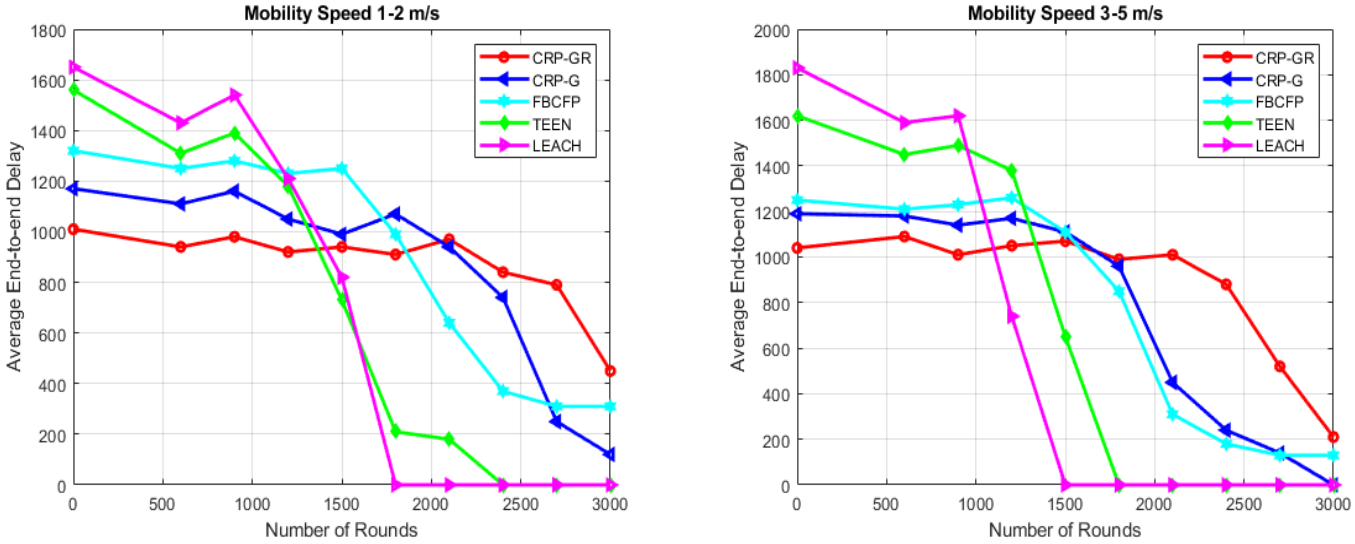


(**a**) (**b**)

Figure 10. Residual energy vs number of rounds.

7.2.4. Average End-to-End Delay

Our proposed scheme has the best CHs selection, which does not need a communication overhead and decreases transmission of redundant information to cluster heads (CHs). The average end-to-end delay of our proposed schemes is minimal compared to FBCFP, TEEN and LEACH as shown in Figure 11a,b with different mobility speeds of nodes. It is because of the early identification of the next hopping node and the remaining energy of nodes for better route selection. While, in the other schemes, the cluster head (CH) selects the same route for data transmission in each round, leading to more congestion and delay. The proposed schemes CRP-GR and CRP-G decrease the average end-to-end delay by 9% and 5%, respectivley, as compared to FBCFP, 18% and 11% as compared to TEEN, and 25% and 17% as compared to LEACH.



(**a**) (**b**)

Figure 11. End-to-End delay vs number of rounds.

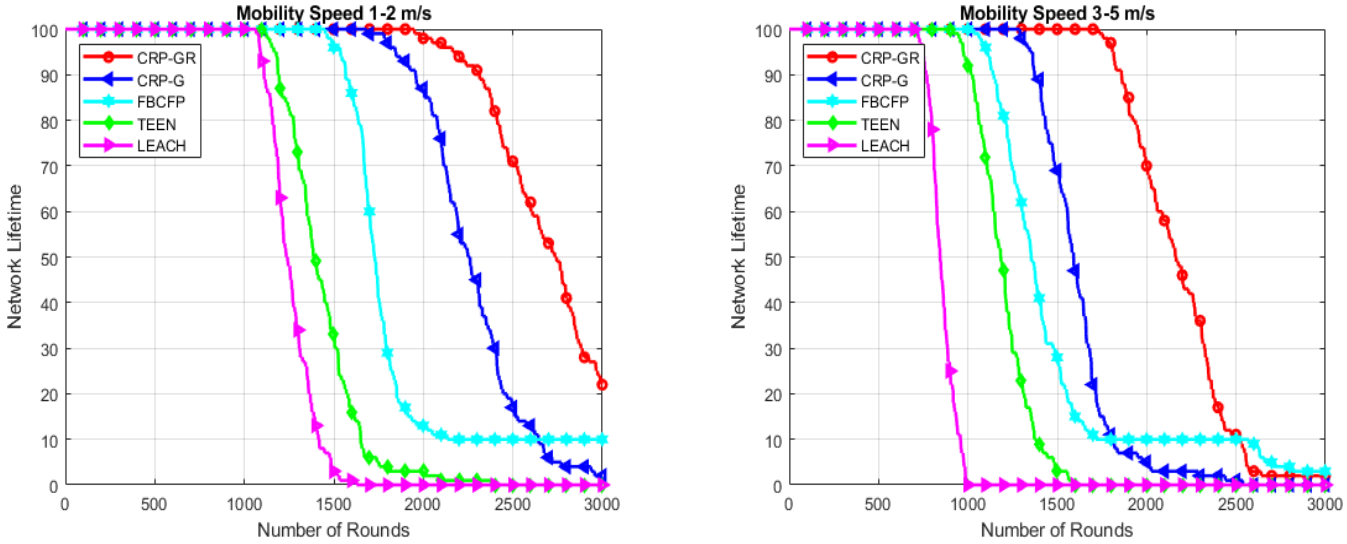
7.2.5. Network Lifetime

The early identification and consideration of various factors (i.e., remaining energy, distance, next hopping node, and node mobility) in our proposed schemes result in the optimal cluster head selection for each cluster and data transmission, thereby increasing the network lifetime. It reduces clustering overhead, supporting the network nodes to stay for more time and stabilize the network. Our schemes remove extra duties from cluster heads, improve energy efficiency, and decrease the number of death nodes. Our proposed scheme network lifetime is better compared to FBFCP, TEEN and LEACH, as shown in Figure 12a,b with different mobility speeds. The average performances of the proposed schemes CRP-GR and CRP-G, with respect to network lifetime, were 84% and 71%. The existing schemes, such as FBCFP, TEEN and LEACH had average performances of 58%, 45% and 39%, accordingly.

7.3. Discussion

Previous works presented different techniques for cluster head (CH) selection and routing paths for transmission. Different protocols, such as FBFCP, TEEN and LEACH, considered various factors (e.g., minimum distance, available connections between nodes, and residual energy) to prolong the lifetime of the network. However, due to more processing and disconnection of links, these protocols consume more energy, decreasing the network lifetime and packet delivery ratio. In our schemes, early identification of the next hopping node with minimum distance, high energy level with reinforcement learning, and optimum cluster head (CH) selection with game theory improves network lifetime, packet delivery ratio, and decreased energy consumption and end-to-end delay.

In our proposed schemes, each cluster head (CH) selection is based on a distance between the node and base station, node mobility, link connection with other nodes, and residual energy. In this scheme, the optimum cluster head (CH) selection with game theory and best route path selection with reinforcement learning (RL) were used to prevent overhead in communication and allow smooth transmission of data between clusters heads (CHs) and the base station. In terms of throughput, residual energy, packet delivery ratio, end-to-end delay, and network lifetime, our scheme outperforms the FBFCP, TEEN, and LEACH schemes. This is because our scheme selects the optimal cluster head (CH) and optimal route path for data transmission from nodes to cluster head (CH) and from the cluster head to the base station. On the other hand, the random cluster head (CH) and route path selection need more calculation and a long transmission range, leading to energy wastage, which is prevented in our proposed schemes. Therefore, the proposed scheme minimizes the energy consumption of the nodes in the network. In this manner, the energy dissipation reduces and improves the success of the energy-saving of the nodes.



(**a**) (**b**)

Figure 12. End-to-End delay vs number of rounds.

8. Conclusions

In heterogeneous 5G-based smart healthcare, the clustering-based routing protocol plays a vital role to transmit data from the source to the base station without delay. This paper proposed a clustering-based routing protocol based on game theory and reinforcement learning (i.e., Q-learning) for the heterogeneous 5G-based smart healthcare network. The cluster head (CH) selection probability is calculated using different attributes, such as the distance between nodes, nodes and the base station, mobility speed and remaining energy, using the symmetric game with mixed strategies. The multipath routing based on the Q-learning defined energy-efficient paths and distance with the help of derived iterative formula for the Q-table. The simulation result shows that our proposed clustering-based routing protocol improves QoS and energy optimization of the network compared to the existing schemes, i.e., FBCFP, TEEN and LEACH. Furthermore, network performance can be adjusted using the learning rate α and regulatory factor β. Choosing an appropriate value of α and β improves network lifetime and reduces the end-to-end delay for realistic demands. In the future, we will extend our research to investigate how to expand the network to adopt medical healthcare sensors to construct a flexible network to deal with actual medical data for emergencies.

Author Contributions: All the authors contributed to the work and wrote the article. A.A. proposed the idea, designs, and writing. M.T., M.A.S., K.I.A. and A.M. suggested directions for the detailed

designs as well as coordinated the work. All authors have read and agreed to the published version of the manuscript

Funding: Sunway University PGR Support.

Institutional Review Board Statement: The study was conducted according to the guidelines of Sunway University, Malaysia.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pirbhulal, S.; Zhang, H.; Alahi, E.; Eshrat, M.; Ghayvat, H.; Mukhopadhyay, S.C.; Zhang, Y.-T.; Wu, W. A novel secure IoT-based smart home automation system using a wireless sensor network. *Sensors* **2017**, *17*, 69. [CrossRef] [PubMed]
2. Ahad, A.; Tahir, M.; Alvin Yau, K.-L. 5G-Based Smart Healthcare Network: Architecture, Taxonomy, Challenges and Future Research Directions. *IEEE Access* **2019**, *7*, 100747–100762. [CrossRef]
3. Kristoffersson, A.; Lindén, M. A Systematic Review on the Use of Wearable Body Sensors for Health Monitoring: A Qualitative Synthesis. *Sensors* **2020**, *20*, 1502. [CrossRef] [PubMed]
4. Wu, W.; Pirbhulal, S.; Sangaiah, A.K.; Mukhopadhyay, S.C.; Li, G. Optimization of signal quality over comfortability of textile electrodes for ECG monitoring in fog computing based medical applications. *Future Gener. Comput. Syst.* **2018**, *86*, 515–526. [CrossRef]
5. Pirbhulal, S.; Zhang, H.; Wu, W.; Mukhopadhyay, S.C.; Zhang, Y.-T. Heart-beats based biometric random binary sequences generation to secure wireless body sensor networks. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 2751–2759.
6. Magsi, H.; Sodhro, A.H.; Chachar, F.A.; Abro, S.A.K.; Sodhro, G.H.; Pirbhulal, S. Evolution of 5G in Internet of medical things. In Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 3–4 March 2018; pp. 1–7.
7. Ahad, A.; Faisal, S.A.; Ali, F.; Jan, B.; Ullah, N. Design and performance analysis of DSS (dual sink based scheme) protocol for WBASNs. *Adv. Remote. Sens.* **2017**, *6*, 245. [CrossRef]
8. Wu, W.; Pirbhulal, S.; Zhang, H.; Mukhopadhyay, S.C. Quantitative assessment for self-tracking of acute stress based on triangulation principle in a wearable sensor system. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 703–713. [CrossRef]
9. Farahani, B.; Firouzi, F.; Chang, V.; Badaroglu, M.; Constant, N.; Mankodiya, K. To- wards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and health-care. *Future Gener. Comput. Syst.* **2018**, *78*, 659–676. [CrossRef]
10. Pirbhulal, S.; Zhang, H.; Mukhopadhyay, S.C.; Li, C.; Wang, Y.; Li, G.; Wu, W.; Zhang, Y.-T. An efficient biometric-based algorithm using heart rate variability for securing body sensor networks. *Sensors* **2015**, *15*, 15067–15089. [CrossRef]
11. Xu, B.; Xu, L.D.; Cai, H.; Xie, C.; Hu, J.; Bu, F. Ubiquitous data accessing method in IoT-based information system for emergency medical services. *IEEE Trans. Ind. Inform.* **2014**, *10*, 1578–1586.
12. Khan, M.F.; Yau, K.-L.A.; Noor, R.M.D.; Imran, M.A. Survey and taxonomy of clustering algorithms in 5G. *J. Netw. Comput. Appl.* **2020**, *154*, 102539. [CrossRef]
13. Ahad, A.; Ullah, Z.; Amin, B.; Ahmad, A. Comparison of energy efficient routing protocols in wireless sensor network. *Am. J. Netw. Commun.* **2017**, *6*, 67–73. [CrossRef]
14. Duan, J.; Gao, D.; Yang, D.; Foh, C.H.; Chen, H.-H. An energy-aware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications. *IEEE Internet Things J.* **2014**, *1*, 58–69. [CrossRef]
15. Kao, C.-C.; Lin, Y.-S.; Wu, G.-D.; Huang, C.-J. A comprehensive study on the internet of underwater things: Applications, challenges, and channel models. *Sensors* **2017**, *17*, 1477. [CrossRef]
16. Sodhro, A.H.; Pirbhulal, S.; Sangaiah, A.K. Convergence of IoT and product lifecycle management in medical health care. *Future Gener. Comput. Syst.* **2018**, *86*, 380–391. [CrossRef]
17. Deng, R.; Lu, R.; Lai, C.; Luan, T.H.; Liang, H. Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption. *IEEE Internet Things J.* **2016**, *3*, 1171–1181. [CrossRef]
18. Wang, K.; Wang, Y.; Sun, Y.; Guo, S.; Wu, J. Green industrial Internet of Things architecture: An energy-efficient perspective. *IEEE Commun. Mag.* **2016**, *54*, 48–54. [CrossRef]
19. Kaur, N.; Sood, S.K. An energy-efficient architecture for the Internet of Things (IoT). *IEEE Syst. J.* **2015**, *11*, 796–805. [CrossRef]
20. Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, Hawaii, 4–7 January 2000; p. 10.
21. Mechta, D.; Harous, S.; Alem, I.; Khebbab, D. LEACH-CKM: Low energy adaptive clustering hierarchy protocol with K-means and MTE. In Proceedings of the 2014 10th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, 9–11 November 2014; pp. 99–103.

22. Ye, M.; Li, C.; Chen, G.; Wu, J. EECS: An energy efficient clustering scheme in wireless sensor networks. In Proceedings of the PCCC 2005: 24th IEEE International Performance, Computing, and Communications Conference, Phoenix, AZ, USA, 7–9 April 2005; pp. 535–540.
23. Mammu, A.S.K.; Hernandez-Jayo, U.; Sainz, N.; De la Iglesia, I. Cross-layer cluster-based energy-efficient protocol for wireless sensor networks. *Sensors* **2015**, *15*, 8314–8336. [CrossRef]
24. Sarma, H.K.D.; Mall, R.; Kar, A. E 2 R 2: Energy-efficient and reliable routing for mobile wireless sensor networks. *IEEE Syst. J.* **2015**, *10*, 604–616. [CrossRef]
25. Boukerche, A.; Zhou, X. A Novel Hybrid MAC Protocol for Sustainable Delay-Tolerant Wireless Sensor Networks. *IEEE Trans. Sustain. Comput.* **2020**, *5*, 455–467. [CrossRef]
26. Manchanda, R.; Sharma, K. SSDA: Sleep-Scheduled Data Aggregation in Wireless Sensor Network-Based Internet of Things. In *Data Analytics and Management*; Springer: Singapore, 2021; pp. 793–803.
27. Komuro, N.; Hashiguchi, T.; Hirai, K.; Ichikawa, M. Development of Wireless Sensor Nodes to Monitor Working Environment and Human Mental Conditions. In *IT Convergence and Security*; Springer: Singapore, 2021; pp. 149–157.
28. Ahad, A.; Tahir, M.; Sheikh Sheikh, M.A.; Hassan, N.; Ahmed, K.I.; Mughees, A. A Game Theory Based Clustering Scheme (GCS) for 5G-based Smart Healthcare. In Proceedings of the 2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT), Piscataway, NJ, USA, 9–11 November 2020; pp. 157–161.
29. Papachary, B.; Venkatanaga, A.M.; Kalpana, G. A TDMA Based Energy Efficient Unequal Clustering Protocol for Wireless Sensor Network Using PSO. In *Recent Trends and Advances in Artificial Intelligence and Internet of Things*; Springer: Cham, Switzerland, 2020; pp. 119–124.
30. Li, Z.; Chen, R.; Liu, L.; Min, G. Dynamic resource discovery based on preference and movement pattern similarity for large-scale social Internet of Things. *IEEE Internet Things J.* **2015**, *3*, 581–589. [CrossRef]
31. Qiu, T.; Luo, D.; Xia, F.; Deonauth, N.; Si, W.; Tolba, A. A greedy model with small world for improving the robustness of heterogeneous Internet of Things. *Comput. Netw.* **2016**, *101*, 127–143. [CrossRef]
32. Yang, L.; Lu, Y.-Z.; Zhong, Y.-C.; Wu, X.-Z.; Xing, X.-J. A hybrid, game theory based, and distributed clustering protocol for wireless sensor networks. *Wirel. Netw.* **2016**, *22*, 1007–1021. [CrossRef]
33. Liu, Q.; Liu, M. Energy-efficient clustering algorithm based on game theory for wireless sensor networks. *Int. J. Distrib. Sens. Netw.* **2017**, *13*, 1550147717743701. [CrossRef]
34. Lin, D.; Wang, Q.; Lin, D.; Deng, Y. An energy-efficient clustering routing protocol based on evolutionary game theory in wireless sensor networks. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 409503. [CrossRef]
35. Thandapani, P.; Arunachalam, M.; Sundarraj, D. An energy-efficient clustering and multipath routing for mobile wireless sensor network using game theory. *Int. J. Commun. Syst.* **2020**, *33*, e4336. [CrossRef]
36. Osborne, M.J. *An Introduction to Game Theory*; Oxford University Press: New York, NY, USA, 2004; Volume 3.
37. Rani, S.; Solanki, A. Data Imputation in Wireless Sensor Network Using Deep Learning Techniques. In *Data Analytics and Management*; Springer: Singapore, 2021; pp. 579–594.
38. Thangaramya, K.; Kulothungan, K.; Logambigai, R.; Selvi, M.; Ganapathy, S.; Kannan, A. Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT. *Comput. Netw.* **2019**, *151*, 211–223. [CrossRef]