SDFMUCR: Software-defined and fuzzy-based multi-hop unequal cluster routing

Lu Sixu, Wu Muqing and Zhao Min

Abstract
Recently, wireless sensor networks play an important role in our life. Researchers give more and more attention to the design of the routing algorithm. Cluster routing is the most popular routing protocol in wireless sensor networks. However, hotspot problem always exists. Unequal cluster routing is an effective method to solve this problem. In these protocols, nodes closer to base station have the smaller cluster size than other sensor nodes. It will reduce the relay pressure of the node which is near to the base station. The development of the software-defined network optimizes the performance of wireless sensor networks. In this article, we propose a software-defined and fuzzy-based multi-hop unequal cluster routing protocol. It adopts three methods to make the routing protocol outstanding. First, it uses a novel software-defined network architecture to reduce computation overhead and energy overhead in sensor nodes. Then, the fuzzy logic–based intracluster and intercluster routing is designed to balance energy consumption. Finally, residual energy–based rotating scheme is proposed to reduce most of the control overhead. According to the simulation results, the proposed protocol balances and reduces the energy consumption, extends the lifetime, and reduces the control overhead and latency of the network.

Keywords
Wireless sensor network, software-defined network, unequal clustering, fuzzy logic, network lifetime

Date received: 31 October 2019; accepted: 8 September 2020
Handling Editor: Jasmine Seng

Introduction
In the past years, wireless sensor networks (WSNs) have gained more and more attention because of the development of the micro-electro-mechanism system (MEMS), wireless communication techniques, and so on. WSNs play an important role in environmental sensing, military defiance, habitat, tracking, monitoring, and so on.

WSN is the network which consists of a large number of wireless nodes. Each node is equipped with sensors and a battery. They can sense and transmit information to the base station (BS) through single hop or multi-hops. Generally, WSNs are deployed in harsh environment which makes it difficult to recharge the battery of the sensor nodes. This reduces the network lifetime and leads to an energy consumption unbalance. Many studies show that the wireless transmission unit consumes more energy than other units. As a consequence, designing an efficient routing protocol is a
great way to increase the lifetime of the entire network. Cluster routing has been paid more and more attention because of the advantage of a balanced energy consumption. In cluster-based WSNs, cluster members (CMs) gather and transmit the sensor data to the cluster heads (CHs) through the schedule of time-division multiple access (TDMA). Then, CHs aggregate the sensor data and transmit it to the BS. Clustering algorithm reduces the amount of data transmitted in the network because of the aggregation in the CHs. As a result, the network energy consumption is reduced.

In most cluster routing algorithms, CHs consume more energy than CMs. This is the hotspot problem in WSNs. Unequal clustering approach is a good way to overcome the hotspot problem. In the unequal clustering approach, the size of cluster which is closer to BS is smaller. And the cluster size of CH which is far away from the BS is larger. CHs with less CMs consume less energy in the phase of aggregate and transmit aggregated data. They use the redundant energy to relay data packets far from the BS.

When determining the optimal size of clusters, the distance to BS, remain energy, density, and so on are considered. When there are lots of uncertainties, fuzzy logic (FL) is one of the best approaches to prevent this problem.

FL is a logical system which consists of four parts: (a) fuzzifier, (b) inference system, (c) de-fuzzifier, and (d) fuzzy rule. Crisp values are input to the FL system and are converted into fuzzy linguistic variables. Then, fuzzy linguistic variables are input to the inference system which includes the fuzzy rule. Then, inference system uses if-then rules to convert fuzzy linguistic variables into fuzzy variables. Finally, the fuzzy variables are input to the de-fuzzifier and converted into a crisp value.

As it is well known, the first-order radio model is used as the energy consumption model for WSNs. So, it can consume more energy in a large wireless network if CHs transmit data directly to the BS. Therefore, CHs far from the BS can transmit data to the CH which is near to the BS to reduce energy consumption.

The WSNs need to balance the energy by rotating the CHs among sensor nodes. However, it usually re-clustering in each round makes it consume more time, energy, and calculation resources. Hence, it should both consider the network lifetime and energy consumption balance. In this article, we propose an algorithm which considers multiple factors comprehensively to design the best rotating rounds after each re-clustering.

In most distributed routing protocols, cluster routing protocols for example, each node should collect the information of other nodes in the CH selection phase. Meanwhile, CHs need dispatched messages to find the routing between CHs and BS. However, it can be difficult for an edge node to find the best routing to the BS when using distributed routing in a large network. It will increase the routing overhead in order to guarantee the success in routing discovery phase.

In conclusion, the issues of the existing work are given as follows:

1. In most distributed routing protocols, sensor nodes need a lot of dispatched messages to find the routing to BS. The routing is calculated in the sensor nodes, which makes the computation overhead high.
2. In most cluster routing algorithms, nodes far from the BS die quicker than other nodes when CHs transmit data to the BS directly. However, the nodes which are near to the BS consume energy quicker than other nodes when CHs transmit data to BS through multi-hops.
3. The WSNs need to balance the energy by rotating the CHs among sensor nodes. However, it usually re-clustering in each round makes it consume more time, energy, and calculation resources. Hence, it should both consider the network lifetime and energy consumption balance.

In this article, we propose a software-defined and fuzzy-based multi-hop unequal cluster routing (SDFMUCR) algorithm. The main contributions of this article are given as follows:

1. According to issue 1, a novel SDN framework is used in this algorithm to reduce the energy consumption of transmitting redundant messages in the cluster setup phase. This method reduces the computation overhead and the routing overhead in the sensor nodes which prolongs the network lifetime. The BS acts as controller which uses fuzzy-based multi-hop unequal cluster routing and residual energy-based rotating scheme to calculate the routing of the entire
network. Sensor nodes only need to gather and transmit the data to the next hop which has been designed in controller.

2. According to issue 2, the proposed protocol uses fuzzy-based multi-hop unequal cluster routing to ensure the unequal clustering and choose the best relay node. Different from other proposed protocols, we propose the “relative intercluster cost,” “relative intracluster cost,” and “relative relay cluster cost” innovatively to design the best routing of the whole network. Unequal cluster scheme reduces the intracluster cost for the CHs which are near to the controller. A novel algorithm is proposed to find the optimum CH for CMs. This algorithm considers three factors innovatively in order to find the best CH for CMs. Multi-hop scheme reduces the intercluster cost for the CHs which are far from controller. These methods balance the energy consumption and prolong the network lifetime.

3. According to issue 3, the proposed residual energy–based rotating scheme gives a high–performance network. This method reduces the number of times to re-cluster the routing and reduces the control overhead in each round. Only when WSN satisfies the conditions which are proposed in section “Residual energy–based rotating scheme,” it will re-calculate the routing of the entire network. This scheme reduces the energy consumption in sensor nodes as well as balances the energy of the entire network.

The rest of the article is organized as follows. Section “Related works” reviews the related works of routing protocols in WSNs. The network assumptions and energy model are described in section “System model.” Section “SDFMUCR protocol” introduces the framework of the SDWSN which is used in this article. Meanwhile, this section describe the intracluster routing and the intercluster routing which use the FL system. Section “Theoretical analysis of the protocol” analyzes the protocol at the theoretical level. Simulation results and analysis are expounded in section “simulation results.” We will conclude the article in section “Conclusion.”

Related works

In recent years, the design of low-energy consumption routing protocols has gained more and more attention. At the mean time, FL and the software-defined network are usually used to extend the network lifetime and optimize other performances of the WSNs. The overview of the SDWSN and FL-based unequal cluster routing are given as follows.

SDWSN

For traditional routing protocols in WSNs, sensor nodes need to collect the related information such as residual energy, neighbor information, and so on. Then, sensor nodes use the information to calculate the routing. This method increases computation overhead in sensor nodes. However, only the controller needs to calculate the routing of the entire network in SDWSN. This will reduce computation overhead in sensor nodes. Meanwhile, this architecture reduces the consumption of energy and prolongs the network lifetime.

Many SDWSN-based frames have been proposed earlier. These frames usually have advantages like energy efficiency, robustness, scalability, or adaptability. However, many of them only contain one or two advantages as mentioned above. The proposed architecture includes all the features mentioned above, and the main contribution is to make the network energy efficient.

T Luo et al. proposed the SDWSN, which synergizes SDN and WSN to solve WSN-inherent problems. S Bera et al. designed a software-defined wireless sensor network (Soft-WSN) architecture which consists of the device and network management. The proposed scheme is able to meet real-time requirements of internet of things (IoT). These two architectures did not consider designing the effect flow table which is used to manage the energy consumption. We proposed the SDN framework overall consideration of the topology information module, controller module, flow table, and data management module. It well reduces the energy consumption of the network.

W Liao et al. designed an energy-efficient algorithm–based multi-energy-space (MES) in Soft-WSN. They designed the architecture of SDWSNs and proposed an MES-based algorithm. SDN-based routing protocol reduces the routing overhead as well as balances the energy consumption and increases the network lifetime.

P Li et al. used game theoretic to improve the energy consumption performance in SDWSN. The architecture of the SDWSN is same as Liao et al.’s and it also improves network performance. In Liao’s and Li’s algorithm, sensor nodes transmit the data packets to the controller by multi-hops and without data aggregate. This way increases consumption of energy and network overhead.

J Aparicio et al. designed an algorithm to improve the energy efficiency of SDN-based mobile wireless sensor network (MWSN). It adopts the sensor OpenFlow
protocol to achieve the SDN architecture. Meanwhile, they proposed a cross-over net algorithm between the network layer and transport layer. This algorithm increases the network lifetime as well as reduces the overhead in the central processing unit (CPU) and message retransmission.

X Wei et al. proposed a particle swarm optimization (PSO)-based algorithm in SDWSN. This protocol divides the network into clusters. And each cluster consists of one control node and a number of sensor nodes. The algorithm is suitable for a variety of scenarios and overcomes the NP-hard problem. The SDN-based protocols perform better than distributed routing protocols. But the huge number of control nodes will increase the cost of network.

N Abdolmaleki et al. designed the fuzzy topology discovery protocol (FTDP) which is according to the software-defined network solution for wireless sensor networks (SDN-WISE). FTDP separates the control logic from sensor node. This method reduces the routing overhead in sensor nodes. It increases the packet delivery ratio and the network lifetime.

**FL-based unequal cluster routing**

Unequal cluster routing is one of the best ways to solve the hotspot problem in WSNs. FL is a great method to implement unequal cluster-based routing protocols. FL can consider many factors (such as residual energy, distance to BS, etc) synthetically and make the protocol efficiently. Therefore, FL-based unequal cluster routing is a good method to improve the performance of the network.

Energy-aware unequal clustering algorithm (EAUCF) is an FL and unequal-based routing algorithm which is proposed earlier. Candidate CHs are elected randomly in WSNs first. The FL’s input parameters are the distance to BS and residual energy. The FL’s output parameter is competition radius. The final CHs are determined among the candidate CHs by comparing residual energy with each other. CMs transmit the data to CH and CHs transmit the aggregation data to BS directly. The lifetime is extended and the energy is balanced. Nevertheless, the algorithm does not consider the density of each node and if the final CHs with high density are selected, the energy consumption will be huge for these CHs.

The FL approach is used for unequal clustering in distributed load balancing unequal clustering (DUCF) too. It uses the residual energy, node degree, and distance to BS as input variables to fuzzy inference system (FIS). The output parameters are the chance to be CH and the number of CMs. It ensures the load balance among clusters and improves the network lifetime. However, it increases the transmission information in the cluster setup phase because of using the number of CMs as the output of fuzzy system.

Fuzzy-based unequal clustering algorithm (FUCA) uses the FL system to form unequal clusters. It prolongs the network lifetime and balances the energy consumption. It elects the candidate CHs randomly, and only candidate CHs can be the final CHs which reduces the CPU usage and the energy consumption in sensor nodes. The inputs of the fuzzy system are residual energy, node density, and distance to BS. The outputs of the fuzzy system are competition radius and rank. If a node has the highest rank within its competition radius, it will be selected as a CH. CHs collect the sensor data of the CMs and then transmit the aggregation data to BS. It does not consider the relay mechanism which makes the network consumption unbalanced.

Multi-objective fuzzy clustering algorithm (MOFCA) solves both the energy hole and the hotspot problem successfully using the FUCA. This algorithm has excellent performance in the total remaining energy (TRE), the lifetime of the WSNs, and the first node dies (FND). The protocol selects candidate CHs the same as the FUCA routing algorithm. Unfortunately, it causes the energy unbalance if low residual energy nodes are chosen as the candidate CHs.

Distributed FL-based energy-aware and coverage preserving unequal clustering algorithm (DEUC) considers the coverage parameter of the sensor nodes innovatively. It prolongs the coverage lifetime of the target area as well as the lifetime of the network. This routing protocol uses the FL system to ensure the unequal cluster and choose the optimum CHs. Then, non-CH nodes also use the FL system to select which cluster to join in. This scheme increases the calculation amount and energy consumption of sensor nodes. It uses the delay time when choosing CHs and when sensor nodes decide to join which cluster. The disadvantages are extending the time in CHs selection phase and the cluster formation phase.

In FL-based unequal clustering (FBUC) routing protocol, the CH selection is the same as the aforementioned routing protocols. The innovation of this protocol is to use the FL system makes CMs decide which CH to join. It reduces the transmission delay and power consumption. However, it does not design the intercluster routing which will not apply for a large network.

Some cluster-based algorithms design the reasonable intercluster routing such as ant colony optimization (ACO)-based relay routing protocol. Fuzzy- and ACO-based combined mac, routing, and unequal clustering cross-layer protocol (FAMACROW) and fuzzy-based unequal clustering and ACO-based routing hybrid protocol (FUCHAR) both use the FL for CHs selection and ACO for intercluster routing. FAMACROW improves performance of energy consumption, network lifetime and packet delivery rate. In order to reduce the
time to re-clustering, CMs rotate to act as CHs for each round. However, protocol should ensure that each CMs can communicate with the CH, and this requirement causes the size of cluster decreased. FUCHAR increases the network lifetime and solves the hotspot problem. In the cluster setup phase, it uses five values such as residual energy and node degree as the input to the FL system to calculate the cluster size and CH probability. This method increases the CPU and energy overhead in sensor nodes.

Distributed fuzzy approach to unequal clustering and routing algorithm (DFCR)\textsuperscript{29} is designed base on the unequal cluster and multi-hop intercluster routing algorithm. This routing protocol also uses the FL to guarantee the unequal cluster. It can lower the inter-cluster cost by proposing the intercluster communication rationally. In the distributed unequal cluster-based routing algorithm (DUCR), CHs find the routing to BS by broadcasting control messages and comparing the level information. However, it will increase the broadcast range continually until they find the optimum relay CH to the BS. It increases control messages that increments the probability of message collision and energy consumption.

Intra and intercluster communication (IICC)\textsuperscript{30} uses the fuzzy-based unequal cluster and multi-hop routing algorithm. It reduces the average energy consumption, reduces the end-to-end delay and increases the packet delivery ratio. However, the process of the intracluster routing and intercluster routing are both handled in sensor nodes which increases the energy consumption of the network.

DUCR\textsuperscript{31} reduces the energy consumption and prolongs the network lifetime. However, it does not consider the density in clustering phase. If a node which has a high density is selected as CH, it will unbalance the energy consumption of the network. Meanwhile, CH may broadcast more messages in order to search the optimal relay CH.

Shawkat K. Guirguis et al. proposed the hybrid fuzzy multi-hop unequal clustering algorithm (HFMUC) algorithm in order to increase the network performance in network lifetime and energy balance.\textsuperscript{32} This protocol uses the FL system for unequal cluster in heterogeneous dense WSNs. And CHs transmit the data packets to BS through multi-hop.

Hybrid unequal energy-efficient clustering (HEEC)\textsuperscript{33} uses the unequal cluster and multi-hop mechanism for intracluster routing. It also designs the multi-hop mechanism for intercluster routing. First, sensor nodes determine the layer of itself based on the distance to BS. Next, nodes calculate the radius and delay time based on the density, residual energy, and the distance to BS. Nodes broadcast the head message until the end of delay time. If a node receives a head message before the end of delay, it quits the CHs election. Then, CMs join the CH with the shortest distance. Finally, CHs select the next hop based on the layer information. CMs transmit the data packet to CH, and CHs transmit aggregation packet to BS hop by hop. In the next round, the node which has the highest residual energy will be selected as a new CH. It reduces network overhead, increases network stability, energy balance, and network lifetime.

The comparison of some of the algorithms with our proposed method is listed in Table 1.

### System model

#### Network assumption

The proposed algorithm adopts the SDN to reduce the energy consumption in the network. It uses the FL for intracluster and intercluster routing. Some assumptions of the network are listed as follows:

1. Sensor nodes act as switches at random distribution and stationary after deployment, and only exit the network when a node is dead due to energy exhaustion.
2. The BS acts as a controller.
3. Sensor nodes and controllers are equipped with global positioning system (GPS) to get their own accurate position.
4. Network links are symmetrical and sensor nodes are homogeneous.
5. The energy of the sensor nodes are limited, and the energy of the controller is sufficient.
6. Power of transmit data and calculation speed of the controller are stronger than those of sensor nodes.

In WSNs, sensor nodes are usually deployed in the harsh environment. So it is hard to recharge the battery of sensor nodes which makes the energy of the sensor nodes limited. A controller is used to collect topology information, calculate routing, collect sensor data, and so on. Therefore, it does not need to deploy in the harsh environment. It is convenient to recharge the battery. So the energy of the controller is sufficient. The controller must process more information and calculate the routing of the entire network. So the calculation speed of the controller must be higher. The transmitted power of the controller can be higher in order to reduce the packet loss in multi-hop routing.

#### Energy model

In WSNs, the total energy consumed of the sensor nodes includes the transmission, reception, and aggregation. The transmission and reception energy model uses the first-order radio model. The energy model
| Protocol                  | Years | Method   | Cluster organization | Cluster size | Intra com. | Inter com. | CH selection               |
|--------------------------|-------|----------|----------------------|--------------|------------|------------|----------------------------|
| NWPSO based              | 2016  | Centralized | ✓                   | Equal        | 1-hop     | 1-hop     | Based on NWPSO             |
| MES                      | 2017  | Centralized | ✗                   | ×            | 1-hop     | ×         |                           |
| Shortest path            | 2017  | Centralized | ✗                   | ×            | 1-hop     | ×         |                           |
| FTDP                     | 2017  | Centralized | ✗                   | ×            | 1-hop     | ×         |                           |
| Cross-over-net           | 2017  | Centralized | ✓                   | Equal        | 1-hop     | k-hop     | Based on usefulness        |
| Soft-WSN                 | 2018  | Centralized | ✗                   | ×            | 1-hop     | ×         |                           |
| EAUCF                    | 2013  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| MOFCA                    | 2015  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| DUCF                     | 2016  | Distributed| ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic       |
| FBUC                     | 2016  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| FAMACROW                 | 2016  | Distributed| ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic       |
| DECUC                    | 2017  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| FAMUC                    | 2017  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| DUCR                     | 2017  | Distributed| ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic       |
| FUCA                     | 2018  | Distributed| ✓                   | Unequal      | 1-hop     | 1-hop     | Based on fuzzy logic       |
| FUCHAR                   | 2018  | Distributed| ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic and ACO |
| DFCR                     | 2018  | Distributed| ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic       |
| HEEC                     | 2019  | Distributed| ✓                   | Unequal      | k-hop     | k-hop     | Based on attribute         |
| Propose protocol         | 2020  | Centralized | ✓                   | Unequal      | 1-hop     | k-hop     | Based on fuzzy logic and SDN framework |

NWPSO: nonlinear weight particle swarm optimization; MES: multi-energy-space; FTDP: fuzzy topology discovery protocol; Soft-WSN: software-defined wireless sensor network; EAUCF: energy-aware unequal clustering algorithm; MOFCA: multi-objective fuzzy clustering algorithm; DUCF: distributed load balancing unequal clustering; FBUC: FL-based unequal clustering; FAMACROW: fuzzy- and ACO-based combined MAC, routing, and unequal clustering cross-layer protocol; DECUC: distributed FL-based energy-aware and coverage preserving unequal clustering; HFMUC: hybrid fuzzy multi-hop unequal clustering algorithm; DUCR: distributed unequal cluster-based routing algorithm; FUCA: fuzzy-based unequal clustering algorithm; FUCHAR: fuzzy-based unequal clustering and ACO-based routing hybrid; DFCR: distributed fuzzy approach to unequal clustering and routing algorithm; HEEC: hybrid unequal energy-efficient clustering; ACO: ant colony optimization; SDN: software-defined networking.
consumption in transmitting $l$ bits packet over a distance $d$ is given by equation (1)

$$
E_{Tx}(l, d) = \begin{cases} 
  l \times E_{elec} + l \times e_{fs} \times d^2 & d \leq d_0 \\
  l \times E_{elec} + l \times e_{mp} \times d^4 & d > d_0
\end{cases}
$$

(1)

where $E_{elec}$ is the electronic energy dissipated by transmitter or receiver, $e_{fs}$ and $e_{mp}$ are the energy dissipated for free space and multipath fading. The parameter $d_0 = \sqrt{e_{fs}/e_{mp}}$ is the distance threshold. The energy dissipated is proportional to the $d^2$ when transmission distance is less than distance threshold. Otherwise, it is proportional to the $d^4$.

The energy consumed in receiving $l$ bits packet is given by equation (2)

$$
E_{Rx}(l) = l \times E_{elec}
$$

(2)

The energy consumed in aggregating data is given by equation (3)

$$
E_{DA} = N \times E_{agg}
$$

(3)

$N$ is the sensor nodes belong in a cluster. $E_{agg}$ is the aggregation energy dissipated for each packet.

The total energy consumed of the CHs includes the transmission, reception, and aggregation. Nevertheless, it only includes the transmission in the non-CHs. The total energy consumption for the CHs is given by equations (4) and (5). Where $agg$ is the data aggregation rate in the CH. $d_c$ is the distance between CH and controller. $d_{non-CH}$ is the distance between CH and non-CH

$$
E_{CH} = E_{Tx}(lN \times agg, d_c) + NE_{Rx}(l) + E_{DA}(N)
$$

(4)

$$
E_{non-CH} = E_{Tx}(l, d_{non-CH})
$$

(5)

**SDFMUCR protocol**

The SDFMUCR protocol can be divided into four phases. It includes the topology discovery phase, the routing table generate phase, the routing table delivery phase, and the data transmission phase. The phases of the proposed protocol are given as Figure 1.

In the topology discovery phase, the BS acts as a controller to broadcast control message. Sensor nodes send necessary information to the controller in order to let the controller calculate the routing. The details of the topology discovery phase are given by section “SDN framework for the proposed protocol.” Next, the controller designs the FL-based intracluster routing and intercluster routing. The details of the routing protocol are given by section “Fuzzy-based multi-hop unequal cluster routing.” Then, the controller delivers the flow table by multi-hop. Finally, sensor nodes transmit sensor data to the controller and start the topology maintain phase. If WSNs need to re-compute the flow tables in the topology maintain phase, it returns to the topology discovery phase. The trigger cases for re-clustering are given by section “Residual energy-based rotating scheme.”

**Residual energy–based rotating scheme**

For most cluster routing protocols, the network re-computes the routing after each round. It increases the calculation pressure and dissipates the energy resource in the cluster setup phase. In this article, residual energy-based rotating scheme is used to prevent this problem. The network updates the cluster routing until an interrupt occurs. The interval between interruptions should not be too long or too short. If too long, it causes energy unbalance in the network. On the contrary, it transmits more control messages and consumes more energy in the phase of topology discovery. Hence, it is necessary to design the best conditions in order to return topology discovery phase and to update the routing again. Some conditions to re-calculate the routing are listed as follows. It will be executed in the controller as long as one happens.

1. The number of alive nodes changes in the network.
2. The network has run more than $r_{max}$ after update the routing. Where $r_{max}$ is a constant, which is the maximum running rounds after the controller delivers the flow tables.
3. It satisfies the inequality (6). Where $E_{min}$ is the minimum energy of the whole network, and $E_{th}$ is the energy threshold for the controller to re-calculate the routing

$$
E_{min} < E_{th}
$$

(6)

We define the FND as the network lifetime. The purpose of this condition is to prolong the lifetime of the entire network. And $E_{min}$ is a key parameter of the entire network. If a sensor node with less energy still acts as a CH, it will run out of energy rapidly. When the energy of the sensor node is less than $E_{th}$, the controller will re-calculate the routing. This method
prolongs the lifetime of the entire network and avoids the sensor nodes with less energy continuously act as CHs.

4. It satisfies the inequality (7). Where $E_{\text{ave}}$ is the average energy of the entire network

$$E_{\text{ave}} - E_{\text{min}} < E_{\text{th}} \quad (7)$$

The purpose of this condition is to balance the energy consumption of the whole network. We hope energy consumption of all the sensor nodes in the network as consistent as possible. When the residual energy of a sensor node is far less than the average energy, the protocol can reduce the energy consumption of this node by re-calculating the routing of the entire network.

**SDN framework for the proposed protocol**

In the proposed protocol, we adopt the SDN framework in order to reduce the consumed energy. The network consists of sensor nodes and controller. Sensor nodes act as switches which only need to collect the position, residual energy, and other information of itself and inform the controller. The BS acts as a controller that computes the routing table of the whole network and transmits forward table to the sensor nodes. The proposed protocol uses a controller instead of sensor nodes to compute the routing of the entire network. This method reduces computation overhead and energy consumption in sensor nodes.

The protocol structures of controller and sensor nodes are given as Figure 2. Both of them have three layers: physical layer, network operating system (NOS) layer, and application layer. In order to implement function, the NOS layer replaces the transport layer, network layer, and data link layer. The physical layer is responsible for the signal processing. The application layer is responsible for application management. For controller, the NOS layer is used for topology information management, load balance management, energy efficiency management, virtualization management, and generating the flow table. For sensor nodes, the NOS layer is used for topology management as well as processes the flow table which is received from the controller.

The proposed SDWSN architecture for SFMUCR protocol is shown in Figure 3. It includes the application layer, control plane, data plane, northbound interface, and southbound interface. The application layer includes a variety of applications. The southbound interface is between the control plane and data plane. It is used to exchange data between the controller and sensor nodes. The northbound interface is between the application layer and control plane. The application layer uses this interface to get the resources of the underlying network. The controller works on the control plane and sensor nodes work on the data plane. Data plane consists of the sensor module, sending module and receiving module. Sensor module is used to sense and generate the sensor data. Sending and receiving module is used to transmit and receive the massage. Control plane consists of the topology information module, controller module, flow table, and data management module. Topology information is managed in the topology information module. These messages are processed in the controller module and then controller generates the flow table. Sensor data are managed in the data management module. The phases of the
The SDFMUCR protocol are topology discovery, topology maintain, flow table generation, and flow table delivery. The details of the phases are introduced below.

**Topology discovery phase.** The main purpose in the topology discovery phase is to collect the key information such as the location and the residual energy of the sensor nodes. The key information is collected by the controller and then is used to calculate the routing of the entire network.

The topo_rq message and the topo_rp message are used in the topology discovery phase. The controller broadcasts the topo_rq message in order to collect the topology information at the start of the topology discovery phase. Sensor nodes which receive the topo_rq message will packet the location (including x and y coordinates) with the residual energy into the topo_rp message. For the sensor nodes which distance to controller is far from the transmission range, a multi-hop routing mechanism is established to solve this problem. When a node receives the topo_rq message, it establishes a temporary one-way secure channel to the controller. Next, it rebroadcasts this topo_rq message until it gets to the edge of the network. The nodes which far from the controller then establish a one-way secure channel to the node which transmits this topo_rq message. So far, the multi-hop routing to controller has been established. Finally, sensor nodes send the topo_rp message through the secure channel. If a node receives the same topo_rq message of the same node in a round, it discards this message. This mechanism ensures the controller receives the topo_rp message from the nodes which are far from the controller. The controller records the topology information to the topology information module as long as it receives the topo_rp message.

In this phase, sensor nodes and controllers are equipped with GPS to get their own accurate position. Although the use of GPS consumes more energy than the GPS-free localization methods, it can get the exact location of the node which makes the routing of the entire network generated earlier. Meanwhile, the GPS goes into sleep mode once the sensor node gets its own position. What's more is that, the proposed protocol uses the residual energy-based rotating scheme to further reduce the work time of the GPS. These methods reduce the energy consumption of the GPS which makes the use of GPS energy efficient. A sensor node typically consists of a computational sensing element, a processing unit, a storage unit, and a power source. The GPS module is only a small part of the sensor node; its cost is much lower than other parts of the sensor node. Compared to the scene of GPS-free methods, this method does increase the cost of the network. Since most GPS modules cost less, the price of this approach is only a little more than the scene of GPS-free. However, this approach greatly improves the performance of the network. So this method is cost-efficient.

**Flow table generation and delivery phase.** The main purpose of the flow table generation phase is to calculate the routing of the entire network. The operations take place in the controller’s control module. The control module calculates the routing using the location and residual energy information which are in the topology information module. The fuzzy-based multi-hop unequal cluster routing is used to decrease the energy consumption and to extend the network lifetime which will be discussed in detail in section “Fuzzy-based multi-hop unequal cluster routing.”

The sensor nodes are divided into CHs and CMs in the cluster routing. Control module generates the flow table after designing the optimal routing. The flow table includes the CH flag (flagCH), precursor list, TDMA schedule list and next hop. For CHs, the CH flag is “1” and the precursor list lists the nodes whose data packets or aggregate packets need to be relayed by local node. The CH flag is “0” and the precursor list is empty for the CMs. The next hop domain is used to transmit data packets. The TDMA schedule list consists of the time slot number and the node number. When a CH receives the flow table, it generates a TDMA schedule for its precursor nodes. The precursor nodes number will be added to the TDMA schedule list from the smallest to the biggest. So far, the TDMA schedule list has been filled in completely. Then, the control module establishes two-way secure channels to deliver the flow table. Finally, the flow table will be delivered to the sensor nodes hop by hop. So far, the routing for the entire network is established.

**Data transmission and topology maintain phase.** The main purposes in the topology maintain phase are to transmit the data packets and agg_data packets and detect changes of the network in order to update the routing timely. Agg_data packet is the packet which is aggregated by CHs. The operations are different in different nodes. For CMs, they only awake in the time slot which is allotted by their CHs. In this time slot, CMs packet the sensor data with the residual energy in the data packet and transmit it to the controller according to the flow table in each round. Then they sleep in the remaining time. The residual energy is used in the control module to decide whether to update the flow table. It does not need the location information because the nodes are stationary after being deployed. As for CHs, they will be awake at any moment in order to receive the data packets from CMs. Once a CH receives all the data packets from precursor nodes, they aggregate the
data packets and transmit the agg_data packet to the next hop. They will also detect the precursor list in each round. The precursor list contains the CH and CMs which are the precursor nodes of the local node. If it does not receive the data packet from any precursor nodes, that means the precursor node dies and the precursor list has changed. If the precursor list changes, the CH establishes a one-way secure channel and transmits the topo_repair message to the controller.

In the topology maintain phase, the controller updates the topology information when it receives the data packet. It returns to the topology discovery phase when it receives the topo_repair message. Besides, some other reasons to let the network into the topology discovery phase are listed in section “Residual energy-based rotating scheme.”

The process of SDFMUCR protocol is shown in Algorithms 1 and 2.

Secure channel. In the proposed protocol, we design the dynamic secure channel. The design of the secure channel makes the control message of the entire network and the sensor data of the nodes transmitted in two channels, respectively. This method reduces the traffic load in one channel and avoids the collision between these two types of information. Normal channel and control channel are built at different frequencies. The secure channel is only established when the sensor nodes transmit the topo_rp message, controller delivers the flow_table to the entire network, or the running time of each round.

Algorithm 1: Process of SDFMUCR protocol in controller.

```plaintext
1: Initialization
2: Set node state as topology discovery phase
3: Set timer t1 to collect the topology information
4: Set timer t2 to collect the data message
5: t ← The running time of each round
6: t3 ← T \times r_{max}
7: Broadcast a topo_rq message
8: while t do
9: Wait until an interrupt occur
10: if Timer t1 runs out then
11: Establish two-way secure channels
12: Using algorithm 3 compute and deliver the flow_table
13: Set node state as topology maintain phase
14: else if timer t2 runs out then
15: if Residual energy of the entire network satisfies the inequality (6) and (7)
16: Set node state as topology discovery phase
17: end if
18: else if timer t3 runs out then
19: Set node state as topology discovery phase
20: else if On receiving a message then
21: switch Type of the message
22: case: On receiving the topo_rq message
23: Drop the topo_rq message
24: case: On receiving the topo_rp message
25: Update the topology information module which includes the location and the residual energy of the nodes
26: case: On receiving the topo_repair message
27: Set node state as topology discovery phase
28: case: On receiving the data packet
29: Update the energy information in the topology information module
30: end switch
31: end if
32: end while
```

Algorithm 2: Process of SDFMUCR protocol in sensor nodes.

```plaintext
1: Initialization
2: Set node state as topology discovery phase
3: flagCH ← 0
4: t ← The running time of each round
5: Get the location and residual energy of node
6: while t do
7: Wait until an interrupt occur
8: if On receiving a message then
9: switch Type of the message
10: case: On receiving the topo_rq message
11: if The first time receive topo_rq message then
12: Set node state as topology discovery phase
13: Update the next hop node
14: Establish a one-way secure channel
15: Response topo_rp message
16: Rebroadcast the topo_rq message
17: else
18: Drop the topo_rq message
19: end if
20: end case
21: Forward topo_rp message to the next hop
22: case: On receiving the flow_table
23: Response the ACK message
24: Update flagCH, the precursor list, the TDMA schedule list and the next hop
25: if flagCH = 1 then
26: Forward flow_table to next hop
27: Generate TDMA schedule for precursor nodes
28: Update the precursor list
29: Set node state as topology maintain phase
30: Reset timer t
31: else
32: Transmit the data packet to the next hop according to the TDMA
33: end if
34: Forward the flow_table to the nodes in the precursor list
35: case: On receiving the topo_repair message
36: Forward topo_repair message to the next hop
37: case: On receiving the agg_data packet
38: Forward agg_data packet to the next hop
39: end switch
40: else if Timer t runs out then
41: if precursor list has changed then
42: Send topo_repair message to the next hop
43: end if
44: Aggregate intra cluster data
45: Send the agg_data packet to the next hop
46: Reset timer t
47: end if
48: end while
```
sensor nodes transmit the topo_repair message. The temporary secure channel consists of one-way secure channel and two-way secure channel. The establishment of secure channel is given as follow.

In the topology discovery phase, sensor nodes establish the one-way secure channel to controller when receiving the topo_rq message. The next hop is the node which transmits this topo_rq message. Then local node responses the topo_rp message to controller through this secure channel. So far, the one-way secure channel has been established.

In the flow table generation and delivery phase, the controller uses the information of the whole network to calculate the optimal secure channels to sensor nodes. The secure channel aims at reliable, security, and less energy consumption. Then controller module establishes the two-way secure channels to each sensor nodes. Next, controller delivers the flow_table hop to hop through these two-way secure channels. When sensor nodes receive the flow_table, they will use the two-way secure channel to response the ACK message to controller and forward the flow_table to next hop. So far, the one-way secure channel has been established.

In the data transmission and topology maintain phase, the controller will re-calculate the two-way secure channels for the whole network when receiving the topo_repair message from sensor nodes.

The proposed framework designs the flow_table as simple as possible. So the overheads of the control message are low. In WSNs, the topology of the whole network is not stationary which makes the secure channels not absolute secure. So the secure channels should be used as less as possible.

**Fuzzy-based multi-hop unequal cluster routing**

The detail of the SDFMUCR will be discussed in this section. The architecture of the SDFMUCR is shown in Figure 4. The proposed routing protocol includes the intracluster routing and the intercluster routing. They are both calculated in the controller. This mechanism reduces the calculation pressure and energy consumption in sensor nodes.

Fuzzy-based unequal cluster routing is used to calculate the intracluster routing. The intracluster routing includes the CHs selection process and the cluster formation process. The FL system for intracluster routing is shown in Figure 5. The inputs to the FL system are relative residual energy, relative intercluster cost, and relative intracluster cost. These three parameters are used to calculate the CH probability and the competition radius.

**Relative residual energy.** The responsibilities of the CHs in WSNs are to receive and aggregate the sensor data from the CMs and transmit the agg_data packets to the controller or the relay CH. So its energy consumption is heavier than CMs. Therefore, the node with high residual energy should have a high CH probability and a large competition radius. The relative residual energy is given by equation (8)

\[
E_{re}^{r} = \frac{E_{re} - E_{min}}{E_{max} - E_{min}} \tag{8}
\]

$E_{re}^{r}$ is the relative residual energy. $E_{re}$ is the residual energy of the node. $E_{min}$ and $E_{max}$ are the minimum and the maximum residual energy of the entire network, respectively. These two parameters are learned from the control module.
Before introducing the relative intercluster cost and relative intracluster cost, we assume node $i$ is the CH and analyze the total energy consumption of its cluster in each round. The total energy consumption is given as equation (9)

$$E_{\text{total}} = E_{\text{CH}} + \sum_{j \in N_{\text{non-CH}}} E_{\text{non-CH}}$$  \hspace{1cm} (9)

where $N_{\text{nei}}$ is the set of neighbor nodes of node $i$. Neighbor node means the distance to node $i$ which is less than the node communication radius. And according to equation (1)–(5) which are defined in section “Energy model.” We can get the equation (10)

$$E_{\text{total}} = E_{\text{CH}} + \sum_{j \in N_{\text{non-CH}}} E_{\text{non-CH}}$$

$$= E_{\text{TX}}(l \times N \times \text{agg}, d_c) + N \times E_{\text{Rx}}(l) + E_{\text{DA}}(N) + \sum_{j \in N_{\text{non-CH}}} E_{\text{TX}}(l, d_{\text{non-CH}})$$

$$= (2N - 1 + N \times \text{agg}) \times l \times E_{\text{elec}} + N \times E_{\text{agg}} + l \times E_{\text{fj}} \times \left( N \times \text{agg}d_c^2 + \sum_{j \in N_{\text{nei}}} d_{ij}^2 \right)$$  \hspace{1cm} (10)

where $N$, $l$, $E_{\text{elec}}$, $E_{\text{agg}}$, $\text{agg}$, and $E_{\text{fj}}$ are constants. And only the $d_c^2$ and $\sum_{j \in N_{\text{nei}}} d_{ij}^2$ can influence the $E_{\text{total}}$. $d_c$ is the distance between node $i$ and controller. $d_{ij}$ is the distance between node $i$ and node $j$. The purpose in our protocol is to minimum the “$E_{\text{total}}.”$ So we define $d_c^2$ as the intercluster cost and $\sum_{j \in N_{\text{nei}}} d_{ij}^2$ as the intra cluster cost. The relative intercluster and the relative intracluster costs are defined as follows.

**Relative intercluster cost.** Relative intercluster cost $C_{\text{inter}}$ is the normalization of the intercluster cost which given as (11). $d_{\text{max}}$ is the maximum distance between controller and sensor nodes which is learned from control module. For CHs, the intercluster cost depends on the distance to controller. In a multi-hop routing, the nodes near to controller also undertake the relay task that consume more intercluster cost. So the sensor nodes with low $C_{\text{inter}}$ should have low CH probability and small competition radius

$$C_{\text{inter}} = \frac{d_c^2}{d_{\text{max}}^2}$$  \hspace{1cm} (11)

**Relative intra cluster cost.** Relative intra cluster cost $C_{\text{intra}}$ is the normalization of the intra cluster cost which given as (12). $R$ is the communication radius of the sensor nodes. The nodes with high $C_{\text{intra}}$ increase the average energy consumption within the cluster. Hence, the sensor nodes with high $C_{\text{intra}}$ should have low CH probability and small competition radius

$$C_{\text{intra}} = \frac{\sum_{j \in N_{\text{nei}}} d_{ij}^2}{(N - 1)R^2}$$  \hspace{1cm} (12)

**CH probability.** CH probability is the chance to be a CH. It is a decimal between 0 and 1. The nodes with high CH probability are more likely to be selected as the CH.

**Competition radius.** Competition radius is the broadcast radius in the CHs selection process. The nodes with higher competition radius are more likely to be selected as the CHs.

The crisp values input to the FL system should be converted into a fuzzy linguistic variable. And the handle of the input and output variables are given as follows.

The first input variable “relative residual energy” has three fuzzy linguistic variables. They are low (L), medium (M), and high (H). Low and high have trapezoidal membership function and medium has the triangular membership function. The membership functions base on the related work and our simulation results.

The fuzzy set for “relative residual energy” is shown in Figure 6.

The second input variable “relative intercluster cost” has three fuzzy linguistic variables. They are low (L), medium (M), and high (H). Low and high have trapezoidal membership function and medium has the triangular membership function. The fuzzy set for “relative intercluster cost” is shown in Figure 7.

The third input variable “relative intracluster cost” has three fuzzy linguistic variables. They are low (L), medium (M), and high (H). Low and high have
trapezoidal membership function and medium has the triangular membership function. The fuzzy set for "relative intra cluster cost" is shown in Figure 8.

The first output variable "CH probability" has 11 fuzzy linguistic variables. They are very very low (VVL), very low (VL), low (L), rather low (RL), low medium (LM), medium (M), high medium (HM), rather high (RH), high (H), very high (VH), and very very high (VVH). VVL and VVH have trapezoidal membership function and other variables have the triangular membership function. The fuzzy set for "CH probability" is shown in Figure 9.

The second output variable "competition radius" has 11 fuzzy linguistic variables. They are very very small (VVS), very small (VS), small (S), rather small (RS), medium small (MS), medium (M), medium large (ML), rather large (RL), large (L), very large (VL), and very very large (VVL). VVS and VVL have trapezoidal membership function and other variables have the triangular membership function. The fuzzy set for "competition radius" is shown in Figure 10.

In the CHs selection process, the control module learns the location and residual energy from topology information module first. Then, the control module calculates the relative residual energy, relative intercluster cost, and relative intracluster cost for each sensor nodes according to equations (8), (11), and (12). These three input variables fuzzify into linguistic variables. The linguistic variables are processed by fuzzy if-then rules which are developed by Mamdani method shown in Table 2. The total of the fuzzy if-then rules are 27 and center of area (COA) method is used to de-fuzzify the output linguistic variables into crisp variables. So far, controller generates the CH probability and competition radius for each sensor node.

Next, each sensor node compares the CH probability to other sensor nodes which are among the competition radius. The node which has the highest CH probability
is selected as the final CH; otherwise it quits the CH selection. This process decides the final CH and is operated in the control module because the control module knows the location of each node. In this process, the controller also acts as the CH. Hence, the non-CH nodes which are closer to the controller will select controller as its CH.

Finally, non-CHs select its CH. And the selected CH should be in its communication range. It considers three factors, the distance to CH $d_{CH}$, the residual energy of CH $E_{res}^{CH}$, and the number of neighbor nodes of CH $N_{net}$. $N_{net}^{max}$ is the number of neighbor nodes of the CH which has the max $N_{net}'$ of the entire network. So the $N_{net}'/N_{net}^{max}$ is the density of the CH. If the number of nodes in a cluster is large, it will increase the intracluster cost of CH. Therefore, the high-density CH should have a low probability to be selected as CH. The probability to select a CH for the non-CH $P_e$ is given as equation (13). Then, the non-CH selects the CH which has the maximum $P_e$. So far, the intracluster routing is finished.

$$P_e = 0.8 \left( 1 - \frac{d_{CH}^2}{R_2^2} \right) + 0.1 \left( \frac{E_{res}^{CH}}{E_{max}} \right) + 0.1 \left( 1 - \frac{N_{net}'}{N_{net}^{max}} \right)$$  \tag{13}

The purpose of the first part in this equation is to reduce the energy consumption for non-CH. The purpose of the last two parts is to reduce the energy consumption for CH. The energy consumption in transmit is higher than receipt. So the first part is the most important and need to have a high weight. Based on the experience, the weight is 0.8, 0.1, and 0.1.

A FL-based intercluster routing is proposed to reduce the intercluster energy consumption. Distance threshold $d_{th}$ is defined in this process. A CH whose distance to the controller is greater than $d_{th}$ should find a relay CH as its next hop. The relay CH should satisfy the inequalities (14), (15), and (16). $E_r$ is the residual energy of the relay CH. $d_{re}$ is the distance between the relay CH and controller. $d_r$ is the distance to the relay node.

$$E_{re} > E_{th} \tag{14}$$

$$d_{re} < d_c \tag{15}$$

$$d_r < d_{th} \tag{16}$$

It selects set of relay CHs through above process. Then, we use the FL system to calculate the join CH chance. The two input parameters to the FL system are relative residual energy and relative relay cluster cost.

Table 2. Fuzzy if-then rules for intracluster routing.

| No. | Relative residual energy | Relative intercluster cost | Relative intracluster cost | CH probability | Competition radius |
|-----|--------------------------|----------------------------|---------------------------|----------------|--------------------|
| 1   | L                        | L                          | LM                        | S              |
| 2   | L                        | L                          | RL                        | VS             |
| 3   | L                        | H                          | L                         | VVS            |
| 4   | L                        | M                          | L                         | VLS            |
| 5   | L                        | M                          | M                         | MS             |
| 6   | L                        | M                          | H                         | VL             |
| 7   | L                        | H                          | M                         | VL             |
| 8   | L                        | H                          | H                         | VL             |
| 9   | M                        | L                          | H                         | RL             |
| 10  | M                        | L                          | M                         | RL             |
| 11  | M                        | L                          | H                         | VH             |
| 12  | M                        | H                          | M                         | VH             |
| 13  | M                        | H                          | M                         | VVH            |
| 14  | M                        | M                          | M                         | LS             |
| 15  | M                        | M                          | H                         | VS             |
| 16  | M                        | H                          | H                         | VVH            |
| 17  | M                        | L                          | M                         | VH             |
| 18  | M                        | L                          | M                         | VVL            |
| 19  | M                        | M                          | M                         | VL             |
| 20  | M                        | M                          | H                         | VH             |
| 21  | M                        | H                          | H                         | VH             |
| 22  | M                        | H                          | H                         | VH             |
| 23  | M                        | H                          | H                         | VH             |
| 24  | M                        | M                          | H                         | VL             |
| 25  | M                        | H                          | M                         | VVL            |
| 26  | M                        | H                          | M                         | VL             |
| 27  | M                        | H                          | H                         | HL             |

CH: cluster head; LM: low medium; RL: rather low; VS: very small; VVS: very very small; VL: very low; RS: rather small; VVL: very very low; RH: rather high; HM: high medium; VVH: very very high; MS: medium small; ML: medium large.
The equation and fuzzy set of “relative residual energy” are given as equation (8) and Figure 6.

Relative relay cluster cost. Relative relay cluster cost $C_{relay}$ is the normalization of the relay cluster cost which is given as equation (17). The nodes with high $C_{relay}$ increase the intercluster energy consumption of CH. So the sensor nodes with high $C_{relay}$ should have low join CH chance

$$C_{relay} = \frac{d_r^2 + d_{rc}^2}{2d_{th}^2} \quad (17)$$

Relative relay cluster cost has three fuzzy linguistic variables. They are low (L), medium (M), and high (H). Low and high have trapezoidal membership function and medium has the triangular membership function.

The fuzzy set for “relative relay cluster cost” is shown in Figure 11.

Join CH chance is the probability of a non-CH node join a CH. The output variable “join CH chance” has five fuzzy linguistic variables. They are very low (VL), low (L), medium (M), high (H), and very high (VH). VL and VH have trapezoidal membership function and other variables have the triangular membership function. The fuzzy set for “join CH chance” is shown in Figure 12.

In the intercluster routing process, the control module selects the CHs which need relay to the controller. Then, CHs generate the set of relay CHs which satisfy the inequalities (14), (15), and (16). Next, control module calculates the relative residual energy and relative relay cluster cost for each node in the set of relay CHs. These two input variables are fuzzified into linguistic variables. The linguistic variables are processed by fuzzy if-then rules which are developed by Mamdani method shown in Table 3. The total of the fuzzy if-then rules are 9 and COA method is used to de-fuzzify the output linguistic variables into crisp variables. So far, it generates the “join CH chance” for CHs. The CHs select the relay CH which has the maximum “join CH chance” as its next hop. So far, the process of the intercluster routing has finished.

The media access control (MAC) mechanism of the proposed protocol is given as follow. SDFMUCR consists of the normal channel and the secure channel. These two channel work on the different frequencies which reduces the collision of the control message. Sensor nodes contain two modes: sleep mode and wakeup mode. In the topology discovery phase, sensor nodes work on the wakeup mode. The topo_rq message is transmitted in the normal channel and the topo_rp message is transmitted in the secure channel. In the data transmission phase, CHs allocate the time slots to the CMs according to TDMA. CMs awake and

![Figure 11. Fuzzy set for “relative relay cluster cost.”](image1)

![Figure 12. Fuzzy set for “join CH chance.”](image2)

Table 3. Fuzzy if-then rules for intercluster routing.

| No. | Relative residual energy | Relative relay cluster cost | Join CH chance |
|-----|--------------------------|------------------------------|----------------|
| 1   | Low                      | Low                          | Medium         |
| 2   | Low                      | Medium                       | Low            |
| 3   | Low                      | High                         | Very Low       |
| 4   | Medium                   | Low                          | High           |
| 5   | Medium                   | Medium                       | Medium         |
| 6   | Medium                   | High                         | Low            |
| 7   | High                     | Low                          | Very High      |
| 8   | High                     | Medium                       | High           |
| 9   | High                     | High                         | Medium         |

CH: cluster head.
transmit the data packets to CH in its own time slot. The CMs are in the sleep mode at other time slots. Then, CHs receive and aggregate the data packet and transmit the agg_data packets to the next hop until the agg_data packets arrive at the controller. In this phase, the CHs always in the wake mode. The process of the forwarding rules generation in controller is shown in Algorithm 3.

The flowchart of SDFMUCR protocol is given as Figure 13. In the topology discovery phase, controller broadcasts the topo_rq message. When a sensor node receives the topo_rq message, it establishes the two-way secure channels, responds the topo_rp message, and rebroadcasts the topo_rq message. Once controller receives all topo_rp message of the network, it uses the fuzzy-based multi-hop unequal cluster routing mechanism to calculate the routing of the entire network. Controller also works as CH in the proposed protocol. Controller calculates the $E_{\text{re}}$, $C_{\text{inter}}$, and $C_{\text{intra}}$ and then uses fuzzy system to calculate competition radius $C_R$ and CH probability $CH_p$. A node which has the highest $CH_p$ within $C_R$ will be selected as CH. CMs select CH which has the highest $E_{\text{re}}$. If a CH’s distance to controller is less than $R$, the next hop of itself is controller. Otherwise, it finds the nodes which satisfy equations (14)–(16) as the relay node set. Join CH chance is calculated by fuzzy system. The next hop is CH which has the highest $CH_p$. Finally, sensor nodes add itself to the precursor list of next hop and the flow table is generated.

In the flow table delivery phase, the controller delivers the flow table. When a node receives the flow table, it updates the $CH_{\text{flag}}$, precursor list, the TDMA schedule list and the next hop. CH will forward the flow table to the nodes which are in the precursor list. In the data transmission and topology maintain phase, CMs transmit data packet to CH and CHs aggregate the data packets and transmit the agg_data packet to the next hop. If a node’s residual energy satisfies the conditions to re-calculate the routing, it sends the topo_repair message to controller. The network will return to the topology discovery phase when the controller finds the number of alive node change, the network has run more than $r_{\text{max}}$ or receives the topo_repair message.

### Theoretical analysis of the protocol

#### Time complexity

The time complexity of SDFMUCR is given as follow. We define $n$ as the number of sensor nodes in the entire network. First, controller processes the topology information of each sensor nodes through FL system. Next, it determines which nodes are CHs. The time complexity is $O(n)$. Then, controller calculates the intracluster routing. The number of CHs in the network is $C$ and the time complexity is $O(C(n-C)) = O(Cn - C^2)$. Finally, controller calculates the intercluster routing. The time complexity is $O(C^2)$. So the time complexity is $O(n + Cn - C^2 + C^2) = O(n)$.

#### Memory complexity

The flow table includes the CH flag ($CH_{\text{flag}}$), precursor list, TDMA schedule list, and next hop. So the memory complexity of SDFMUCR is $O(n)$.

---

**Algorithm 3: Process of the forwarding rules generation.**

1: $N \leftarrow$ Total sensor number of node in network  
2: $c \leftarrow$ The sequence number of controller  
3: $S = \{1, 2, \ldots, N\}$  
4: $CH = CM = \phi$  
5: $Pe_j = J_{\text{CH}} = \phi$  
6: for $i = 1 \rightarrow N$ do  
7:     Using equation (8), (11), (12) calculate $E_{\text{re}}$, $C_{\text{inter}}$  
8:         and $C_{\text{intra}}$  
9:         $(C_R, CH_p) = \text{fuzzy}(E_{\text{re}}, C_{\text{inter}}, C_{\text{intra}})$  
10: end for  
11: for $i = 1 \rightarrow N$ do  
12:     if $\exists j : S_{d_j < C_R} * CH_p_i > CH_p_j$  
13:         $i_{\text{flagCH}} = 0$  
14:     else  
15:         $i_{\text{flagCH}} = 1$  
16:     end if  
17:     $CH = CH \cup \{i\}$  
18:     end if  
19: end for  
20: for $i = CM$ do  
21:     for $j = CH$ do  
22:         Using (13) calculate  
23:         $P_{\text{e}} = P_{\text{e}} \cup \{P_{\text{e}_j}\}$  
24:     end for  
25:     $\text{next hop} = \{j : CH \cap P_{\text{e}_j} = \max(P_{\text{e}})\}$  
26:     end for  
27: $\text{pre list}_i = \phi$  
28: for $i = CH$ do  
29:     $\text{pre list}_i = \phi$  
30:     $\text{flagCH}_i = \phi$  
31: end if  
32: end for  
33: if $d_i < R$ then  
34:     $\text{pre list}_i = \text{pre list}_i \cup \{i\}$  
35: else  
36:     $\text{flagCH}_i = CH$  
37:     if $i$ satisfy inequality (14), (15) and (16) then  
38:         Using (17) calculate $C_{\text{relay}}$  
39:         $J_{\text{CH}} = \text{fuzzy}(E_{\text{re}}, C_{\text{relay}})$  
40:         $J_{\text{CH}} = J_{\text{CH}} \cup \{J_{\text{CH}_i}\}$  
41:     end if  
42:     $\text{next hop}_i = \{j : CH \cap \text{next hop}_i = \max(J_{\text{CH}})\}$  
43:     end if  
44: end for  
45: end if  
46: end for  
47: end if  
48: end for  
49: return flow_table
Control overhead

Control overhead is the average number of the control messages to be transmitted in the network within a round. In the topology discovery phase, the nodes within the transmission range of controller response the topo_rp message to controller directly. Other nodes response the topo_rp by multi-hops. We define $N_t$ as the number of nodes within the transmission range and $k$ as the maximum hops of the entire network. The maximum number of the control message in the topology discovery phase is $N_t + k(n - N_t) = kn + (k - 1)N_t$. In the flow table generation and delivery phase, CHs
need to response the ACK message and retransmit the flow table to the precursor nodes. CMs only need to response the ACK message. The number of the control message in the flow table generation and delivery phase is \((n - C) + 2C = n + C\). Therefore, the maximum number of the control message is \(C_m = kn + (k - 1)N_r + n + C\). In the proposed protocol, it only need control message to re-calculate the routing when the network satisfy the conditions proposed in section “residual energy based rotating scheme.” So the control overhead of SDFMUCR is between \(C_m/r_{max}\) and \(C_m\).

**Simulation results**

The proposed SDFMUCR protocol is compared with the EAUCF,23 FUCA,7 and HEEC33 under four scenarios. EAUCF and FUCA are the distributed FL-based unequal clustering protocols for WSNs. HEEC is the multi-hop-based unequal clustering protocol for WSNs. The different network scenarios are shown in Figure 14 and the simulation parameters are given as Table 4.

The nodes are random distributed in the 100 m \(\times\) 100 m location area. The initial energy of each node is 0.5 J. The length of the control message is 32 bits and the length of the data packet is 4000 bits. The data aggregation rate in the CHs is 0.1. The controller is located in the center of network in Scenarios 1 and 2. In Scenarios 3 and 4, the controller is located in the corner of network. The numbers of the deployed nodes in four scenarios are 100, 200, 100, and 200, respectively.

We use the MATLAB to simulate and evaluate the proposed protocol and other well-known protocols. Each scenario is simulated 50 times to reduce error in simulation. Five metrics are considered as the performance parameters to evaluate the routing protocols. They are the TRE, the lifetime of the network, the number of alive nodes of communication round, control overhead of the network, and the latency of the sensor nodes. In each scenario, the average of the simulated results is shown as follows.

**Lifetime and energy consumption of the network**

One goal in our research is to reduce the energy consumption of the entire network. If the residual energy of the sensor node reduces slowly, the nodes will have more time for collecting and transmitting the data to the BS. It prolongs the lifetime of the network which makes the protocol more suitable for deployment in harsh environments. For the lifetime of the network, first node die is as the metric to evaluate it. The number of alive nodes of communication round represents the trend of the number of alive nodes in the network. With the decrease in the number of alive node, the nodes which are far from BS will have a hard time to transmit the data packet to the BS. So when a node dies, some of the alive nodes may stop working because there is no routing to the BS. So we define the FND as the lifetime of the network, and the number of alive nodes of communication round reflects the connectivity of the network.

Table 5 and Figure 15 show the lifetime for all protocols. They express that the proposed protocol performs well in each scenario. The lifetime in each scenarios are 1144 rounds, 1140 rounds, 714 rounds, and 784 rounds. The details for all protocol simulation results in each scenario are shown in the subsequent paragraphs.

In Scenario 1, the controller is located in the center of the network and 100 sensor nodes are deployed in the area. The total of remain energy of communication rounds is shown in Figure 16. In simulation result, SDFMUCR consumes energy slower than other protocols which indicates the proposed protocol reduces the energy consumption.
energy consumption in whole network. The number of alive nodes of the communication rounds is shown in Figure 16. For the lifetime of network, the SDFMUCR performs better than EAUCF by 222.2%, FUCA by 175% and HEEC by 23.9%. In EAUCF, FUCA, and HEEC, CM nodes transmit the packet to CH and then CH aggregates and transmits the agg_data packet to BS. In SDFMUCR, the sensor nodes which are near to controller transmit the data packets to controller directly. This method reduces the energy consumption of aggregation, reception, and transmission in CH nodes which are closer to controller. Meanwhile, SDFMUCR proposes the residual energy–based rotating scheme and this scheme reduces the average control overhead and energy overhead in each round. Finally, SDN architecture is used in this protocol and reduces the computation overhead in sensor nodes.

In Scenario 2, the position of the controller remains unchanged and the node density is twice than that in Scenario 1. The simulation results are shown in Figures 18, 19 and Table 5, which are nearly the same as Scenario 1 shown in Figures 17 and 18. Therefore, the node density does not influence the performance of the proposed protocol. For the lifetime of network, the SDFMUCR performs better than EAUCF by 208.1%, FUCA by 210.6%, and HEEC by 32.9%.

In Scenario 3, the controller is located in the corner of the network and the number of deployed nodes is 100. The simulation results are shown in Figures 20 and 21 and Table 5. It can be seen that the proposed protocol performs better than the comparison algorithms in energy consumption. Meanwhile, the SDFMUCR performs better than EAUCF by 142.0%, FUCA by 59.4%, and HEEC by 50.3%. For SDFMUCR, CHs find the optimum CH as its next hop if necessary. This way avoids the large energy consumption for the CHs which are far away from the controller. The EAUCF and FUCF do not have the scheme for relay, so their performance is worse than SDFMUCR. In HEEC, distributed intercluster routing increases the control

### Table 5. The results of the FND for all protocols (round).

| Protocol     | EAUCF | FUCA | HEEC | SDFMUCR |
|--------------|-------|------|------|---------|
| Scenario 1   | 355   | 416  | 923  | 1144    |
| Scenario 2   | 367   | 370  | 858  | 1140    |
| Scenario 3   | 295   | 448  | 475  | 714     |
| Scenario 4   | 153   | 180  | 719  | 784     |

EAUCF: energy-aware unequal clustering algorithm; FUCA: fuzzy-based unequal clustering algorithm; HEEC: hybrid unequal energy-efficient clustering; SDFMUCR: software-defined and fuzzy-based multi-hop unequal cluster routing.

**Figure 15.** FND of the different protocol with different scenario.

**Figure 16.** Scenario 1: Total remain energy of communication rounds.

**Figure 17.** Scenario 1: Number of alive nodes of communication rounds.
overhead which increases the energy consumption of the entire network. This simulation makes sure that the controller position does not influence the performance of network. The designed protocol satisfies the scalability needs in WSNs. In SDFMUCR, it does not have the scheme for select the candidate CHs. All the sensor nodes participate in the selection for final CHs. Comparing to the protocols which generate the random number to select the CHs or candidate CHs. This way avoids the nodes which have low residual energy are selected as the final CHs and extends the lifetime of the network.

In Scenario 4, the location of the controller is as same as Scenario 3 and the deployed nodes are 200. The simulation results are shown in Figures 22 and 23 and Table 5. In Scenario 4, the SDFMUCR also performs well in energy consumption and the lifetime of the network. The SDFMUCR performs better than EAUCF by 412.4%, FUCA by 335.6%, and HEEC by 9.0% for the lifetime of the network. This simulation results show that the position of the controller and the network density do not influence the performance of the proposed protocol.

According to the simulation results, SDFMUCR performs better than the comparison protocols. It reduces the energy consumption and prolongs the network lifetime. First, the proposed SDN framework reduces the control overhead and calculation overhead of the sensors. Second, residual energy-based rotating scheme reduces the average number of control message transmitted in the network. Third, fuzzy-based unequal cluster routing and multi-hop scheme balance the
energy consumption of the intercluster routing which prolongs the lifetime of the entire network. Fourth, CMs join the CH based on the distance to CH, residual energy of CH, and the density of CH. This mechanism balances the energy consumption of the intracluster routing. Fifth, controller also works as CH which balances the energy of the node which is near to controller.

In the above reasons, the proposed SDN framework plays an important role in reducing the energy consumption of the entire network. Three reasons are listed as follows. First, SDFMUCR reduces the number of the control message. For sensor nodes, it only needs to transmit the topo rq message, topo rp message, and flow_table to establish the routing of the entire network. However, the distributed routing protocols need many control message for intercluster routing and more control message for intracluster routing which increases the control message and increases the energy consumption. Second, SDFMUCR transmits the control message through two channels which reduces the collision of the message, reduces the number of times to retransmit the message and reduces the energy consumption. Third, instead of calculating the routing in the sensor nodes, SDFMUCR generates the routing in the controller which reduces the calculation overhead in sensor nodes. So the proposed SDN framework in SDFMUCR reduces the energy consumption in the network.

Control overhead
The control overhead for all protocols is given as Table 6. Control overhead is the average number of the control messages to be transmitted in the network within a round. If the protocol’s control overhead is low, the energy consumption for control message and the time for generating the routing is low. According to simulation, SDFMUCR performed well than all comparison protocols in each scenario. The control overhead of EAUCF and FUCA is relatively low because they do not have the mechanism to relay which reduces the number of control message in the entire network. HEEC has the highest control overhead because the relay mechanism increases the number of control message of the network. SDFMUCR proposes the residual energy–based rotating scheme. It only need control message to re-calculate the routing when satisfy the conditions to re-establish the routing. This mechanism reduces the average control message in each round and the control overhead is very low.

Latency
The latency for all protocols is given as Figure 24. Latency is the average time between when the data were sensed and when the data were transmitted to BS for a sensor node. If the latency is low, the network will be more suitable for deployment scenarios with low latency requirements. EAUCF and FUCA almost have

![Figure 22. Scenario 4: Total remain energy of communication rounds.](image)

![Figure 23. Scenario 4: Number of alive nodes of communication rounds.](image)

| Table 6. Control overhead for all protocols. |
|-----------------|----------|---------|---------|
| Scenario 1    | EAUCF    | FUCA    | HEEC    | SDFMUCR |
| Scenario 2    | 107      | 109     | 185     | 63      |
| Scenario 3    | 216      | 219     | 368     | 214     |
| Scenario 4    | 101      | 110     | 187     | 80      |
| Scenario 4    | 205      | 209     | 405     | 197     |

EAUCF: energy-aware unequal clustering algorithm; FUCA: fuzzy-based unequal clustering algorithm; HEEC: hybrid unequal energy-efficient clustering; SDFMUCR: software-defined and fuzzy-based multi-hop unequal cluster routing.
the same latency because they do not have the relay mechanism and the intracluster routing is similar for these two protocols. HEEC has the relay mechanism in order to balance the energy of the sensor nodes. It needs more time to relay the aggregate data packet, so the latency of the HEEC is the most highest. In SDFMUCR, controller also works as a CH in the network. The nodes which are closer to controller can transmit the sensor data to controller directly, so the latency of the SDFMUCR is low.

Conclusion

In this article, the proposed protocol combines SDN architecture with FL-based unequal cluster routing. The intra and intercluster routing are calculated in the control module that reduces calculation and energy overhead in sensor nodes. Meanwhile, all nodes participate in the CHs selection process which avoids the low residual energy nodes that are selected as the CHs. Simultaneously, the residual energy–based rotating scheme is proposed to reduce the control message in each round. Finally, FL-based unequal cluster routing and multi-hop intercluster routing are proposed to ensure the optimum performance of the designed algorithm. Via simulation, SDFMUCR performs better than the well-known algorithms, such as EAUCF, FUCA, and HEEC. The proposed protocol reduces the energy consumption, prolongs the lifetime of the network, balances the energy consumption, reduces the control overhead, and reduces the latency of the network. The proposed SDWSN framework and protocol can be used in the barren areas, forests, or farms for a long time to monitor environment variables. Further studies will focus on the design of the routing with mobile BS.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the National Natural Science Foundation of China (NSFC) (grant no. 61872401).

ORCID iD

Lu Sixu https://orcid.org/0000-0003-2908-0174

References

1. Sun X. Discussion of networking mode and application technology on the internet of things. In: 2010 Asia Pacific youth conference on communication (APYCCCT), Kunming, China, 7–8 August 2010. Wuhan, China: Scientific Research Publishing.
2. Mazumdar N and Om H. Distributed fuzzy logic based energy-aware and coverage preserving unequal clustering algorithm for wireless sensor networks. Int J Commun Syst 2017; 30(13): e3283.
3. Mazumdar N and Om H. A GA-based fault-aware routing algorithm for wireless sensor networks. In: De D, Mukherjee A, Kumar Das S, et al. (eds) Nature inspired computing for wireless sensor networks. Singapore: Springer, 2020, pp.21–38.
4. Chen L, Dong Y, Xu Z, et al. An adaptive routing strategy for cluster-based wireless sensor networks. In: 2015 27th Chinese control and decision conference, Qingdao, China, 23–25 May 2015. New York: IEEE.
5. Kim E. Cluster based routing method in a wireless sensor network. In: International conference on grid and distributed computing (GDC), Jeju Isl, South Korea, 25–28 November 2015. New York: IEEE.
6. Li P and Yan J. A cluster-based routing tree construction algorithm for wireless sensor networks. In: Information technology for manufacturing systems (ITMS), Shanghai, China, 7–8 May 2011. Dürnten-Zurich, Switzerland: Trans Tech Publications Ltd.
7. Agrawal D and Pandey S. FUCA: fuzzy-based unequal clustering algorithm to prolong the lifetime of wireless sensor networks. Int J Commun Syst 2018; 31(2): e3448.
8. Adhikary DRD and Mallick DK. A congestion aware, energy efficient, on demand fuzzy logic based clustering protocol for multi-hop wireless sensor networks. Wireless Pers Commun 2016; 97(1): 1445–1474.
9. Purkait R and Tripathi S. Energy aware fuzzy based multi-hop routing protocol using unequal clustering. Wireless Pers Commun 2017; 94(3): 809–833.
10. Neamatollahi P and Naghibzadeh M. Distributed unequal clustering algorithm in large-scale wireless sensor networks using fuzzy logic. J Supercomput 2018; 74(6): 2329–2352.
11. Neamatollahi P, Naghibzadeh M and Abrishami S. Fuzzy-based clustering-task scheduling for lifetime enhancement in wireless sensor networks. *IEEE Sens J* 2017; 17(20): 6837–6844.

12. Mostafaei H and Menth M. Software-defined wireless sensor networks: a survey. *J Netw Comput Appl* 2018; 119: 42–56.

13. Musa N, Gerhard H and Adnan AM. Software defined networking for improved wireless sensor network management: a survey. *Sensors* 2017; 17(5): 1031.

14. Kobo HI, Abu-Mahfouz AM and Hancke GP. A survey on software-defined wireless sensor networks: challenges and design requirements. *IEEE Access* 2017; 5: 1872–1899.

15. Luo T, Tan HP and Querk TQS. Sensor OpenFlow: enabling software-defined wireless sensor networks. *IEEE Commun Lett* 2012; 16(11): 1896–1899.

16. Bera S, Misra S, Roy SK, et al. Soft-WSN: software-defined WSN management system for IoT applications. *IEEE Syst J* 2018; 12(3): 2074–2081.

17. Liao W, Wu M and Wu Y. Design of multi-energy-space-based energy-efficient algorithm in novel software-defined wireless sensor networks. *Int J Distrib Sens N* 2017; 13(7): 1–18.

18. Li P, Wu M, Liao W, et al. A game-theoretic and energy-efficient algorithm in an improved software-defined wireless sensor network. *IEEE Access* 2017; 5: 13430–13445.

19. Aparicio J, Echevarria JJ and Legarda J. A software defined networking approach to improve the energy efficiency of mobile wireless sensor networks. *KSII T Internet Inf* 2017; 11(6): 2848–2869.

20. Wei X, Yuan Z and Ning W. An energy-efficient routing algorithm for software-defined wireless sensor networks. *IEEE Sens J* 2016; 16(20): 7393–7400.

21. Abdolmaleki N, Ahmadi M, Malazi HT, et al. Fuzzy topology discovery protocol for SDN-based wireless sensor networks. *Simul Model Pract Th* 2017; 79: 54–68.

22. Mazumdar N, Roy S and Nayak S. A survey on clustering approaches for wireless sensor networks. In: *International conference on data science and business analytics (ICDSBA)*, Changsha, China, 21–23 September 2018. New York: IEEE.

23. Bagci H and Yazici A. An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Appl Soft Comput* 2013; 13(4): 1741–1749.

24. Balakrishnan B and Balachandran S. DUCF: distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach. *Appl Soft Comput* 2016; 40: 495–506.

25. Sert SA, Bagci H and Yazici A. MOFCA: multi-objective fuzzy clustering algorithm for wireless sensor networks. *Appl Soft Comput* 2015; 30: 151–165.

26. Logambigai R and Kannan A. Fuzzy logic based unequal clustering for wireless sensor networks. *Wirel Netw* 2016; 22(3): 945–957.

27. Gajjar S, Sarkar M and Dasgupta KS. FAMACROW: fuzzy and ant colony optimization based combined MAC, routing, and unequal clustering cross-layer protocol for wireless sensor networks. *Appl Soft Comput* 2016; 43: 235–247.

28. Arjunan S and Sujatha P. Lifetime maximization of wireless sensor network using fuzzy based unequal clustering and ACO based routing hybrid protocol. *Appl Intell* 2018; 48(8): 2229–2246.

29. Mazumdar N and Om H. Distributed fuzzy approach to unequal clustering and routing algorithm for wireless sensor networks. *Int J Commun Syst* 2018; 31(12): e3709.

30. Goyal N, Dave M and Verma AK. Energy efficient architecture for intra and inter cluster communication for underwater wireless sensor networks. *Wireless Pers Commun* 2016; 89(2): 687–707.

31. Mazumdar N and Om H. DUCR: distributed unequal cluster-based routing algorithm for heterogeneous wireless sensor networks. *Int J Comput Int Syst* 2017; 30(18): 1–14.

32. Guirguis SK, Abdou MA and Elhaggar AA. A hybrid fuzzy multi-hop unequal clustering algorithm for dense wireless sensor networks. *Int J Comput Int Syst* 2017; 10(1): 951–961.

33. Mostafa BS and Massoud BA. HEEC: a hybrid unequal energy efficient clustering for wireless sensor networks. *Wirel Netw* 2019; 25: 4751–4772.