Applied Soft Computing Journal 89 (2020) 106115

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com/locate/asoc

A new fuzzy multi-hop clustering protocol with automatic rule tuning for wireless sensor networks

Fakhrosadat Fanian [a], Marjan Kuchaki Rafsanjani [b,*]

[a] Department of Computer Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran
[b] Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

ARTICLE INFO

Article history:
Received 1 October 2018
Received in revised form 28 September 2019
Accepted 20 January 2020
Available online 23 January 2020

Keywords:
Wireless Sensor Network (WSN)
Clustering
Shuffle Frog Leaping Algorithm (SFLA)
Multi-hop
Fuzzy Inference System (FIS)

ABSTRACT

In today's world, a major challenge is to conserve and make optimal use of energy. This is a critical matter in wireless sensor networks due to their wide application in different areas. More importantly, scant attention has been paid to the use of node energy for certain applications in such networks. This study used the Shuffled Frog Leaping Algorithm (SFLA) to propose a Fuzzy Multi-hop clustering protocol (FMSFLA). The SFLA is used for automated configuration and optimization of the rule-base table in a fuzzy inference system and five adjustable parameters in two phases, i.e. Cluster Head (CH) selection and parent selection, based on application features. The proposed protocol (FMSFLA) considers effective parameters including energy, distance from the base station (BS), the number of neighboring nodes, real node distance from the BS, mean route load, delay, overlap, and the problem of hot spots, to achieve the best application-based performance. The FMSFLA includes rounds, in each round the phases of CH selection, parent selection, cluster formation, and steady state are performed. In the CH selection phase, CHs are selected from candidate nodes based on the fuzzy output and energy threshold (i.e. a control parameter) with respect to the overlap rate of adjacent CHs. In our protocol, the parent selection phase began by determining the levels of CHs in the network. At the end of this phase, the parent of each CH is determined on the basis of the greatest fuzzy output based on application. In the cluster formation phase, the clusters are formed on the basis of the determined CHs. Finally, the information received by CHs is sent through their parents to the BS in the steady state phase. The FMSFLA is evaluated against the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols in terms of the number of alive nodes, received packets, and cluster heads in addition to their appropriate distribution rates and other parameters pertaining to the network lifetime and protocol scalability using three application-oriented scenarios. According to the simulation results, the FMSFLA functioned far better than the other protocols in all scenarios with respect to goals and application features.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Wireless sensor networks (WSNs) represent a new generation of networks consisting of many inexpensive sensor nodes, which are linked through wireless waves. The goal of WSNs is to collect information from the surrounding environment of network sensors. Energy is a major challenge to WSNs, with numerous methods proposed to reduce the energy of sensor nodes and lengthen the network lifetime due to dischargeable and irreplaceable batteries. Clustering is one of the most efficient methods for this purpose [1–4]. In clustering, sensor nodes are classified into certain groups, called clusters. Every cluster has a cluster head (CH) which is responsible for sending collected information from cluster members to a base station (BS) on single-hop or multi-hop routes through other CHs [5–9]. Finally, information is collected and processed in the BS so that the actual values of the relevant parameters are estimated fairly accurately [10–13]. Clustering offers numerous advantages such as reducing bandwidths, reducing overhead, preventing the redundancy of message exchange among nodes, and implementability of administrative strategies in networks [14–16].

There have been many technological advances in the manufacturing of small nodes, development of wireless sensor networks, and increased applications of such networks in industries [17], medicine [18], military [19,20], agriculture [21], etc. Therefore, such networks have been widely used to increase the welfare and comfort of human life. In this paper, the SFLA is employed to propose a multi-hop fuzzy clustering protocol with the main purpose of reducing network energy and prolonging network lifetime. Given the configurations of the SFLA objective function designed

* Corresponding author.
E-mail addresses: f_fanian@iauk.ac.ir (F. Fanian), kuchaki@uk.ac.ir (M. Kuchaki Rafsanjani).

https://doi.org/10.1016/j.asoc.2020.106115
1568-4946/© 2020 Elsevier B.V. All rights reserved.

on the basis of application, the proposed FMSFLA prolonged the network lifetime based on an application. Then CHs are selected from candidate nodes having the greatest fuzzy outputs based on the overlap rate of adjacent CHs based on application, resulting in the appropriate distribution of CHs in the network. In addition, the redistributor (parent) nodes are selected in the FMSFLA from candidate nodes having the greatest fuzzy outputs based on application. There are also two fuzzy rules base tables developed and optimized by the SFLA prior to the network operations to select CHs and parent nodes along with five parameters controlling the clustering and routing processes. Then they are used in the current network process. The proper distribution of CHs, the right number of CHs, and their appropriate activity radius are determined in the FMSFLA based on application. Finally, they result in a steady workload, reduce the energy consumption, and prolong the network lifetime. In fact, the SFLA determine and optimize the minimum and maximum overlap rates of adjacent CHs and their activity radius based on different scenarios, their distances to the BS, and the threshold determining the energy level of candidate nodes, all regarded as control parameters based on application.

In this paper, Section 2 reviews the related literature. Section 3 presents hypotheses, a network model, an energy consumption model, and the proposed protocol in detail. Section 4 introduces the fuzzy logic model thoroughly with relevant inputs in the FMSFLA. In Section 5, the SFLA is used for describing the details of the FMSFLA optimization process. Section 6 analyzes the complexity of the proposed algorithm. Section 7 evaluates the proposed algorithm. Finally, Section 8 presents the conclusion.

2. Literature review

According to a simple classification, clustering-based routing protocols can be categorized as single-hop and multi-hop methods. Fig. 1 shows the difference between these two categories of methods.

2.1. Single-hop clustering-based routing methods

Classic clustering-based routing methods such as Low Energy-Adaptive Clustering Hierarchy (LEACH) [5] and Threshold-based cluster head replacement-LEACH (T-LEACH) [22] deal with CH selection based on network criteria. In the most basic of such methods, CHs are merely selected randomly based on probability. Thus, nodes with insufficient energy are also likely to be selected as CH nodes, resulting in the early termination of network nodes.

Energy, distance, and node density are the criteria used in LEACH-Energy-based protocol (LEACH-EP) [23], LEACH with Distance-based Thresholds (LEACH-DT) [24], and Density of Sensor-LEACH (DS-LEACH) [25] to balance the energy consumption of nodes and select CHs. In other words, nodes with higher energy levels can be selected as CHs. CH selection can also be based on different probabilities of node distance from the BS or node density. LEACH with Sliding Window Dynamic number of Node (LEACH-SWDN) [26] considers the role of the correct number of CHs in the network in addition to the effect of the optimal number of CHs in reducing energy consumption and prolonging network lifetime. All such methods do not guarantee the uniform distribution of CHs over the network, because they lack a specific parameter or strategy to control the process of distribution CHs. In addition, few parameters are selected to determine the role of CHs. However, there are many parameters affecting the CH role determination. In these methods, no mechanisms are considered for routing or sending information to the BS. In fact, the information is sent from CHs to the BS through single-hop, a method which increases energy consumption and reduces the network lifetime. Hence, these protocols are not proper for large-scale networks. In other words, they are not scalable. Furthermore, fuzzy logic is used for CH selection in many single-hop methods such as LEACH protocol using Fuzzy Logic (LEACH-FL) [27] and Adaptive Multi-Clustering algorithm using Fuzzy Logic (Adaptive MCFL) [28]. In LEACH-FL, every node determines the probability through the fuzzy logic. The resultant probability of every node is then subtracted from one. If the result is smaller than the threshold of this method, the node will be selected as the CH. In addition to the fuzzy logic, many researchers have used meta-heuristic algorithms to select dynamic clusters from sensor nodes in recent years. They have also tried to employ methods such as Heuristic Algorithm for Clustering Hierarchy (HACH) [29], Harmony Search Algorithm Cluster-based Protocol (HSACP) [30], Energy-Aware Evolutionary Routing Protocol (EAERP) [31] and Application Specific Low Power Routing protocol (ASLPR) [32] to minimize energy consumption and increase the network lifetime. An enhanced version of LEACH, named ASLPR, is proposed to prolong the application-based network lifetime. ASLPR focuses on the status information of every node including distance from the BS, residual energy, and distance between CHs. This method combines the genetic algorithm with the simulated annealing algorithm to optimize the parameters used for determining the threshold. Although efficient, ASLPR is fairly complicated due to considering a large number of adjustable parameters which can be adjusted and optimized to achieve certain applications. The function of ASLPR depends on the accurate adjustment of parameters. This method is mainly focused on CH selection process and clustering. It ignores the process of sending information to the BS, protocol scalability for larger networks, and the problem of hot spots. This method sends information to the BS through single-hop and lacked a specific mechanism to control the cluster radius.

However, a combination of fuzzy techniques and meta-heuristic algorithms have been used in a limited number of methods such as Swarm Intelligence based Fuzzy routing protocol (SIF) [33], Centralized cluster-based routing protocol based on Sugeno Fuzzy inference system (LEACH-SF) [34] and Fuzzy Shuffled Frog Leaping Algorithm (FSFLA) [35] in recent years. In SIF and LEACH-SF, unlike other clustering techniques, the Fuzzy C-Means (FCM) algorithm is employed first in these hybrid methods to cluster all sensor nodes. CHs are then selected from cluster members using the fuzzy table (optimized through evolutionary and application-based algorithms). In SIF, the Mamdani Fuzzy Inference System (FIS) is used along with the firefly and simulated annealing algorithms for optimization. In LEACH-SF, the Sugeno FIS and Artificial Bee Colony (ABC) are employed for application-based optimization. Finally, non-CH nodes of every cluster become CHs selected from cluster members based on the highest fuzzy output in both methods. The inputs include residual energy, distance from the BS, and distance from the cluster gravity center. Since the fuzzy clustering method FCM is used in these two methods, computational complexity increased initially in comparison with other methods with no clustering. However, the distribution of CHs is more favorable in these two methods than in other protocols due to the use of FCM. In fact, previous methods lacked a mechanism for the uniform distribution of CHs. However, the FCM cannot form appropriate clusters due to merely considering the intra-cluster distance without paying any attention to the distance from the BS [36]. These methods sending the information directly from CHs to the BS could not work efficiently in large-scale networks, because no emphasis was given to the mechanism of routing information sent to the BS and the problem of hot spots. In other words, the information was sent to the BS through single-hop, and the only mechanism for controlling the activity radius of clusters was the FCM algorithm.

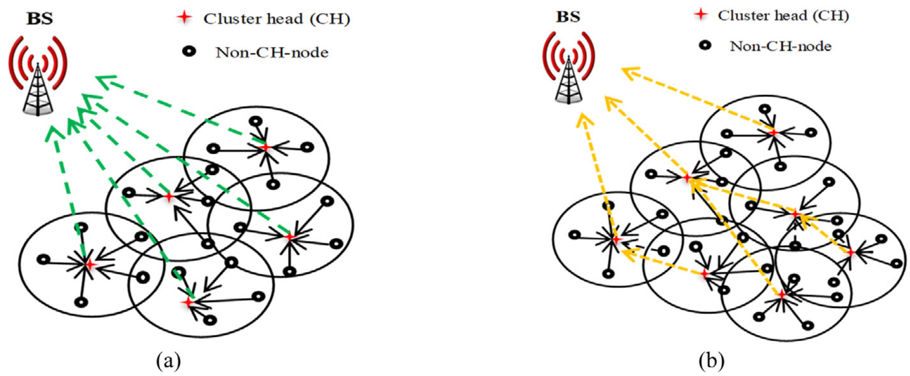


Fig. 1. (a) Sending information in single-hop clustering-based routing methods, b) Sending information in multi-hop clustering-based routing methods.

Due to the significant role of CHs, other parameters should also be taken into account in the selection of the nodes assuming the role of CHs. In FSFLA, combination of fuzzy and Shuffled Frog Leaping Algorithm (SFLA) is used to optimize the fuzzy rule table along with five other parameters. Inputs of the fuzzy inference system, energy, distance to BS, number of neighbor nodes, and node histories are considered. In this method, two thresholds including intra-cluster and inter-cluster, are used for proper distribution of CHs, which can be adjusted according to the need and required number of clusters in the network based on the application. This method uses the single hop mechanism to transfer information from the CHs to the BS.

2.2. Multi-hop clustering-based routing methods

Since wireless sensor nodes access limited sources of energy, the single-hop clustering method can reduce the communication energy of wireless networks. However, this method lacks energy-efficiency when communications are established on longer routes and single-hop communications consume more energy. Thus, multi-hop communications are appropriate for large-scale networks [37,38]. Multi-hop routing is usually associated with two strategies. In some methods, multi-hop routing is implemented among CHs. In other words, the routes are formed among CHs, and the information is sent to the BS through multi-hop routes. A number of such methods are Decentralized energy efficient Hierarchical Cluster-based Routing algorithm (DHCR) [39], Multi-Level Route-aware Clustering algorithm (MLRC) [40], Hierarchical Distributed Management Clustering (HDMC) [41], Flow-Balanced Routing protocol (FBR) [42], Hybrid Hierarchical Clustering Approach (HHCA) [43] and Energy-aware Routing Algorithm (ERA) [44]. However, some methods merely focus on routing among all sensor nodes to send information to the BS [45,46]. In methods with multi-hop routing among CHs, there will definitely be fewer problems in data distribution rate and wave interference or collision due to data aggregation in CHs and the administrative structure of CH roles in the network. As a result, the route length and transfer delay are shortened in reaching the BS. The number of nodes, the number of levels, delay, and other problems are decreased in methods that develop tree structures among nodes. In addition, structure development depends on a predetermined time in a number of such methods. In other words, the multi-hop structure may not develop among all network nodes [47–49].

Due to a large number of effective parameters and the unspecific effect of every parameter on the roles of CHs in wireless sensor networks, numerous methods have used the fuzzy logic [50] and fuzzy-based techniques in recent years to make decisions on selecting the best CHs or determining their activity radiuses to distribute them. Fuzzy based Unequal Clustering Protocol (FUCP) [51], Fuzzy logic Based Unequal Clustering (FBUC) [52], Multi-Objective Fuzzy Clustering Algorithm (MOFCA) [53], and Energy-Aware Unequal Clustering algorithm with Fuzzy (EAUCF) [54] focus on the problem of hot spots or CH selection. In MOFCA and EAUCF, temporary CHs are first selected randomly. The competition radius of every temporary CH is then determined by every node through FIS and certain inputs including the residual energy of every node, distance from the BS, and density. Therefore, the radiuses of temporary CHs are shortened in proximity to the BS. In Distributed Unequal Clustering using Fuzzy logic (DUCF) [55], every node determines its size and chance to become a temporary CH by considering residual energy, node degree, and distance from the BS to generate asymmetric clusters and select proper CHs through FIS. The reviewed methods did not take into account the mechanism of sending information to the BS through middle CHs. Generally, all of the above methods have not paid specific attention to the selection of the right parent node based on the application, although they use the multi-hop mechanism to send information to the BS.

2.3. Our contributions against the existing methods

The proposed clustering protocol (FMSFLA) selected CHs from all of the candidate nodes having higher residual energy rates, shorter distances to the BS, and shorter distances to the member nodes than other candidate nodes. It also selects parents from candidate CHs with respect to higher residual energy rates, shorter real distances, and lower path loads. The goal of the residual energy parameter is to balance the energy consumed by all of the nodes, although the distance to the BS and the real distance reduced the energy consumed to send information to the BS. In addition, the mean path load made the energy consumption steady in the network. In the clustering methods used for WSNs with different performances (e.g. increasing the network lifetime and energy efficiency), there are no tunable parameters to control the network current processes. The fuzzy rules base tables are defined manually. In addition, no optimization processes are implemented for application specifics. In simpler words, they have fixed and nonadjustable performances. They may perform well in one application, but produce dissatisfactory results in others. In the proposed FMSFLA, the SFLA is employed to tune two fuzzy rules base table along with five control parameters of the proposed protocol strategies to present an adjustable performance based on application. The innovations of the proposed FMSFLA are then compared with those of other methods briefly.

1. The goal of the FMSFLA is to maximize the network lifetime and the number of received packets based on the features of every application, which are usually ignored in most methods.

2. In most methods, there are no extensive criteria for CH selection (clustering) or the selection of the right parent node in the multi-hop process. The proposed protocol (FMSFLA), considers effective parameters such as energy, distance from the BS, the number of neighboring nodes, real distance from the BS, mean route load, delay, overlap, and the problem of hot spots.
3. In FMSFLA, application-based FIS is used for CH selection based on the compromise of the effective parameters pertaining to the node status among the nodes exceeding the mean network energy. As a result, the network energy consumption becomes uniform.
4. Parent selection is performed classically in most methods; however, the FMSFLA benefits from the FIS and application features to achieve flexibility and better parameters control.
5. The fuzzy rules of the FMSFLA are concurrently optimized for both CH selection phase and parent selection phase using the SFLA, so that in most methods, fuzzy rules are defined manually and used for all applications.
6. Most methods guarantee no appropriate CH distributions. However, the FMSFLA guarantees the appropriate distribution for space and application by considering minimum and maximum overlap values for adjacent CHs, which is determined optimally through the optimization process.
7. In the objective function, the optimization process of FMSFLA considers the trade-off between delay and energy to optimize the table of fuzzy rules in the CH and parent selection phases. Thus, the objective function of the proposed protocol is adjustable on this basis.

3. FMSFLA clustering protocol

The FMSFLA is proposed to achieve three main goals, the first of which is to develop an energy-efficient clustering and routing protocol. Second, the proposed protocol is supposed to be adjustable and usable based on different applications of the wireless sensor network. Third, it should be scalable. Fig. 2 presents an overview of the proposed FMSFLA protocol phases.

3.1. FMSFLA network and energy model analysis

The following hypotheses are made in the proposed protocol (FMSFLA):

- All sensors have the same level of initial energy, and each sensor with sufficient energy can directly communicate with the BS.
- Sensor nodes send their information to CHs periodically, and CHs are responsible for collecting, combining, and sending the information to the BS either directly or through other CHs.
- All sensors and the BS are motionless, and there are not energy constraints on the BS.
- All sensors are of a single homogeneous type.
- The BS can be located everywhere inside or outside the network, and nodes are distributed randomly and uniformly in the network.
- All sensors are aware of their position and the BS position, and BS has enough information about the network and its nodes.
- The Time Division Multiple Access (TDMA) scheduling algorithm is used to establish communications between CHs and their member nodes.
- All sensor nodes are able to adjust the transfer strength based on the distance from the receiver nodes.

The LEACH protocol energy consumption model [5] is considered in the FMSFLA to facilitate comparison. Eq. (1) shows the model.

$$E_{Tx}\ (l,d)=\begin{cases} lE_{elec}+l\epsilon_{fs}\ d^2 & d<d_o \\ lE_{elec}+l\epsilon_{mp}\ d^4 & d\geq d_o \end{cases} \tag{1}$$

where E_{elec} (nJ/bit) represents the energy consumed in a transmitter or receiver electronic circuit to send or receive one bit of the packet data. Moreover, ϵ_{fs} (pJ/bit/m^2) and ϵ_{mp} (pJ/bit/m^4) show the energy consumed by the radio amplifier of a transmitter node to send one bit of the packet data in a d-long channel between a transmitter node and a receiver node. The transmitter node consumes energy both to set up the necessary radio-electronic circuit (to put data on the channel) and to set up the radio amplifier. However, the receiver node only consumes energy to set up the radio-electronic circuit (to read data on the channel). As a result, $E_{Tx}(l,d)$ of Eq. (1) is used to determine the necessary energy used for sending l bit(s) on d. The first equation of Eq. (1) is used for models of free space and less than the threshold of distance (d_o), and for distances greater than (d_o), the second equation of Eq. (1) is used, where $\mathrm{d_o}=\sqrt{\epsilon_{fs}/\epsilon_{mp}}$. The following equation is used for receiving l bit(s).

$$E_{Rx}\ (l)=lE_{elec} \tag{2}$$

3.2. FMSFLA protocol overview

The FMSFLA is a centralized protocol, all computations of which are performed in a processor in the BS. The proposed FMSFLA consists of a few rounds, each including CH selection, parent selection, cluster formation, and steady state. In each round, the BS collects some information on the node status including residual energy, distance from the BS, the number of neighboring nodes, etc.

Phase 1: CH Selection Phase: In this phase, positional information and status of the nodes in the BS are taken into account to select CHs. In each round, the positional information of every node including residual energy, distance from the BS, and the number of neighboring nodes in $R_{hot\ spots}$, is collected. This information is then regarded as fuzzy inputs in the FIS, which uses the table of fuzzy rules, employed for CH selection, once before setting up the network (with respect to the inputs) along with the table of rules for the parent selection phase and five control parameters based on the optimal application features. Fuzzy computations are performed on every node whose energy level is higher than the mean energy of alive nodes in accordance with the determined threshold (β) to select candidate CHs in the CH selection phase. Eq. (3) is used for this purpose:

$$\mathrm{PT}_{Candidate}{}^{\mathrm{t}}\ (\mathrm{n})=\begin{cases} Out\ crisp_{CH}^{t}\ (n) & if\ \ \mathrm{E^t\ (n)}\geq\beta\times\mathrm{E_{avg}^{Alive}\ (t)} \\ 0 & if\ \mathrm{E^t\ (n)}<\beta\times\mathrm{E_{avg}^{Alive}\ (t)} \end{cases} \tag{3}$$

$$\mathrm{E_{avg}^{Alive}\ (t)}=\frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}\mathrm{E^t\ (i)}\qquad \mathrm{i}\in\mathrm{A} \tag{4}$$

where $E^t\ (n)$ is the residual energy of the node n in round t and N is the number of alive nodes. Moreover, A is the set of alive nodes in the current round t, and β is the energy threshold, used for determining the nodes with appropriate energy levels, compared with the mean energy of alive nodes in round t in Eq. (4), that can be selected as a CH candidate, whose value is determined in the optimization phase based on application. According to Eq. (3), the defuzzified output of $Out\ crisp_{CH}^{t}\ (n)$ determines the capability of n to become a CH in round t. It is also determined for the nodes whose energy levels are β times larger than the mean energy of alive nodes in comparison with the CH selection inputs. The fuzzy

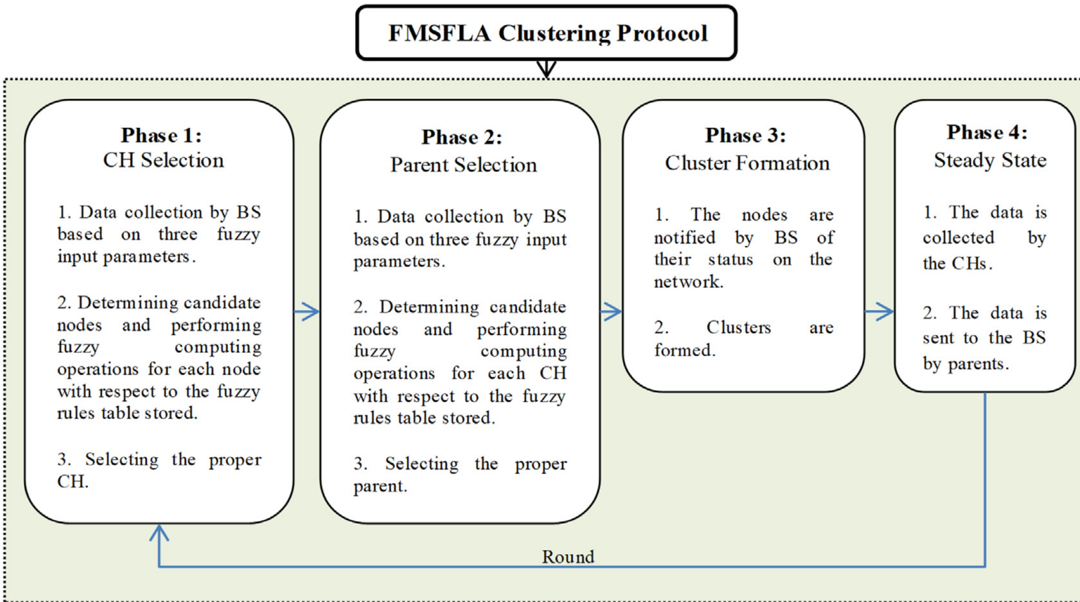


Fig. 2. An overview of the proposed FMSFLA protocol phases.

output values of the candidate node are sorted out in descending order, and the first candidate node with the greatest fuzzy output is selected as the final CH. It is then added to the final selected CHs list. Subsequently, other nodes of the sorted list are added to the selected CHs list if their overlap levels with adjacent CHs are lower than *Max_OV*. This process is continued until the following conditions are met and every candidate node of the selected CHs list is added to the final selected CHs list if it meets the following two conditions.

- The overlap of every CH like i on the selected CHs list with the closest adjacent CH like j on the final selected CHs list should at least be greater than *Min_OV*, shown in Eq. (5).

$$Min\left(OV_{i\,j}\right) \geq Min\_OV \quad \forall\, i \in \Delta \tag{5}$$

where Δ is the set of CHs which have been selected in the current round and put on the final selected CHs list.

- The distance between every member node m and the corresponding cluster head CH_m should be shorter than the $R^{CH_m}_{hot\ spots}$ of that CH, shown in Eq. (6).

$$Max\left(d\left(m, CH_m\right)\right) \leq \mathrm{R}^{CH_m}_{hot\ spots} \qquad \forall\, m \in \nabla \tag{6}$$

where ∇ is the set of alive nodes in the current round, and Eq. (7) is used for determining OV_{ij}, representing the overlap between two CH nodes i and j. The values of *Min_OV* and *Max_OV* represent the minimum and maximum overlaps of CHs based on the constraints imposed by the optimization algorithm on different applications. These values are optimized and stored before initiating the network operations. Then they are used at the time of CH selection.

$$OV_{i\,j} = \begin{cases} 0 & if\ R_{i\,j} > R^{i}_{hotspots} + R^{j}_{hotspots} \\ 1 - \dfrac{R_{i\,j}}{R^{i}_{hotspots} + R^{j}_{hotspots}} & if\ R_{i\,j} \leq R^{i}_{hotspots} + R^{j}_{hotspots} \end{cases} \tag{7}$$

where R_{ij} shows the Euclidean distance between two CH nodes i and j, shown in Fig. 3. $R^{i}_{hotspots}$ and $R^{j}_{hotspots}$ represent the radiuses of hot spots for adjacent nodes i and j, respectively. If OV_{ij} is equal to one, it shows the maximum overlap, meaning that the two CH nodes are almost placed on each other and their distance (R_{ij}) is null. If OV_{ij} is equal to zero, the two radiuses of CH nodes meet each other on their boundaries; therefore, the two CHs do not overlap. Fig. 4 shows the general process of selecting a CH in this phase briefly.

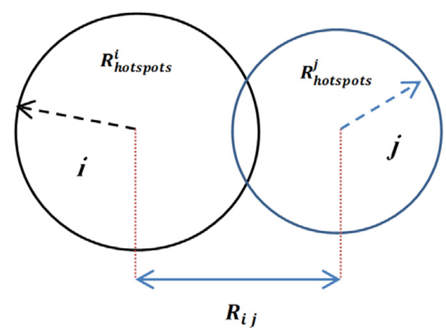


Fig. 3. The overlap between CH nodes i and j.

Phase 2: Parent Selection Phase: A virtual tree structure is developed between CHs and the root (i.e. BS) for routing data towards the BS. First, the BS transmits an *ADV-Construction Tree* message within d_0. The BS becomes the parent to the CHs receiving the message in this range. Then, every recipient CH located within a range shorter than d_0 with specific parents transmits the *ADV-Construction Tree* message within d_0 (such a CH is referred to as x) to every CH like y, which receives the message from CH nodes (x). Then, x is added to the candidate CHs list for the node y, as a CH node like y may receive the *ADV-Construction Tree* message from several nodes like x. Then, for nodes x on the candidate CHs list, a specific node like y that meeting Eq. (8) condition, is selected as the CHs of parent candidate nodes. They are subsequently put into the FIS, in which the input energy levels of candidate CHs are selected as the first fuzzy input parameters of the node x ($I1^{t}_{p}(x)$). The second fuzzy parameter of the node x is the real distance between candidate CHs and the BS ($I2^{t}_{p}(x)$). The third fuzzy parameter is the load mean of the route between the candidate CH (x) and the BS ($I3^{t}_{p}(x)$). Fuzzy computations are

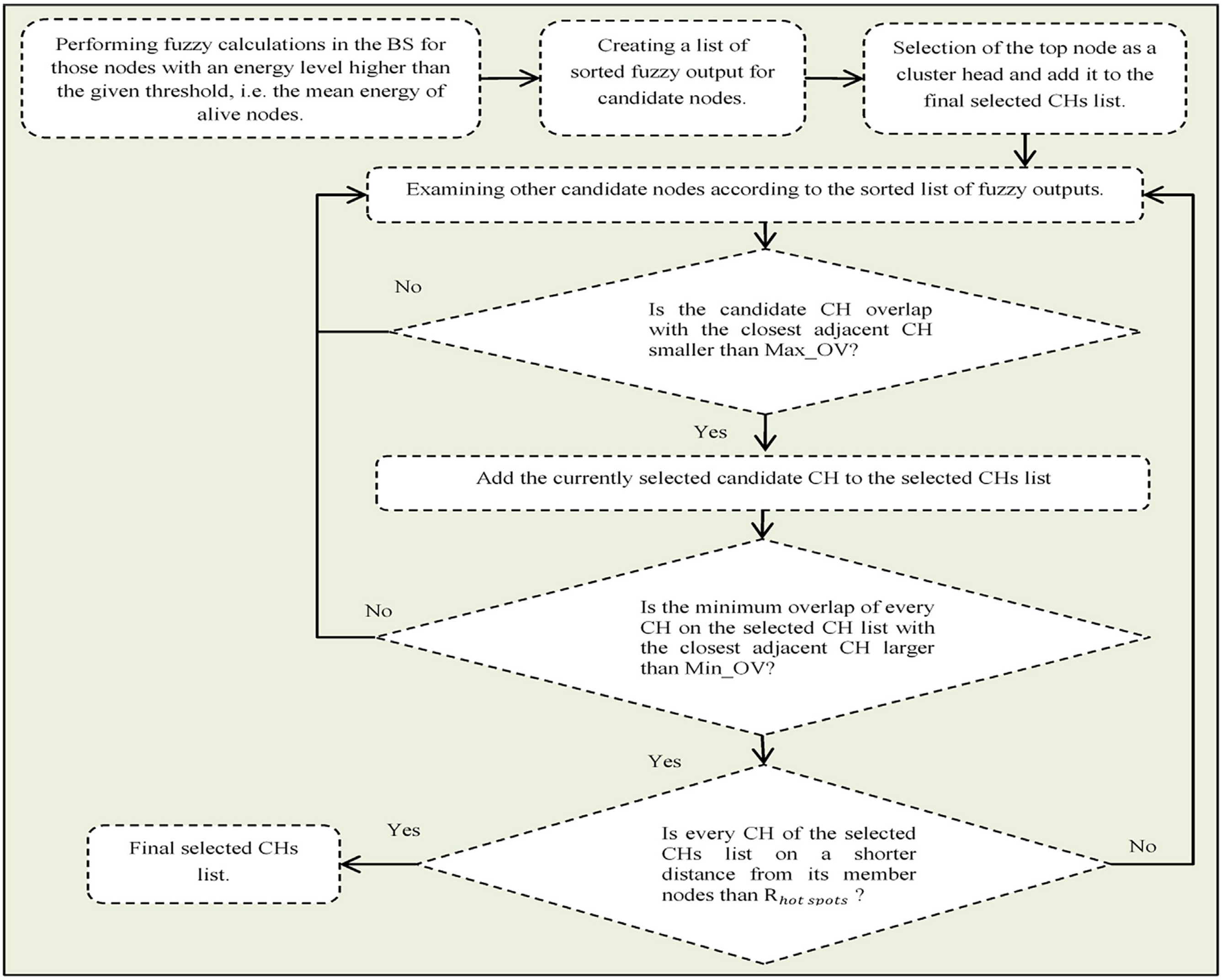


Fig. 4. The CH selection phase flowchart in the FMSFLA protocol.

performed for all candidate CH nodes (x), and the node with the highest fuzzy output is selected as a parent for y. This process is continued until parents are determined for all CHs in this range. These CHs send out the *ADV-Construction Tree* message within d_0 and repeat the process all over again to determine the parents of all CHs. Finally, the BS sends the parent ID to the nodes in the CH-selection phase to notify CHs at the same time as the ADV-BS message is being sent.

For instance, on Fig. 5, the BS becomes the parent of CHs nodes x_1, x_2, and x_3 within d_0 from the BS and also these nodes have received the *ADV-Construction Tree* message from the BS. Then, each of these CHs (x_1, x_2, and x_3) with a specific parent will simultaneously transmit an *ADV-Construction Tree* message within d_0. Assuming that y is a CH receiving the *ADV-Construction Tree* message from x_1, x_2, and x_3, the candidate CHs (x_1, x_2, and x_3) should first meet the condition in Eq. (8) to determine the parent for y. According to Fig. 5, Eq. (8) is analyzed for these three nodes. The condition was met for two nodes x_2 and x_3; however, $x_1$could not meet the condition because $a < c$ but $b < c$. Hence, x_2 and x_3 are added to the list of parent candidate nodes. Considering the fuzzy input parameters pertaining to the positional information on each of these two CHs, fuzzy computations are performed to obtain the defuzzified output of each CH. Eventually, the candidate CH with the greatest fuzzified output will be selected as the final parent of y. Moreover, the FIS uses the table of optimized fuzzy application-based rules, obtained before the network setup phase, for parent selection. Fig. 6 shows the parent selection pseudo code briefly.

$$\begin{aligned} & Dist\left(CH_y, CH_{x_i}\right) < Dist\left(CH_y, BS\right) \\ & \quad \& \; Dist\left(CH_{x_i}, BS\right) < Dist\left(CH_y, BS\right) \end{aligned} \tag{8}$$

Phase 3: Cluster Formation Phase: After determining CHs and parents, the BS sends an ADV-BS message to all nodes selected as CHs in the current round. Based on node IDs, every node is notified if it has been selected as a CH. In addition, every CH receives its relevant parent ID. Then the selected CHs send out ADV-CH messages including their IDs so that the non-CH nodes become the member of the nearest CH within $R_{hot\ spots}$ based on the signal strength of the CH messages.

Phase 4: Steady State Phase: In this phase, every non-CH node sends its information to the corresponding CH only in the time slot sent by the CH. Every node sends only one data packet to a CH in every round. Based on the aggregation strategy, CHs aggregate the received data and send them to the BS either directly if near the BS or through middle CHs if far from the BS. After this phase, the current round finishes, and the new round begins.

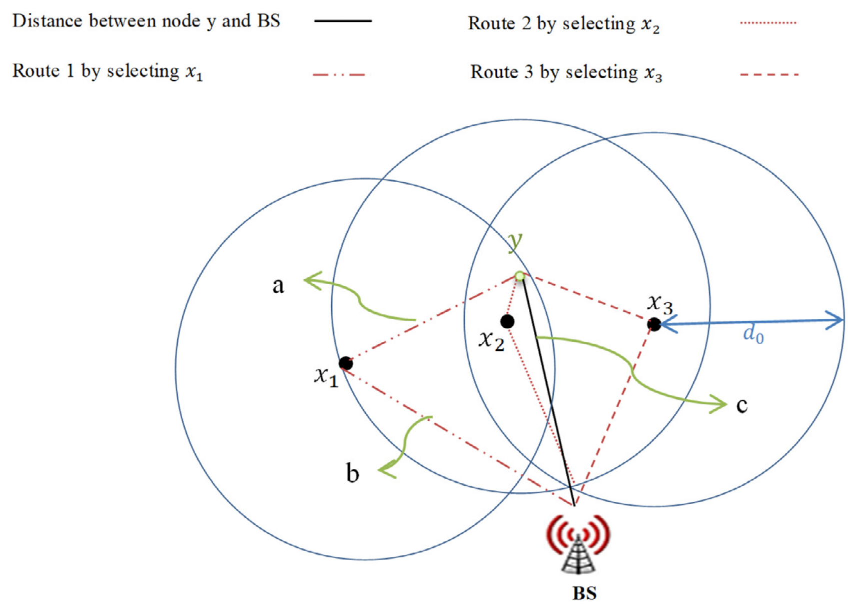


Fig. 5. The candidate CH selection by Eq. (8) in the parent selection phase of the FMSFLA protocol.

Algorithm: Parent Selection

1) **Input:** Number of CHs [Number of CHs]
2) **Output:** Parent Selection
3) Initially, BS broadcasts a ADV-Construct Tree msg in the range d_o
4) For each CH x that receives the ADV-Construct Tree msg.
5) $if\ (Dist(BS,x) \leq d_o\ )then$
6) Parent(x) =BS;
7) End if
8) End for
9) Each CH x, whose parent is specified, broadcasts the ADV-Construct Tree msg.
10) Each CH y (which is less than d_o than the nodes x) may receive multiple ADV-Construct Tree msg from several CH x.
11) If condition Eq (8) exists for each CH x and any CH y that received the msg from CH x.
12) That node x is added to the parent candidate nodes list for that y.
13) For each node x from the parent candidate nodes list
14) Calculate the energy of the node x ($I1_p^t(x)$).
15) Calculate the real distance of the node x (($I2_p^t(x)$).
16) Calculate the mean route load of the node x (($I3_p^t(x)$).
17) function[Out crisp $_p$] = $P\_Select(I1_p^t(x), I2_p^t(x), I3_p^t(x))$.
18) Calculate defuzzification output Out crisp $_p$.
19) List_parent [x] = Out crisp $_p$
20) End for
21) Final_parent= Max(List_parent);

Fig. 6. The pseudo code for the parent selection phase in the FMSFLA protocol.

4. Parameters and fuzzy logic model

In this section, the fuzzy logic model is presented thoroughly along with the input and output parameters of the CH selection and parent selection phases in the proposed protocol (FMSFLA), with details on how to calculate, and the type of membership functions for each, are fully described.

4.1. Fuzzy logic model

The fuzzy logic was introduced by Zadeh [50] in 1965. Expert knowledge can be used in the fuzzy logic to make decisions on how to control a system. The fuzzy logic is primarily employed in the FMSFLA for making decisions on the fuzzy outputs of CH selection and parent selection phases. Given the flexibility and compromise between factors affecting the appropriate CH selection, and parent selection through the FIS, the FMSFLA selects the right candidate for both phases based on application and the optimized fuzzy rule table. The FMSFLA consists of input information normalization, fuzzifier, fuzzy inference system (FIS), and defuzzifier.

- Normalization: This is a method of putting similar data from different domains in the same domain. Normalization prevents the significant effect of very large input data on the objective function. Eq. (9) expresses normalization; therefore, all crisp input variables range in [0, 1].

$$U' = \frac{U_{(n)} - min_A}{max_A - min_A} \times (n\_max_A - n\_min_A) + n\_min_A \quad (9)$$

  Assuming that max_A and min_A are the maximum and minimum of an input A, according to Eq. (9), the value of U of the input A is mapped onto another value U' ranging in $[n\_max_A - n\_min_A]$, equaling [0, 1] here. The value of U can include every input variable pertaining to the CH selection phase or the parent selection phase [56].
- Fuzzifier: Every crisp input value is mapped onto the corresponding fuzzy set by a fuzzifier. Then, an integer or a membership degree is attributed to every fuzzy set.

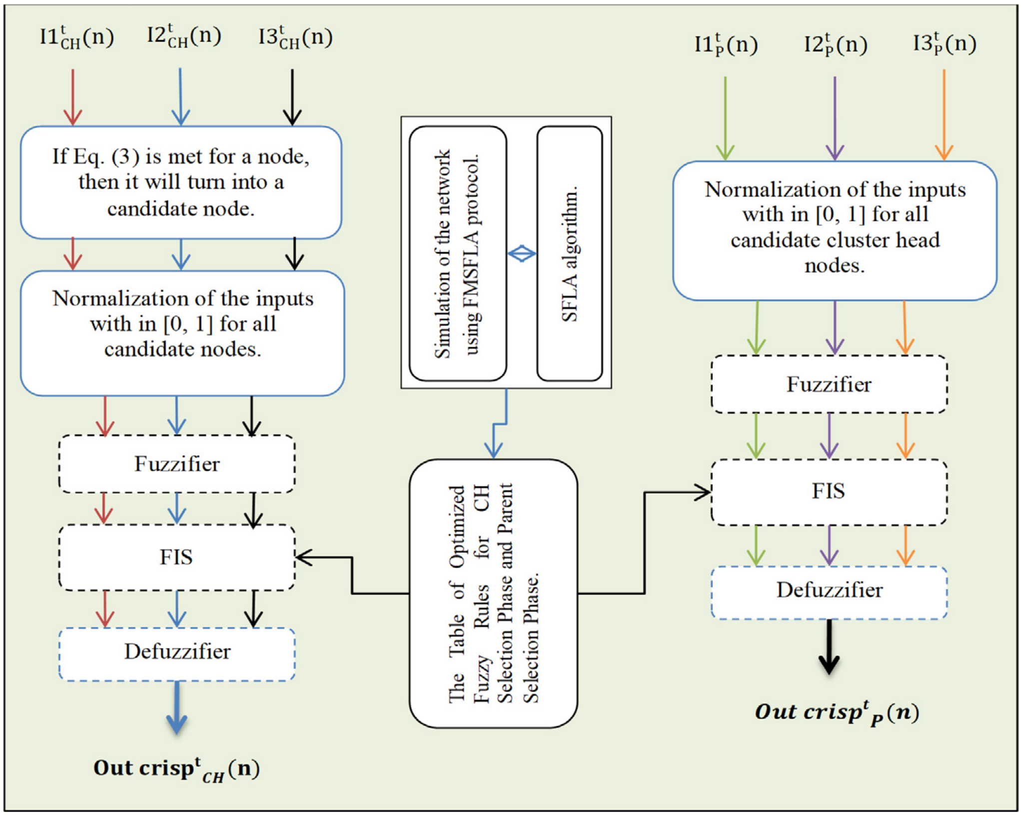


Fig. 7. The FMSFLA in CH selection and parent selection phases.

- Fuzzy Inference System (FIS): The FIS consists of a rule base table and different fuzzy inference methods pertaining to fuzzy values. The rule base table includes a series of rules using linguistic variables to describe and link the fuzzy input variables to output variables. Each of these variables is described by a fuzzy set or a fuzzy operator (AND or OR). In the FMSFLA, the fuzzy rule table is optimized and stored for both CH selection and parent selection phases based on applications and the SFLA. The FMSFLA table of rules is an AND-based Mamdani fuzzy system [57]. For instance, one rule of this table can be expressed as follows for each CH or parent selection phase:

$$\begin{aligned} &If \quad input_1 = a_1 \ \ and\ input_2 = a_2 \ \ and \ldots\ input_n = a_n \ \ , \\ &\ \ then\ output = t \end{aligned} \tag{10}$$

where the linguistic descriptor variables a_1 to a_n, pertaining to $input_1$ to $input_n$, indicate one version of the membership function for the input variables, and t is a linguistic variable of the membership function for outputs.
- Defuzzifier: Defuzzification is performed on the fuzzy solution space. In other words, a crisp output value is found in the fuzzy space. The FMSFLA employs the center of area (COA) method [58,59] to generate the final fuzzy output of the node n in the round t.

$$Out\ crisp^t_i\ (n) = \frac{\sum_{\mathrm{k}=1}^{Rules} \mu\ (\mathrm{k})\ .\mathrm{u}\ (\mu\ (\mathrm{k}))}{\sum_{\mathrm{k}=1}^{Rules} \mu\ (\mathrm{k})} \tag{11}$$

where i in $Out\ crisp^t_i\ (n)$ can be either $Out\ crisp^t_{CH}\ (n)$ or $Out\ crisp^t_P\ (n)$ and $\mu\ (k)$ is the minimum or multiple membership degree of every input in CH selection, and parent selection phases in the corresponding membership function in the rule k. Furthermore, $u(\mu\ (k))$ is the numerical value resulting from the center of the output membership function in the rule k. *Rules* show the number of rules, depending on the number of rules in every CH selection or parent selection phase. Fig. 7 shows the FIS of the FMSFLA protocol in both CH selection and parent selection phases.

4.2. Input parameters

In this section, for all input parameters, triangular membership functions (with different centers due to the behavior of the input parameters) are selected due to the common use of triangular membership functions in routing and clustering methods of wireless sensor networks.

4.2.1. Fuzzy input parameters of the CH selection phase

Input variables have a key role in a fuzzy system. For every node n in every round t in the CH selection phase, there are three fuzzy input variables including the node residual energy ($I1^t_{CH}(n)$), distance from the BS ($I2^t_{CH}(n)$), and the number of neighboring nodes ($I3^t_{CH}(n)$).

- $I1^t_{CH}(n)$: This variable indicates the residual energy of node n (the first input of node n) in round t of the CH selection phase. Due to the importance of energy consumption and more flexibility of fuzzy rules in this phase, five membership functions are considered for this variable. These functions are described by five linguistic variables including *Very Low*, *Low*, *Medium*, *High*, and *Very High* (Fig. 8).
- $I2^t_{CH}(n)$: This variable is the second fuzzy input in the CH selection phase. It indicates the Euclidean distance between node n and the BS. The shorter this distance, the higher the node's chance of becoming a CH. The FMSFLA protocol employs three triangular membership functions, shown by three linguistic variables *Low*, *Medium*, and *High* (Fig. 9).

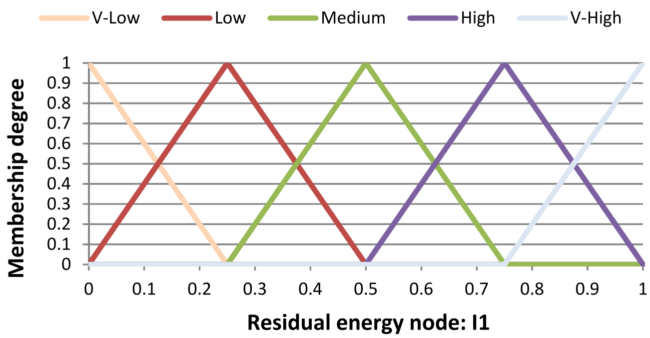


Fig. 8. The membership functions of the first fuzzy input in the CH selection phase.

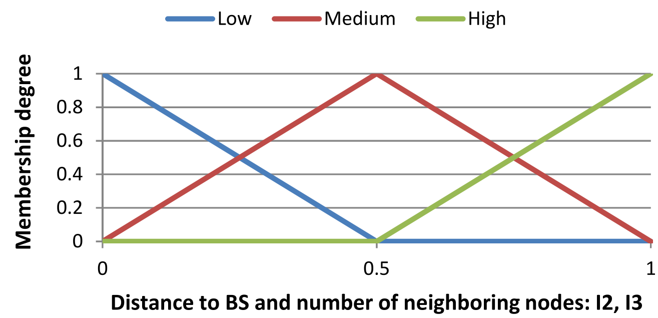


Fig. 9. The membership functions of the second and third fuzzy inputs in the CH selection phase.

- $I3^t_{CH}(n)$: This variable is the third fuzzy input in the CH selection phase. It indicates the number of neighboring nodes in round t. The radius of every node ($R_{hotspots}$) is taken into account to determine this variable. Since information is passed on from higher levels in multi-hop methods, i.e. levels far from the BS, to the CHs which are near the BS, the workload increases at CHs near the BS. Thus, they lose their energy sooner. This is the problem of hot spots. In the FMSFLA, Eq. (12) is used to determine the radius of every node based on its distance from the BS.

$$R_{hot\ spot}(n) = \left(\frac{d(n, BS) - d_{Min}(i, BS)}{d_{Max}(i, BS) - d_{Min}(i, BS)} \times (\mathrm{Max\_}d - \mathrm{Min\_}d)\right) + \mathrm{Min\_}d \quad (12)$$

where $d(n, BS)$ is the Euclidean distance between the node n and the BS. Moreover, $d_{Min}(i, BS)$ and $d_{Max}(i, BS)$ refer to the minimum and maximum Euclidean distances between the node i and the BS, respectively, when $i \in N$ (N is the number of network nodes). In other words, Eq. (12) determines the hot spot radius for every node. The determined ranges for the minimum radius of hot spots (Min_d) and the maximum radius of hot spots (Max_d) (stated as λ and γ which are coefficients of d_0) are optimized and stored by the SFLA based on applications. Fig. 10 shows a schematic view of the hot spots problem in the FMSFLA.
Given the third input, the greater number of neighboring nodes, the better the chance of a specific node to become a CH. Eq. (13) shows the number of neighboring nodes (n).

$$I3^t_{CH}(n) = \sum_{\substack{i=1 \\ i \neq n}}^{N} X(i, n) \quad (13)$$

$$X(i, n) = \begin{cases} 1 & if\ d(i, n) \leq R^n_{hotspots} \\ 0 & otherwise \end{cases} \quad (14)$$

where N refers to the total number of network nodes. $X(i, n)$shows the number of node n's neighbors for every neighboring node i. It is equal to one if $d(i, n)$ representing the Euclidean distance between node i and node n is shorter than $R_{hotspots}$ of node n in Eq. (14). Otherwise, it is null. Fig. 9 shows the triangular membership functions for this parameter. Thus, they are shown by verbal variables including *Low*, *Medium*, and *High*.
- Out crisp$^t_{CH}$(n): This is the fuzzy system output in the CH selection phase for node n in round t. It includes 9 triangular membership functions which are described as *Very Low*, *Low*, *Rather Low*, *Low Medium*, *Medium*, *High Medium*, *Rather High*, *High*, and *Very High*, respectively. Fig. 11 shows the functions. If a node has greater values of this parameter, it has greater chances of becoming a CH.

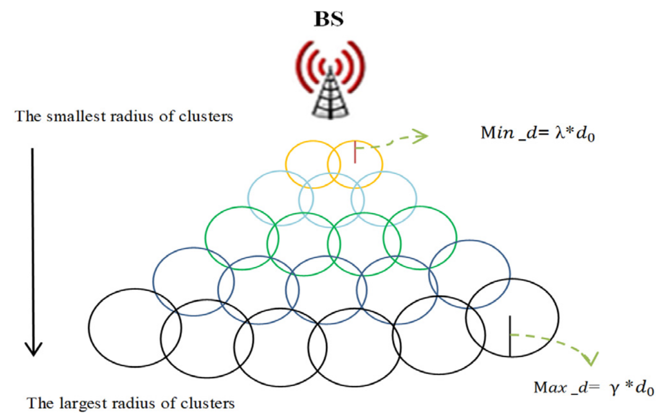


Fig. 10. The schematic view of the problem of hot spots in the FMSFLA protocol.

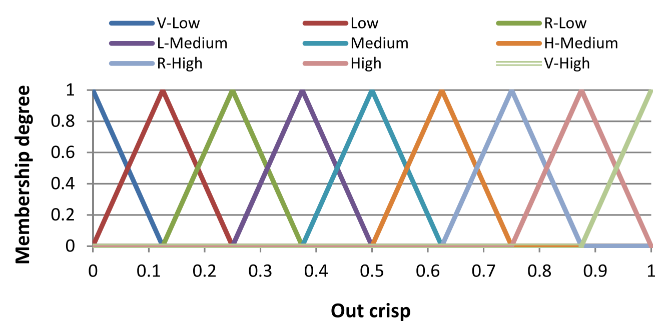


Fig. 11. Fuzzy output membership functions in CH selection and parent selection phases.

4.2.2. Fuzzy input parameters of the parent selection phase

- $I1^t_p(n)$: This is the first fuzzy system input of the parent selection phase, which indicates the necessary energy of the candidate CH node n to become a parent in round t. Given the importance of energy parameter in candidate CH selection for retransmitting and sending information towards the BS, five membership functions are considered for this parameter. These functions are described as *Very Low*, *Low*, *Medium*, *High*, and *Very High*. Fig. 12 shows the membership functions of this parameter.
- $I2^t_p(n)$: This is the second input of the parent selection phase, which shows the real distance of every CH node n in round t. Eq. (15) represents the real distance of the node. The real distance refers to the summation of the Euclidean distance

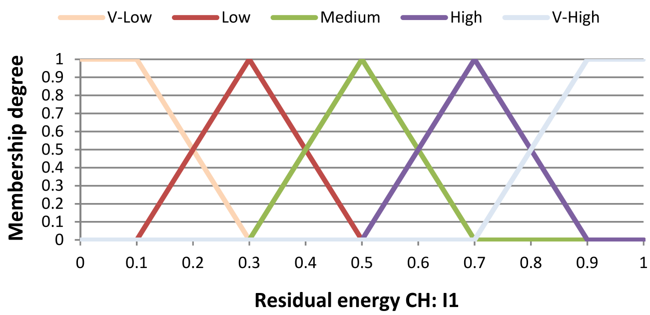


Fig. 12. The membership functions of the first fuzzy input in the parent selection phase.

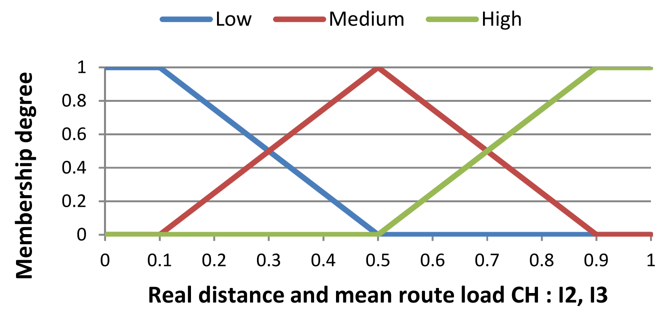


Fig. 13. The membership functions of the second and third fuzzy inputs in the parent selection phase.

between a child node (for which node n is a candidate for becoming its parent) and its parent, and between a parent and its parent, etc. all the way up to the BS.

$$I2_P^t(n) = \sum_{i=1}^{N_{hop}+1} d\left(child_i\ , parent_i\right) \tag{15}$$

where $d(child_i, parent_i)$ is the Euclidean distance between $child_i$ and $parent_i$. Moreover, N_{hop} shows the number of hops between the candidate CH node n and the BS. It is determined by counting the number of hops on the paths created by node n and its parent towards the BS. For instance, Fig. 5 shows that $N_{hop} = 1$ for node y between the candidate nodes x_2 and x_3 by selecting node x_2 for determining the real distance between y and x_2 and the distance between x_2 and the BS. The node x_3 is then selected to determine the distance between y and x_3 and the distance between x_3 and the BS. Three membership functions are considered for this parameter. The functions are described as *Low*, *Medium*, and *High*. Fig. 13 shows the membership functions.

- $I3_p^t(n)$: This is the third input of the parent selection phase, which shows the mean load of the selected route between the candidate CH node n and the BS in round t. Put simply, this parameter shows the mean number of parent cluster members plus the number of its parent members all the way up to the BS (the parents of node n are determined in advance; thus, the n-BS route includes the parents of node n). Initially, the operation starts from the node n, which is the candidate of the parent's role for the child node:

$$I3_P^t(n) = \frac{1}{N_{hop}} \sum_{i=1}^{N_{hop}} N_{parent_i} \tag{16}$$

where N_{parent_i} is the number of parent members of node i, and N_{hop} is the number of hops for the node n on the route created by the node n and the n candidate parents towards the BS. This parameter includes three membership functions, described as *Low*, *Medium*, and *High*. Fig. 13 shows the membership functions of this parameter:

- *Out crisp*$_p^t(n)$: This parameter indicates the fuzzy system output of the parent selection phase for the candidate CH node n in round t. For more flexibility, nine triangular membership functions are used and described as *Very Low*, *Low*, *Rather Low*, *Low Medium*, *Medium*, *High Medium*, *Rather High*, *High*, and *Very High* (Fig. 11). Every candidate CH node with the greatest *Out crisp*$_P^t(n)$ is selected as the final parent.

5. Optimization of the FMSFLA protocol via SFLA

This section presents the process of optimizing the fuzzy rule table of the FMSFLA through the SFLA optimization algorithm thoroughly. Fig. 14 shows the optimization flowchart of the proposed FMSFLA protocol.

5.1. Statement of the problem

In this paper, the shuffled frog leaping algorithm (SFLA) is employed to present a fuzzy multi-hop clustering protocol. This protocol analyzes the positional information of nodes to select optimal and appropriate CHs and parents to retransmit information towards the BS based on application features. To this end, the fuzzy rule table of the CH selection phase and that of the parent selection phase are optimized along with five adjustable parameters including β energy threshold in Eq. (3), maximum and minimum overlaps (*Max_OV* and *Min_OV*) in Eq. (5), and maximum and minimum radiuses of $R_{hotspots}$ (*Max_d* and *Min_d*) in Eq. (12) using the SFLA. The fuzzy rules of CH selection phase included 5, 3, and 3 membership functions for $I1_{CH}^t(n)$, referring to energy, $I2_{CH}^t(n)$ pertaining to the distance between the node and the BS, and $I3_{CH}^t(n)$, indicating the number of neighboring nodes, respectively. Therefore, there are 45 fuzzy rules in the CH selection phase ($Rule_{CH} = 45$). There are three inputs $I1_p^t(n)$ pertaining to the energy of every candidate CH node, $I2_p^t(n)$ pertaining to the real distance between the candidate CH node and the BS, and $I3_p^t(n)$ pertaining to the mean route load for the candidate CH node. Thus, there are 5, 3, and 3 membership functions, respectively. As a result, there are 45 fuzzy rules ($Rule_P = 45$). Generally, the entire search space included all possible solutions (9^{90}). Hence, this is an NP-hard problem, for which evolutionary algorithms are the best solution. In this paper, the SFLA is used for optimization. A possible solution to the SFLA is a meme string consisting of 95 memotypes (Fig. 15). The fuzzy rule table of the CH selection phase is optimized at the same time as that of the parent selection phase. There are 9 fuzzy output functions in the CH selection phase ($NFMout_{CH}$). There are also 9 fuzzy output functions in the parent selection phase ($NFMout_p$). Therefore, each memotype of the first 90 cells included a value selected from $NFMout_{CH} = NFMout_p \in \{1, 2, 3, \ldots, 9\}$. These values are described as *Very Low*, *Low*, *Rather Low*, *Low*, *Low Medium*, *Medium*, *High Medium*, *Rather High*, *High*, and *Very High*. Memotypes 90–95 related to five adjustable parameters including β in Eq. (3), *Min_OV* and *Max_OV* in Eq. (5), and *Min_d* and *Max_d* in Eq. (12). The variable ranges of these parameters are taken into account to determine their correct values in the optimization algorithm with a precision hop of 0.05.

$$0 \le \beta \le 1, \quad 0.2d_0 \le Min\_d \le 0.6d_0, \quad 0.7d_0 \le Max\_d \le 1.3d_0 \tag{17}$$

$$0 \le Min\_OV \le 0.5, \quad 0.5 < Max\_OV \le 1 \tag{18}$$

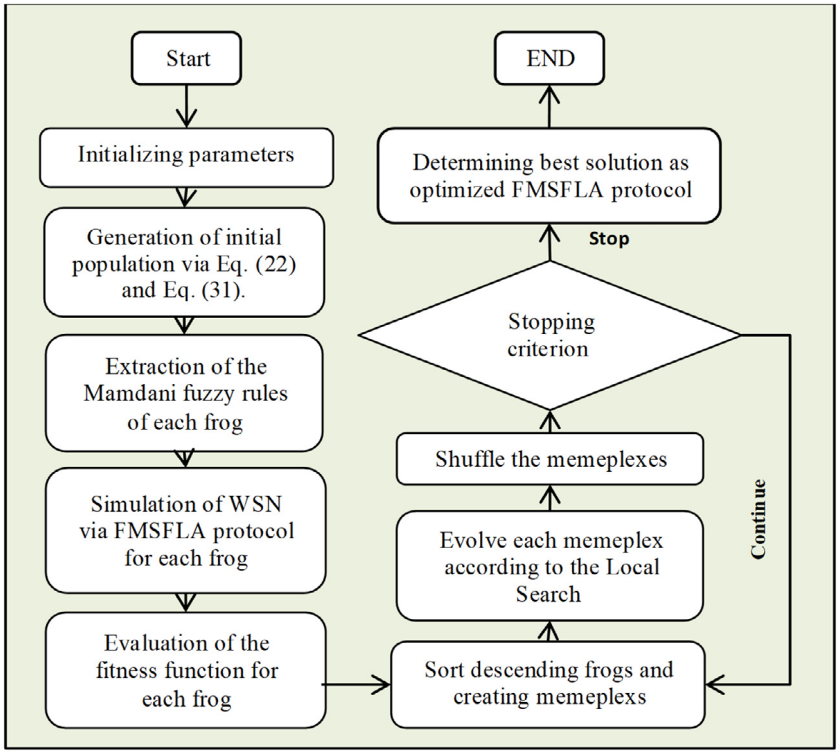


Fig. 14. Overall flowchart for the optimization of the FMSFLA protocol.

1	2	3	…	45	46	47	48	…	90	91	92	93	94	95
2	**1**	**3**	**…**	**8**	**9**	**4**	**3**	**…**	**6**	***β***	***Min_d***	***Max_d***	***Min_OV***	***Max_OV***

Fig. 15. A feasible solution to optimize the fuzzy rules in the FMSFLA protocol.

5.2. The proposed fitness function

Given the application-based strategy of the FMSFLA, the fitness function is a multi-objective function, adjusted to meet the application features.

Maximize:

$$FF = \left((w_1 \times FND + w_2 \times HND + w_3 \times LND + w_4 \times \frac{Data\ FND}{N} + w_5 \times \frac{Data\ HND}{N} + w_6 \times \frac{Data\ LND}{N}) / \left(Delay^{w\_delay}\right) \right)^{-1} \tag{19}$$

$$0 \leq w_i \leq 1\ , \quad \sum_{i=1}^{6} w_i = 1, \quad i = \{1, 2, 3, 4, 5, 6\} \tag{20}$$

Since the SFLA seeks to minimize the objective function, the inverted objective function is regarded as the fitness function (FF) in this algorithm. In Eq. (19), FND is the round in which the first node dies in the corresponding network. Moreover, HND is the round in which half of the nodes die in the corresponding network. Finally, LND is the round in which all of the network nodes die. These three parameters are given w_1, w_2, and w_3 weighted coefficients. $Data\ FND/N$, $Data\ HND/N$, and $Data\ LND/N$ are the average number of packets received by the BS from one node in the corresponding network until FND, HND, and LND, respectively. These parameters are given w_4, w_5, and w_6 weighted coefficients, respectively. It should be emphasized that the number of nodes (N) is used to balance the values of the six objectives within the multi-objective function of Eq. (19). The average delay of the number of hops is formulated in the following way for the CHs of all rounds.

$$Delay = \frac{1}{T * C} \sum_{t=1}^{T} \sum_{c=1}^{C} N_{hop\,tc} \tag{21}$$

where $N_{hop\,tc}$ shows the number of hops in the *cth* CH in round t, and w_delay is the weighted coefficient of delay. There is always a trade-off between energy and delay. In other words, energy consumption is high in methods where CHs send information direction to the BS in large environments, whereas the delay is low. However, delay always depends on the number of hops in information transfer in multi-hop methods [38,60,61]. The delay factor is also considered in the objective function of the FMSFLA, whose main goal is to reduce energy consumption and prolong the network lifetime. In applications where the delay is more important, the weighted coefficient of this parameter can be adjusted in the objective function to tune the table of fuzzy rules for such applications based on the importance of delay. In this objective function, weighted coefficients can be tuned with respect to the expected requirements. Accordingly, the fuzzy rule tables can be optimized along with five control parameters based on application features in both CH selection and parent selection phases.

5.3. FMSFLA protocol optimization procedure

As discussed earlier, an advantage of the FMSFLA is adjustability based on the requirements and features of different applications. Given the optimization of the fuzzy rules table in both CH selection and parent selection phases with along five adjustable parameters, this protocol can achieve the best performance and functionality. The FMSFLA employs the SFLA to perform the optimization process.

The shuffled frog leaping algorithm (SFLA) is a population-based cumulative search model. In this algorithm, the initial population is divided into certain subgroups named memeplex. Every memeplex M is characterized by a size q; thus, it is shown as $M \times q$. Every memeplex can be regarded as a group of frogs seeking the same goal. A frog's leap improves the meme presented by every individual and increases that individual's performance in reaching a goal in the local search. The frogs of every memeplex can be influenced by other frogs. Therefore, the memetic evolution takes place. After the memetic evolution occurs for certain times, the shuffle process is used for exchanging information between memeplexes. In the shuffle process, meme is influenced by certain frogs of different memeplexes and its quality improves. The local search and the shuffle-the-memeplexes process are iterated until a certain convergence criterion is met (the number of iterations is considered in this paper) [62,63].

5.3.1. Generating the initial population for FMSFLA protocol optimization process

CH Selection Phase: In this paper, the initial population is generated for the SFLA by considering the features of every input parameter in the fuzzy system. In other words, the initial population is generated in the CH selection phase by considering the node energy ($I1_{CH}^t(n)$), distance from the BS ($I2_{CH}^t(n)$), the number of neighboring nodes ($I3_{CH}^t(n)$), the direct effect of energy, the number of neighboring nodes on CH selection, and the inverse effect of distance from the BS.

$$W_{CH}\,(i,j,k) = (1 + Z1_{CH}) * (i-1) + (1 + Z2_{CH}) * (NFMi2_{CH} - j) + (1 + Z3_{CH}) * (k-1) \tag{22}$$

$$i = 1, 2, 3, \ldots, NFMi1_{CH} \tag{23}$$

$$j = 1, 2, 3, \ldots, NFMi2_{CH} \tag{24}$$

$$k = 1, 2, 3, \ldots, NFMi3_{CH} \tag{25}$$

$$W_{CH} = \left(round \left(W_{CH} * \frac{NFMout_{CH} - 1}{\max\,(W_{CH})} \right) \right) + 1 \tag{26}$$

$NFMi1_{CH}$, $NFMi2_{CH}$, and $NFMi3_{CH}$ indicate the number of membership functions in every input, respectively. They are equal to 5, 3, and 3. $Z1_{CH}$, $Z2_{CH}$, and $Z3_{CH}$ are uniform random numbers ranging in [−0.5, 0.5]. According to Eq. (26), the output of the initial population, generated by Eq. (22), is considered to range in the domain for the number of fuzzy output functions for the CH selection phase ($NFM\ out_{CH} = 9$).

Parent Selection Phase: The energy of candidate CH node n ($I1_p^t(n)$), real distance from the candidate CH node n ($I2_p^t(n)$), and the mean load of the route of the candidate CH node n ($I3_p^t(n)$) are taken into account in generating the initial population for the table of fuzzy rules in the parent selection phase. The following equations express the direct effect of energy and the inverse effects of other parameters:

$$W_P\,(i,j,k) = (1 + Z1_P) * (i-1) + (1 + Z2_P) * (NFMi2_P - j) + (1 + Z3_P) * (NFMi3_P - K) \tag{27}$$

$$i = 1, 2, 3, \ldots, NFMi\,1_P \tag{28}$$

$$j = 1, 2, 3, \ldots, NFMi\,2_P \tag{29}$$

$$k = 1, 2, 3, \ldots, NFMi\,3_P \tag{30}$$

$$W_P = \left(round \left(W_P * \frac{NFMout_P - 1}{\max\,(W_P)} \right) \right) + 1 \tag{31}$$

$NFMi1_P$, $NFMi2_P$, and $NFMi3_P$ refer to the number of membership functions for each input, considered 5, 3, and 3, respectively. Each $Z1_P$, $Z2_P$, and $Z3_P$ is a uniform random member ranging in [−0.5, 0.5]. According to Eq. (31), the output of the initial population, generated by Eq. (27), existed in the range of the number of output fuzzy functions for the parent selection phase ($NFM\ out_P = 9$). Hence, the initial population includes 95 memotypes based on the knowledge of certain parameter features and slight random changes for 90 cells representing initial solutions pertaining to two rule tables (Fig. 15). The initial solution to five other cells of solutions pertain to five adjustable parameters including β, *Min_d*, *Max_d*, *Min_OV*, and *Max_OV*. According to Section 5.1, these parameters are randomly generated in the predetermined range using the SFLA.

5.3.2. Fitness evaluation

In order to evaluate any solution in the SFLA algorithm, initially, according to inputs for each CH selection and parent selection phases, along with 5 controllable parameters, using the Eq. (22) to Eq. (31), the initial population is generated. Mamdani fuzzy rules are then extracted for each frog. Given the conditions of nodes in the relevant network and scenario, a few rules are fired. Then, their fuzzy outputs are calculated. The CH selection phase strategy (Section 3.2) is adopted to select CH nodes. According to Section 3.2, appropriate parents are selected to send information from CHs to the BS. Finally, the network is clustered, and the information is sent to the BS through the selected parents. Generally, the WSN corresponding to every solution is developed and simulated thoroughly. Then, Eq. (19) is used for evaluating the objective function for every network corresponding to a solution. Therefore, the best solution is selected. Subsequently, frogs are sorted in a descending order based on fitness. According to the SFLA, the initial population is divided into a number of memeplexes. Every memeplex is then evaluated by local search. Finally, memeplexes are shuffled and the best solution is saved. This process is continued until the termination condition is met when the best table of fuzzy rules is saved by considering the best fitness value.

6. Analysis of the complexity of the online process of FMSFLA

The time complexity of the FMSFLA protocol includes the time complexity of the CH selection phase (clustering) and the parent selection phase (multi-hop information transfer to the BS). The fuzzy rule tables, and the five adjustable parameters are optimized and saved for the current process of the network operation (online). Since the FMSFLA is a centralized protocol, all CH selection and parent selection processes are done in the BS. The time complexity of this protocol is discussed here.

The FMSFLA uses the fuzzy logic to select CHs from alive nodes through the BS. Given the execution of the FIS on every alive node N, the time complexity is equal to $O(N \times N_{k_{CH}})$, in which $N_{k_{CH}}$ shows the number of fuzzy rules for the CH selection phase. In the FMSFLA, the exchanged messages include the ADV-BS notification sent by the BS to CHs. This message notifies the nodes of being in the current CH round. This message also includes some information about the parent of every CH, which should be maximally equal to the number of CHs or the number of clusters. If the number of clusters is C, the time complexity is equal to $O(C)$. Another exchanged message is the ADV-CH notification, sent by CHs to their surroundings so that the nodes become the members of a CH in the shortest distance to them. This message is sent to N alive nodes with a complexity of $O(N)$. The JOIN-CH message is sent by member nodes to corresponding CHs to join them. The message length is $(N - C)$; thus, its complexity is $O(N)$. The ADV-TDMA message, sent by CHs to their member nodes to send the TDMA time table, is equal to the number of alive nodes represented by N with a complexity of $O(N)$. After determining the CHs and parents in the proposed protocol, every CH sends its information to its parent, i.e. a middle CH or the BS. The time complexity of the parent selection phase depends on the number of CHs. In much simpler words, it depends on the number of clusters (C) and the number of hops (N_{hop}). Generally, N_{hop} is dependent on the size of the network (L_{net}), which grows the as much as the network size, N_{hop} also increases. Thus, for the parent selection, depending on the number of fuzzy rules (N_{kp}) in the parent selection phase, the time complexity is $O(C \times L_{net} \times N_{kp})$. Moreover, the number of candidate nodes should also be considered in this phase for parent selection. However, it is ignored due to its indefinite nature. Thus, Eq. (32) is used for analyzing the global complexity of the proposed protocol.

$$O\left(\left(N \times N_{k_{CH}}\right) + \left(C \times N_{hop} \times N_{k_p}\right)\right) \tag{32}$$

6.1. The analysis of the complexity of the offline process of FMSFLA

The SFLA is employed in the FMSFLA to perform the optimization process offline before setting up the network. In the SFLA, the initial population consists of two parts (F=M×q). Thus, the complexity of the initial population generation is $O(pop_M \times pop_q)$. The SFLA is iterated for It_{SFLA}. All members of the population evaluate the objective function as many times as the number of iterations in each full simulation of the network. Therefore, the time complexity of the entire network simulation in the FMSFLA with CH selection and parent selection phases during the entire network lifetime is equal to one round of the FMSFLA in the maximum network lifetime, i.e. LND, which is equal to $(parent\ selection + CH\ selection) \times LND$, in brief. The time complexity of five adjustable parameters is negligible in the determined round due to their random selection in the SFLA. Therefore, the time complexity of the entire optimization process in the SFLA can be obtained through the following equation:

$$O\left(\left(pop_M \times pop_q\right) \times \mathrm{It}_{SFLA} \times \left(\left(N \times N_{k_{CH}}\right) + \left(C \times N_{hop} \times N_{k_p}\right)\right) \times LND\right) \tag{33}$$

Moreover, the fuzzy rule table of both CH selection and parent selection phases with five adjustable parameters are optimized only once in the preprocessing phase based on the application features before network operations. Such a time complexity does not affect the network. However, considering its advantages, which optimize the fuzzy rule-base and the five controllable parameters according to a specific application, and comparing them with the manual adjustment process, we may tolerate this time complexity, which in turn provides us with a longer network lifetime and reduce energy consumption.

7. Performance evaluation

This section draws a comparison between the scenarios and details of configuring FMSFLA and other similar protocols in simulations. The simulation results are then compared and a brief account of procedures is given for each protocol.

Table 1
Descriptions of scenarios in different application specifications.

Parameters	Scenario 1	Scenario 2	Scenario 3
Number of nodes	200	200	100
Sensor field	400m × 400m	400m × 400m	300m × 300m
BS location	(200, 200)	(200, 200)	(150, 150)
w_1	0.7	0.2	0.6
w_2	0.2	0.7	0.3
w_3	0	0	0
w_4	0.05	0.05	0.05
w_5	0.05	0.05	0.05
w_6	0	0	0
w_delay	0.01	0.01	0.01

7.1. Simulation settings

In the FMSFLA, different values are evaluated to select the best value in order to configure the parameters of the population-based SFLA. For the purpose of this study, the frog population size is 50 obtained from $f = M \times q$, where M is the number of memeplexes (=5), and q is the size of each memplex (=10). As the termination condition, the maximum number of iterations is 100 in this algorithm. Three scenarios are designed to set up the objective function of the FMSFLA to optimize Mamdani fuzzy rule table in both CH selection and parent selection phases along with five adjustable parameters based on the application. Given the application requirements, these scenarios are designed with different significance degrees for FND and HND parameters in the objective function. In the first and second scenarios, a 400×400 m^2 space with 200 nodes is analyzed when the BS is located in the center of the grid. In the third scenario, a 300×300 m^2 space with 100 nodes is analyzed when the BS is in the center. Regarding these three scenarios, the application-based aspect of the proposed protocol is evaluated optimally with respect to different adjustments of weighted coefficients in addition to changes in the search space. In Eq. (19), w_1 and w_6 are the weighted coefficients of the objective function used for configuring the fuzzy rule table and five adjustable parameters through the SFLA. In each scenario, the fuzzy rule table of both CH selection and parent selection phases are considered along with five adjustable parameters with respect to coefficients and applications. They are then optimized and saved in the relevant scenario. Table 1 shows brief specifications of the three scenarios along with the weighted coefficients used in Eq. (19). These weights can be adjusted with respect to application requirements to achieve the best performance. For instance, FND is the most important parameter in the first scenario. It is of great significance in highly-sensitive areas such as medical applications. Therefore, the significance degree of FND can be adjusted to its significance level in an intended application. In the second scenario, HND is more important than FND. Such configuration can be used in low-sensitive areas such as in agricultural applications. Table 2 shows a brief account of values considered for β in Eq. (3), Min_OV and Max_OV in Eq. (5), and Min_d and Max_d in Eq. (12) after optimizing five control parameters through the SFLA in the FMSFLA.

Tables 3 and 4 shows the fuzzy rules table optimized for both the CH and parent selection phases by the SFLA algorithm for the third scenario, and Fig. 16 shows the graph of surf view based on the positive and negative effects of the fuzzy input parameters on the output. It should be noted that in Fig. 16, since each phases 1 and 2 of the proposed FMSFLA protocol has three inputs and one output, the surf graph of the fuzzy rules table of each phase with respect to the three modes of membership functions of the third input parameter is shown.

Table 5 is a brief demonstration of the energy parameters. The FMSFLA protocol adopts a 10% aggregation rate strategy in

Table 2
Optimized values of control parameters by the SFLA in each scenario.

Controllable parameters	Descriptions	Range of variations	Equation number	Optimized value in Scenario 1	Optimized value in Scenario 2	Optimized value in Scenario 3
β	Energy threshold parameter	[0,1]	Eq. (3)	0.25	0	1
Min_OV	Minimum overlap of CHs	[0, 0. 5]	Eq. (5)	0.25	0.45	0.35
Max_OV	Maximum overlap of CHs	[0.5,1]	–	0.6	0.65	0.6
Min_d	Minimum $R_{hotspots}$ (clusters)	[$0.2d_0$,$0.6d_0$]	Eq. (12)	52.6	26.3	52.6
Max_d	Maximum $R_{hotspots}$(clusters)	[$0.7d_0$,$1.3d_0$]	Eq. (12)	114	105.2	100

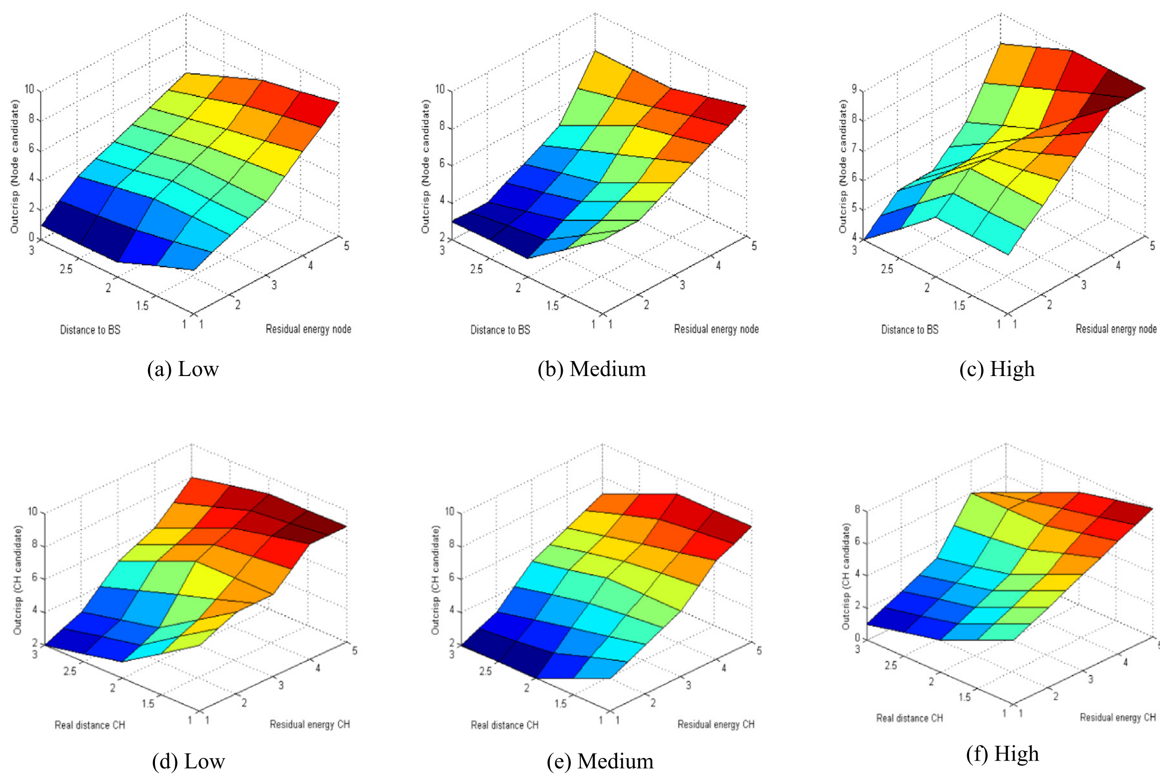


Fig. 16. The surf graph of the fuzzy rules tables of phases 1 and 2 of the FMSFLA protocol with respect to their third input parameter in the third scenario. (a), (b) and (c): Optimized fuzzy rules table graph of phase 1 (CH selection) with respect to three membership functions of parameter number of neighboring nodes. (d), (e) and (f): Optimized fuzzy rules table graph of phase 2 (parent selection) with respect to three membership functions of parameter mean route load CH.

CHs [53]:

$$\mathrm{S_{comp}} = \mathrm{S_{rec}} \times \mathrm{R_{agg}} \times \mathrm{T_{CLUSTER}} \quad (34)$$

$$\mathrm{S_{agg}} = \mathrm{S_{rec}} + \mathrm{S_{comp}} \quad (35)$$

In Eq. (34), S_{comp} shows the amount of compressed data, and S_{rec} refers to the packet data size received from every member node. $\mathrm{R_{agg}}$ refers to the data aggregation rate, and $\mathrm{T_{CLUSTER}}$ refers to the number of member nodes in the cluster except for CH. Furthermore, S_{agg} shows the size of aggregated data obtained from Eq. (35).

The proposed protocol is simulated on three scenarios, in which the nodes are randomly scattered 10 times. The network is simulated each time. Put simply, all compared methods are evaluated against the proposed protocol in 30 scenarios. Table 5 shows the mean performance of each scenario after 10 random node states.

7.2. Used protocols in evaluation

In this section, every protocol is briefly described along with evaluation conditions for the sake of comparison.

- LEACH [5]: In this protocol, every node selects a random number. If it is smaller than the threshold, the node becomes a CH and other nodes join the closest CH. The information is then sent directly to the BS through CHs.
- LEACH-EP [23]: In this method, the nodes with higher energy levels, compared to the mean CHs energy in the previous round, have higher chances of becoming a CH. In other phases, this protocol acts as the LEACH protocol.
- LEACH-FL [27]: In this protocol, every node uses the fuzzy logic to determine its chance and then subtracts it from one. If the result is smaller than the threshold, then the node becomes a CH. In other phases, this protocol acts as LEACH.
- ASLPR [32]: This protocol collects different pieces of information such as distance to the BS, residual energy, and the CH-to-CH distances. Each node then selects a random number. Nodes with smaller thresholds than the determined threshold will become CHs. In this method, the genetic and simulated annealing algorithms are employed to optimize the parameters of the main threshold based on application features.
- SIF [33]: This protocol employs the FCM algorithm for clustering sensor nodes. Then, it selects the appropriate CHs from the cluster members using the fuzzy system and considering certain inputs including residual energy, distance from the BS, distance from the cluster centroid of gravity, and the highest fuzzy output. In this method, the fuzzy rule

Table 3
Optimized fuzzy rule table for phase 1 (CH selection) of the FMSFLA protocol in the third scenario.

Residual energy node	Distance to BS	Number of neighboring nodes	Out crisp$^t_{CH}$(n)
V-Low	Low	Low	R-Low
V-Low	Low	Medium	H-Medium
V-Low	Low	High	H-Medium
V-Low	Medium	Low	V-Low
V-Low	Medium	Medium	R-Low
V-Low	Medium	High	H-Medium
V-Low	High	Low	V-Low
V-Low	High	Medium	R-Low
V-Low	High	High	L-Medium
Low	Low	Low	L-Medium
Low	Low	Medium	H-Medium
Low	Low	High	R-High
Low	Medium	Low	L-Medium
Low	Medium	Medium	L-Medium
Low	Medium	High	R-High
Low	High	Low	R-Low
Low	High	Medium	R-Low
Low	High	High	Medium
Medium	Low	Low	Medium
Medium	Low	Medium	R-High
Medium	Low	High	High
Medium	Medium	Low	Medium
Medium	Medium	Medium	Medium
Medium	Medium	High	R-High
Medium	High	Low	L-Medium
Medium	High	Medium	L-Medium
Medium	High	High	Medium
High	Low	Low	R-High
High	Low	Medium	High
High	Low	High	V-High
High	Medium	Low	H-Medium
High	Medium	Medium	R-High
High	Medium	High	R-High
High	High	Low	Medium
High	High	Medium	Medium
High	High	High	H-Medium
V-High	Low	Low	V-High
V-High	Low	Medium	V-High
V-High	Low	High	V-High
V-High	Medium	Low	High
V-High	Medium	Medium	High
V-High	Medium	High	V-High
V-High	High	Low	H-Medium
V-High	High	Medium	High
V-High	High	High	High

Table 4
Optimized fuzzy rule table for phase 2 (parent selection) of the FMSFLA protocol in the third scenario.

Residual energy CH	Real distance CH	Mean route load CH	*Outcrisp*^t_P (n)
V-Low	Low	Low	H-Medium
V-Low	Low	Medium	L-Medium
V-Low	Low	High	L-Medium
V-Low	Medium	Low	R-Low
V-Low	Medium	Medium	Low
V-Low	Medium	High	Low
V-Low	High	Low	Low
V-Low	High	Medium	Low
V-Low	High	High	V-Low
Low	Low	Low	R-High
Low	Low	Medium	Medium
Low	Low	High	Medium
Low	Medium	Low	L-Medium
Low	Medium	Medium	L-Medium
Low	Medium	High	R-Low
Low	High	Low	R-Low
Low	High	Medium	R-Low
Low	High	High	Low
Medium	Low	Low	R-High
Medium	Low	Medium	H-Medium
Medium	Low	High	H-Medium
Medium	Medium	Low	R-High
Medium	Medium	Medium	H-Medium
Medium	Medium	High	L-Medium
Medium	High	Low	Medium
Medium	High	Medium	Medium
Medium	High	High	R-Low
High	Low	Low	V-High
High	Low	Medium	High
High	Low	High	R-High
High	Medium	Low	High
High	Medium	Medium	R-High
High	Medium	High	H-Medium
High	High	Low	H-Medium
High	High	Medium	H-Medium
High	High	High	H-Medium
V-High	Low	Low	V-High
V-High	Low	Medium	V-High
V-High	Low	High	High
High	Medium	Low	V-High
V-High	Medium	Medium	V-High
V-High	Medium	High	R-High
V-High	High	Low	High
V-High	High	Medium	R-High
V-High	High	High	Medium

Table 5
Parameters and values associated with the energy consumption model.

Parameters	Values
Initial energy of nodes	1J
E_{elec}	50 nJ/bit
ε_{fs}	100 pJ/bit/m^2
ε_{mp}	0.013 pJ/bit/m^4
Control packet size	100 bit
Data packet size	4000 bit
Aggregation ratio	10%

table is based on application features. The network is simulated by combining the firefly and the simulated annealing algorithms before starting operations.

- ERA [44]: This protocol consists of two phases; clustering and routing. Every node selects itself as a CH based on an energy timer. The nodes join the closest CH with the highest level of energy. In the routing phase, the BS is considered to be at the zero level, and the CH nodes receiving a message from the BS within a specific range will become the first-level CH nodes. This process is iterated by the CHs of this level and those of the next levels until the end of the network so that the network levels are determined. Then, every CH of a higher level sends a part of its data to the BS among the CHs of lower levels, which are closer to the BS with energy levels, which are greater than or equal to the mean residual energy of all CHs, from which a message has been received.

The LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols are evaluated against the proposed protocol (FMSFLA) in completely identical conditions. The FMSFLA is a multi-hop protocol. The A-STAR algorithm [64] is added to all single-hop methods including LEACH, LEACH-EP, LEACH-FL, ASLPR, and SIF, except for the ERA algorithm which is a multi-hop protocol, to improve and multi-hop them. In much simpler words, CHs always send information directly to the BS after clustering. In the methods discussed in this paper, CHs select the next CH to send information to the BS in a multi-hop way through the A-STAR algorithm. The A-STAR search algorithm is widely used in graph search algorithms. It is regarded as an effective heuristic algorithm for finding the shortest route. Since there must be several routes towards the BS in the above cluster-based protocols, the A-STAR algorithm is used by every CH node to find the next parent on

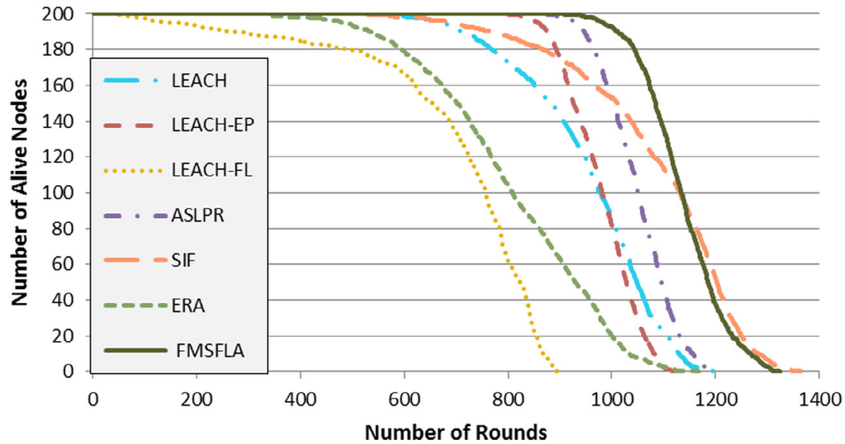


Fig. 17. Distribution of alive nodes with respect to different rounds for the first scenario.

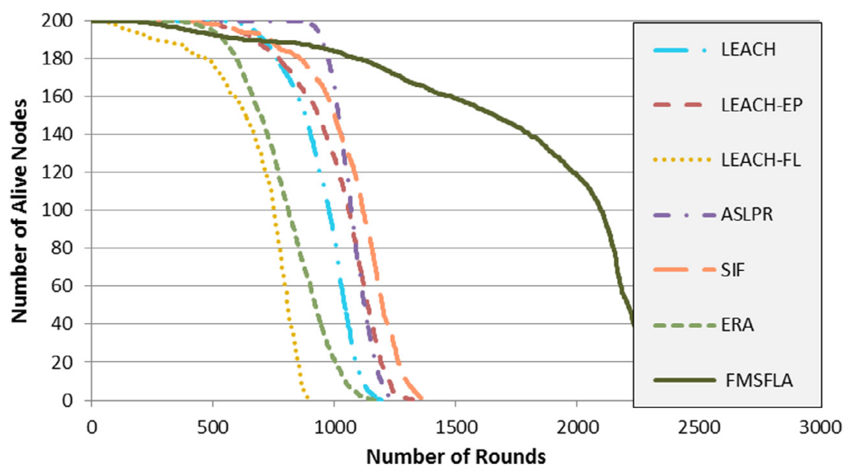


Fig. 18. Distribution of alive nodes with respect to different rounds for the second scenario.

the way up to the BS. The evaluation function of this algorithm is described as the following:

$$f(n) = g(n) + h(n) \quad (36)$$

where $g(n)$ indicates the route cost or the Euclidean distance between the source CH node and the destination node n, and $h(n)$ shows the route cost or the Euclidean distance between the node n and the BS. Therefore, every CH selects another CH node with a smaller $f(n)$ through Eq. (36). Put simply, the target CH should be on a simpler and straighter route towards the BS [65].

In addition, the ASLPR and SIF protocols are application-based methods, the objective functions of which can be adjusted to meet certain applications. Thus, the objective functions of these methods are the same as the objective function of the proposed protocol (FMSFLA). They have actually been optimized by adding the data aggregation rate through multi-hop routing. In the ERA protocol, it is assumed that $R = d_0/2$. Accordingly, all evaluated methods are multi-hop protocols, and the application-based methods are configured to meet the conditions of each scenario with respect to the proposed protocol. As a result, the FMSFLA is evaluated under completely reasonable conditions compared with the other methods. All simulations are run in MATLAB R2014 as it provides an integrated environment for clustering algorithms. Therefore, it facilitates simulation, execution, and comparison.

7.3. Comparison of the simulation results

Figs. 17–19 show the number of alive nodes in different rounds in all scenarios. Accordingly, the FMSFLA showed better functions in all scenarios. For instance, Fig. 17 indicates the number of alive nodes in different rounds in the first scenario with respect to its goal, showing the importance of FND and relevant increases in FND. The FMSFLA is evidently more stable than other protocols, as the nodes started to die later than usual. Node deaths continued at an almost linear rate until all nodes died.

Fig. 18 shows the number of alive nodes in different rounds of the second scenario with respect to its goal, showing the importance of HND and increases in HND. The first node evidently died sooner in the FMSFLA in the second scenario; however, this protocol could intelligently manage energy and node deaths with respect to the scenario goal. Therefore, HND increased. After half of the nodes died, the death of nodes lasted evenly until LND is reached. Fig. 19 indicates the number of alive nodes in different rounds of the third scenario. Accordingly, this scenario is designed to increase FND in a space with a small number of nodes. The FMSFLA is more stable than the other protocols.

Given the importance of FND in the first scenario and the weighted coefficients in the relative objective function, Table 6 and Fig. 20 indicate that the proposed protocol resulted in a better FND value in comparison with LEACH (61%), LEACH-EP (13%), LEACH-FL (270%), ASLPR (7%), SIF (56%), and ERA (155%).

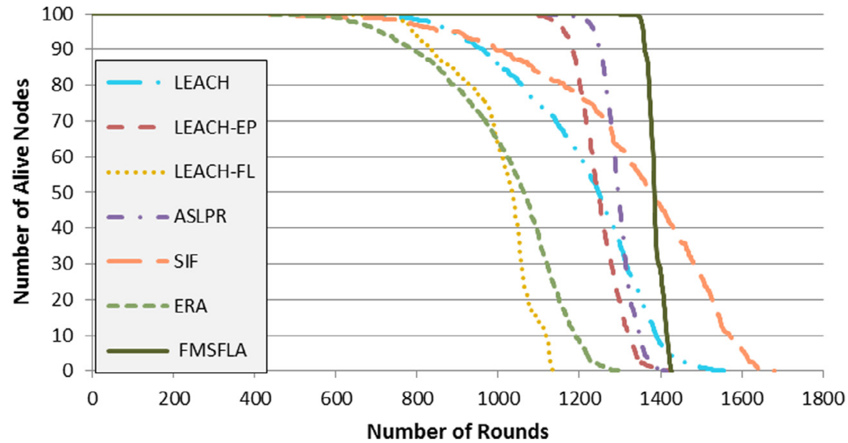
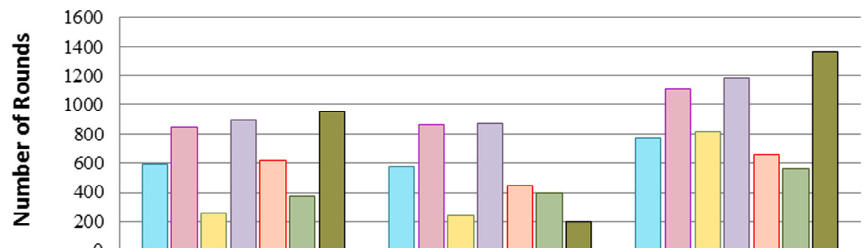


Fig. 19. Distribution of alive nodes with respect to different rounds for the third scenario.



	Scenario1	Scenario2	Scenario3
LEACH	593.8	577.8	776.4
LEACH-EP	849.7	862.8	1106.8
LEACH-FL	259.2	240.6	811.7
ASLPR	897.5	874.4	1187.7
SIF	615.2	441.6	657.2
ERA	375.2	395.7	559.7
FMSFLA	957.9	196	1368.6

Fig. 20. FND values for all protocols on different scenarios.

Table 6
Network lifetime in FND and HND for the first scenario.

Scenario 1	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND
1	651	1074	960	1063	663	842	998	1131	697	1201	540	927	1030	1212
2	520	1010	783	1015	102	764	905	1089	573	1155	411	796	973	1163
3	670	1035	971	1033	368	841	981	1103	705	1153	427	903	1044	1199
4	650	994	857	978	83	744	936	1051	561	1149	480	827	937	1143
5	580	959	847	952	495	713	877	1036	684	1061	386	776	949	1102
6	545	955	777	941	313	721	880	1017	676	1139	345	784	946	1102
7	589	981	860	976	33	668	847	1055	579	1125	279	743	966	1141
8	579	960	825	939	230	728	851	1010	509	1066	335	739	878	1094
9	592	973	837	983	112	806	810	1048	536	1149	343	814	964	1133
10	562	920	780	907	193	693	890	985	632	1078	206	765	892	1077
AVE	**593.8**	**986.1**	**849.7**	**978.7**	**259.2**	**752**	**897.5**	**1052.5**	**615.2**	**1127.6**	**375.2**	**807.4**	**957.9**	**1136.6**

According to Table 1, the weighted coefficient of FND is considered 0.2 in the second scenario, and the energy threshold (β) is nulled by the optimization algorithm in Eq. (3) (β affects the FND). Therefore, the FMSFLA intelligently seeks to increase HND. In other words, the death of the first node, FND, is not of great significance in the FMSFLA. As a result, FND in this protocol is smaller than other methods, however, for two application-based methods, i.e. ASLPR and SIF, (discussed in Section 7.2) the FND of these methods are decreased in the second to the first scenario, 23.1 and 173.6 rounds for ASLPR and SIF, respectively, while in the proposed FMSFLA protocol, the 761.9 round has dropped. Thus, since these methods are application-based, they change less in this scenario than in the FMSFLA. Accordingly, the FMSFLA employs more appropriate factors than the other two methods for CH selection and parent selection mechanisms. Moreover, this protocol also benefits from an ideal objective function and adjusts the fuzzy rule-base table automatically through optimization. Table 7 and Fig. 20 show the application-based functions of the FMSFLA protocol in comparison with other protocols.

This third scenario benefits from a smaller space than the other two scenarios. At the same time, it employs half the number of nodes used by the other two scenarios. The FND is more

Table 7
Network lifetime in FND and HND for the second scenario.

Scenario 2	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND
1	637	1032	964	1020	366	837	978	1124	591	1193	442	896	152	2112
2	498	1011	856	1014	102	756	897	1095	494	1119	406	832	148	2162
3	637	1032	964	1020	366	837	978	1124	538	1173	442	896	247	2093
4	642	989	890	966	189	742	938	1077	568	1145	500	843	200	2167
5	587	972	819	950	568	772	885	1050	421	1116	427	788	244	2108
6	579	930	805	937	324	674	828	1042	437	1101	332	810	187	2064
7	647	978	842	973	33	667	846	1064	399	1113	364	739	221	2091
8	547	957	814	950	156	698	784	1035	438	1072	383	780	141	2138
9	489	988	874	987	111	769	725	1067	310	1113	385	806	184	2039
10	515	915	800	917	191	682	885	1013	220	1071	276	728	236	2096
AVE	**577.8**	**980.4**	**862.8**	**973.4**	**240.6**	**743.4**	**874.4**	**1069.1**	**441.6**	**1121.6**	**395.7**	**811.8**	**196**	**2107**

Table 8
Network lifetime in FND and HND for the third scenario.

Scenario 3	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND	FND	HND
1	813	1245	1100	1229	699	970	1232	1281	605	1362	665	1064	1378	1385
2	853	1281	1157	1294	394	907	1240	1335	830	1411	633	1081	1408	1420
3	685	1251	1102	1226	724	1005	1202	1277	718	1371	626	1044	1290	1384
4	685	1226	1090	1253	963	1031	1232	1308	790	1399	560	1049	1386	1403
5	847	1251	1169	1245	944	1040	1196	1303	464	1343	432	1088	1366	1374
6	746	1319	1120	1293	955	1088	1245	1329	427	1410	656	1171	1402	1412
7	687	1247	931	1252	1079	1119	1014	1300	865	1389	379	1070	1378	1386
8	790	1232	1119	1214	667	951	1163	1287	495	1354	617	1030	1362	1370
9	845	1230	1149	1237	1006	1051	1228	1285	776	1373	522	1065	1370	1382
10	813	1223	1131	1223	686	991	1125	1269	602	1300	507	1019	1346	1357
AVE	**776.4**	**1250.5**	**1106.8**	**1246.6**	**811.7**	**1015.3**	**1187.7**	**1297.4**	**657.2**	**1371.2**	**559.7**	**1068.1**	**1368.6**	**1387.3**

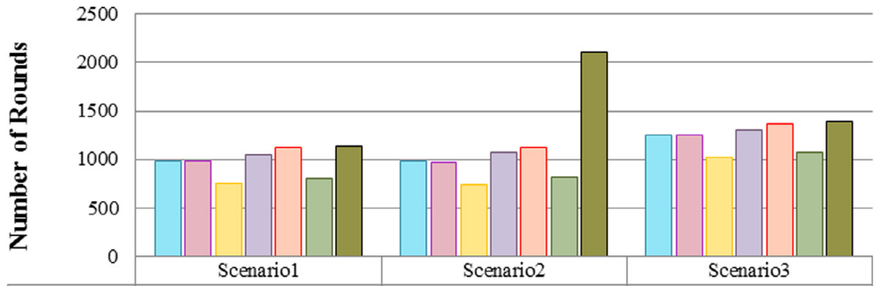
important in methods based on application features. Thus, all methods show higher FND values in this scenario than the other two. According to Table 8 and Fig. 20, the LEACH-FL protocol performed better in the third scenario, which shows that CH selection is not properly managed by the fuzzy rule table in the large scale networks despite the fact that the A-STAR algorithm is employed to send information through multi-hop routes. However, the FMSFLA showed a better FND than the other methods. It is more stable than LEACH (592.2), LEACH-EP (261.8), LEACH-FL (556.9), ASLPR (180.9), SIF (711.4), and ERA (808.9) algorithms, the round is more. In other words, these methods are 76%, 24%, 69%, 15%, 108%, and 145% more stable, respectively. Table 8 shows the performance of FMSFLA in the third scenario compared with the other three methods. Fig. 20 indicates the value of the most important parameter of the network lifetime, i.e. FND, in comparison with different rounds.

Tables 6 and 7 and Figs. 20 and 21 show the changes in the LEACH, LEACH-EP, LEACH-FL, and ERA protocols in the first and second scenarios. In a number of evaluated methods (LEACH, LEACH-EP, and LEACH-FL), the probability of nodes is used for CH selection. In the ERA protocol, a timer is employed for a part of the clustering phase. Given the correspondence of the first and second scenarios in space specifications, BS location, number of nodes, and node distribution in addition to their differences in adjusting weighted coefficients in the objective function Eq. (19), the conditions are constant in the first and second scenarios for the LEACH, LEACH-EP, LEACH-FL, and ERA methods.

Tables 6–8 also show the second parameter of the network lifetime, i.e. HND, for different methods of every scenario. Accordingly, Fig. 21 shows the mean HND for the 10 executions of every scenario in different rounds. As the figure obviously shows, the goal of the first scenario is to increase FND. In the FMSFLA, the HND parameter is greater than those of LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols by 150.5, 157.9, 384.6, 84.1, 9, and 329.2 in different rounds. The goal of the second scenario is increase HND, based on application features. According to Fig. 21, the FMSFLA increased HND intelligently. In fact, it increased the HND by 1126.6, 1133.6, 1363.6, 1037.9, 985.4, and round 1295.2 in the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols, respectively. In much simpler words, the FMSFLA is better than the other methods of this scenario by 115%, 116%, 183%, 97%, 88%, and 160%. Given the goal of the third scenario and the adjustment of FND to more important coefficients and HND to smaller weighted coefficients, the proposed protocol increased HND more than the other methods.

Tables 9–11 show the history of dead nodes clearly for all of the three scenarios (in 10% death of the nodes; after the death of 10% of the network nodes, the round will obviously show 10 random cases for nodes in each scenario as mean values). The greatest values are shown in bold. These values belong to the FMSFLA in the first and third scenarios with respect to their goals. After that, they belong to the SIF. As discussed earlier, the goal of the second scenario is to increase HND. The first node dies quickly; however, after the death of only 20 nodes (10%), the proposed protocol saves the highest number of alive nodes in the network, and assumes the highest values until LND is reached.

Figs. 22 and 23 show the number of packets received by the BS in each round before reaching FND and HND in the three scenarios. This parameter is important because different protocols may reach FND or HND in the same round. However, the nodes of every protocol may deliver different numbers of packets to the BS. According to Fig. 22, the proposed protocol delivered 72,820, 21,640, 139,740, 12,080, 68,540, and 116,540 more packets to the BS than the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA, respectively, in the first scenario. The goal of the second scenario is to increase HND. In the FMSFLA, the first node died early and intelligently. Therefore, a fewer number of packets are sent before FND. In the third scenario, the CHs of the FMSFLA



	Scenario1	Scenario2	Scenario3
LEACH	986.1	980.4	1250.5
LEACH-EP	978.7	973.4	1246.6
LEACH-FL	752	743.4	1015.3
ASLPR	1052.5	1069.1	1297.4
SIF	1127.6	1121.6	1371.2
ERA	807.4	811.8	1068.1
FMSFLA	1136.6	2107	1387.3

Fig. 21. HND values of all protocols on different rounds.

Table 9
History of dead nodes (on average) for the first scenario.

Scenario 1											
% Dead nodes	FND	%10	%20	%30	%40	HND	%60	%70	%80	%90	LND
LEACH	593.8	775.6	857.1	911.3	955.9	986.1	1012.7	1035.3	1059.3	1081.5	1144.2
LEACH-EP	849.7	920.4	936.3	950.5	965.2	978.7	994.5	1010.5	1022.8	1039.5	1065.5
LEACH-FL	259.2	494	615.2	664.2	709.2	752	779.9	797.9	813.1	827.1	839.3
ASLPR	897.5	985.9	1008.7	1022.9	1037.7	1052.5	1065	1078.3	1096	1114	1130.8
SIF	615.2	863.5	970.1	1048.3	1095.2	1127.6	**1165.5**	**1188.8**	**1214.4**	**1245.4**	**1288**
ERA	375.2	612.5	680.3	734.1	774.7	807.4	849.4	894.6	944.4	997.3	1099.1
FMSFLA	**957.9**	**1053.8**	**1088.1**	**1107.7**	**1123.1**	**1136.6**	1150.4	1167.9	1191.4	1216.6	1255.7

Table 10
History of dead nodes (on average) for the second scenario.

Scenario 2											
% Dead nodes	FND	%10	%20	%30	%40	HND	%60	%70	%80	%90	LND
LEACH	577.8	777.7	855.2	908.3	952.7	980.4	1006.7	1027.7	1055.8	1081.8	1131.4
LEACH-EP	862.8	913.7	932.1	946	960.7	973.4	989.5	1001.8	1018.4	1035	1065.5
LEACH-FL	240.6	486.2	599.2	647.7	699	743.4	771.1	787.9	806	822.1	835.6
ASLPR	**874.4**	987.8	1013.5	1035.6	1052	1069.1	1083.7	1108.8	1130.6	1166.9	1192.8
SIF	441.6	860.4	976.6	1038.1	1090.2	1121.6	1159.9	1187.5	1221.8	1256.6	1293.3
ERA	395.7	607.9	680.3	732.1	771.2	811.8	847.1	887	940.2	1000.8	1106.9
FMSFLA	196	**1116**	**1470.3**	**1805.2**	**1996.1**	**2107**	**2162.4**	**2204.8**	**2263.4**	**2297.1**	**2307.7**

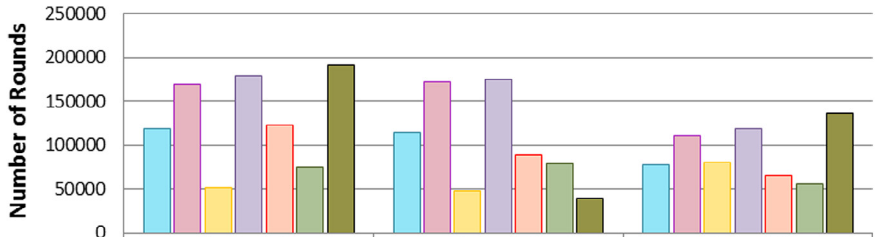
Table 11
History of dead nodes (on average) for the third scenario.

Scenario 3											
% Dead nodes	FND	%10	%20	%30	%40	HND	%60	%70	%80	%90	LND
LEACH	776.4	965.3	1069.4	1139.2	1203.6	1250.5	1287.4	1323.4	1346.4	1381.2	1476.3
LEACH-EP	1106.8	1188.1	1208.4	1223.5	1235	1246.6	1259	1274.7	1289.9	1314.7	1355.4
LEACH-FL	811.7	913.8	944.2	967.8	989.3	1015.3	1036.6	1053.5	1064.5	1074.2	1081
ASLPR	1187.7	1258.2	1271	1279.9	1287.3	1297.4	1306	1314.3	1323.2	1337.7	1362.6
SIF	657.20	1015.8	1166	1260.6	1320	1371.2	**1441.1**	**1490.8**	**1518**	**1569.9**	**1616.9**
ERA	559.70	795.9	903.8	968.1	1029.4	1068.1	1098.4	1125.3	1145	1181.5	1233.8
FMSFLA	**1368.6**	**1382.1**	**1384**	**1385.3**	**1386.4**	**1387.3**	1388.4	1389.1	1390.3	1391.7	1394.6

delivered 59,220, 26,180, 55,690, 18,090, 71,140, and 80,890 more packets to the BS than the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols, respectively.

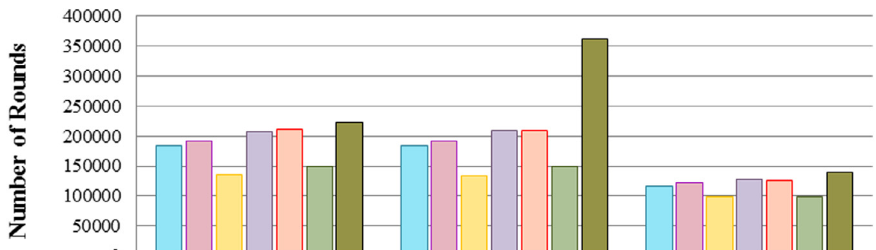
Fig. 23 shows the number of received data packets before reaching HND in the three scenarios. In the first and third scenarios, HND is less important in scenario design. Nevertheless, the FMSFLA sent more packets to the BS. In fact, it delivered 37315.4, 30333.4, 86401.1, 15,994, 12024.5, and 72770.9 more packets to the BS than the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols, respectively, in the first scenario. The goal of the second scenario is to increase HND. More packets are received by the BS in this scenario than the other two. In fact, the BS received 178,244, 171,288, 228601.6, 153536.6, 153959.1, and 212407.9 more packets in the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA, respectively.

The FMSFLA considers the delay parameter to some extent. There is also a tradeoff between delay and energy. Given the fact that the main goal of this paper is to increase the network lifetime and decrease energy consumption, the protocol should become single-hop in every space to send information to the BS when a delay is supposed to decrease. Thus, energy consumption increases in large spaces. However, the FMSFLA adjusts the value



	Scenario1	Scenario2	Scenario3
LEACH	118560	115360	77540
LEACH-EP	169740	172360	110580
LEACH-FL	51640	47920	81070
ASLPR	179300	174680	118670
SIF	122840	88120	65620
ERA	74840	78940	55870
FMSFLA	191380	39000	136760

Fig. 22. Number of received data packets to the BS till FND in various scenarios.



	Scenario1	Scenario2	Scenario3
LEACH	184985.5	184019.3	116418.8
LEACH-EP	191967.5	190975.3	122843.2
LEACH-FL	135899.8	133661.7	98197.9
ASLPR	206306.9	208726.7	128444.1
SIF	210276.4	208304.2	126506.1
ERA	149530	149855.4	98718.8
FMSFLA	222300.9	362263.3	138579.6

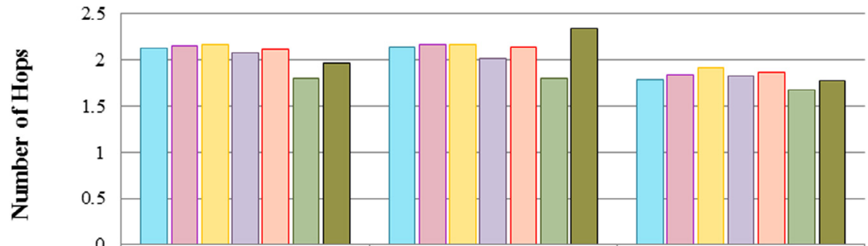
Fig. 23. Number of received data packets to BS till HND in various scenarios.

of delay in the objective function Eq. (19) as much as 0.01. According to Fig. 24, the proposed protocol could take fewer hops than the other methods except for the ERA protocol in the first and third scenarios. Thus, it could result in a shorter delay in sending information to the BS. The goal of the proposed protocol is to increase HND in the second scenario. In fact, it increased the number of CHs intelligently. Thus, the delay increased slightly more than other methods. The proposed protocol functioned more favorably than the other methods in increasing the network lifetime, energy consumption management, and all evaluation parameters. This little bit of delay can therefore be ignored.

Fig. 25 shows the mean number of CHs for every protocol in different scenarios. In fact, this figure indicates the intelligent function of the FMSFLA with respect to the goals of every scenario in CH selection regarding optimized parameters (Table 2) in comparison with other protocols. Moreover, the number of CHs is independent of node distribution conditions. Given the application features and the proper distribution of CHs in the network, different scenarios are considered for the proposed protocol. The first and second scenarios are almost identical. They merely differ in weighted coefficients of the objective function in Eq. (19).

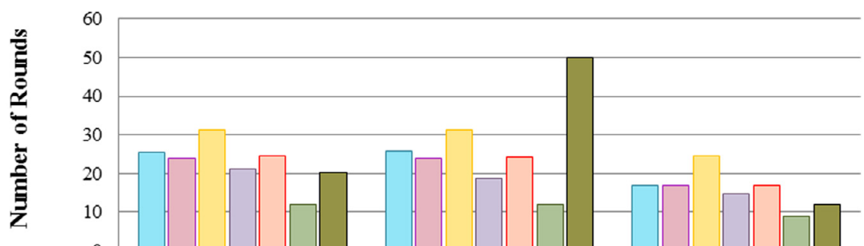
Fig. 26 indicates the value of the fitness function in the three scenarios. The fitness function Eq. (19) is used in Section 7.1 to produce proper results for every protocol. Accordingly, the FMSFLA shows greater values of fitness than the other protocols in all three scenarios.

Tables 12–14 show the minimum and maximum distances between two adjacent CHs during the entire network lifetime for all of the evaluated methods in three scenarios until the FND on average. Accordingly, the FMSFLA resulted in better distribution rates of CHs in every scenario. Put simply, the FMSFLA (see Fig. 25) guarantees the proper distribution of CHs in the network based on applications for energy management and prolonged network lifetime. Considering the number of CHs selected by the proposed FMSFLA protocol and the effect of the number of CHs and parameters of Min_OV, and Max_OV in Eq. (5) (The Min_OV and Max_OV parameters are optimized and stored for each scenario), on the factors of the minimum and the maximum distance between adjacent CHs, the proposed FMSFLA protocol has a better difference (there is a shorter distance between the maximum and a minimum distance between the two CHs), thereby ensuring the proper distribution of CHs in the network, according to its application, in terms of energy management and longer network lifetime. For instance, the average distances between CHs are 89.9, 81.1, 66, 90.1, 66.5, 104.7, and 44.9 for the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, ERA, and FMSFLA protocols, respectively, in the first scenario. In the second scenario, the results are 89.2, 90.6, 65.1, 90.5, 69.9, 106.3, and 53.3, respectively. In the third



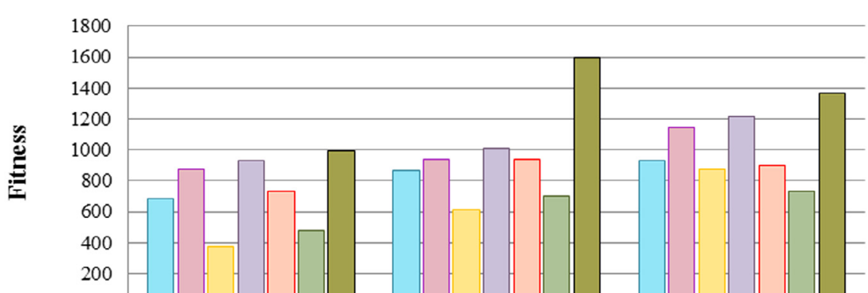
	Scenario1	Scenario2	Scenario3
LEACH	2.13	2.14	1.79
LEACH-EP	2.15	2.16	1.84
LEACH-FL	2.17	2.17	1.91
ASLPR	2.08	2.01	1.83
SIF	2.11	2.14	1.86
ERA	1.8	1.8	1.67
FMSFLA	1.97	2.34	1.78

Fig. 24. Delay value for all protocols in different scenarios.



	Scenario1	Scenario2	Scenario3
LEACH	25.58	25.72	16.98
LEACH-EP	23.98	23.82	16.74
LEACH-FL	31.13	31.29	24.38
ASLPR	21.16	18.75	14.77
SIF	24.54	24.3	16.69
ERA	12.01	11.92	8.97
FMSFLA	20.08	49.9	11.98

Fig. 25. Mean number of CHs generated in three scenarios.



	Scenario1	Scenario2	Scenario3
LEACH	683.57	870.02	932.47
LEACH-EP	874.22	937.52	1147.66
LEACH-FL	375.73	609.13	875.47
ASLPR	928.323	1011.97	1217.96
SIF	733.92	940.34	896.12
ERA	477.37	700.42	729.76
FMSFLA	994.5	1600.71	1367.05

Fig. 26. Value of fitness function for all protocols in different scenarios.

scenario, the results are 63.5, 65.5, 63, 56.8, 29.2, 67.8, and 13.5, respectively.

To evaluate the scalability of the FMSFLA, all protocols are simulated under similar conditions by adjusting weighted coefficients, such as the first scenario for application-based protocols

Table 12
Minimum and maximum distance between adjacent clusters in the first scenario.

Scenario 1	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
1	11.4	98.1	10.3	98.3	6.5	87.3	14.1	101.2	21.1	88.6	16.7	117.6	55.3	101.7
2	11.6	99.9	10.8	95.8	4.4	51.4	15.1	103.9	24.1	87.1	19.6	120.1	51.4	96
3	11.9	101.3	10.4	100.6	2.9	71.8	15.4	104.9	24.6	90.4	17.7	123.9	54.7	100.4
4	12.4	104	11.5	101.4	3.9	77.9	15.2	103.4	27.3	85.4	17.7	121.7	56.5	98.3
5	12.9	103.1	10	102.5	5.8	76.7	14.4	104.7	26.7	99.8	17.7	127.3	55.2	99.3
6	12.3	101.1	10.8	100.6	3.5	77.4	15	105.1	25	85.2	18.3	124.4	51.8	99.9
7	12.3	106.8	11	103.9	7.7	60.8	15.3	110.6	26.2	103.8	17.5	127.9	53.5	102.1
8	12.4	103.6	11	103	4.3	63.7	15.2	105.4	22.6	95.4	16	117.5	54	98.6
9	13.8	99.1	106	101.4	4.2	66.8	14.9	103.8	26.8	91.2	18	117.7	54.8	97.2
10	12.7	106.2	10.7	106.2	3.4	73.3	16	108.7	27.4	90.8	17.9	126.5	54.9	98.5
AVE	**12.37**	**102.32**	**20.25**	**101.37**	**4.66**	**70.71**	**15.06**	**105.17**	**25.18**	**91.77**	**17.71**	**122.46**	**54.21**	**99.2**

Table 13
Minimum and maximum distance between adjacent clusters in the second scenario.

Scenario 2	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
1	11.3	98.3	10.2	98.2	3.7	76.3	17.7	105.3	23.3	93.2	16.7	117.6	22.6	77.1
2	12.3	99.7	11	97.8	4.3	51.4	18.4	107.8	26.4	98.1	18.6	121	23.9	76.8
3	12.7	100	10.5	100.4	2.4	67.3	19.7	108.1	25.8	94.8	18.7	123.6	22.3	76.1
4	12.8	100.7	11.3	101.6	3.9	78.8	19.8	107.4	26.2	90.3	17.6	122	23.8	76.1
5	11.4	101.5	10.2	101.4	4.1	81.9	19	111	26.9	98	16.3	126	24.6	79.1
6	12.1	103.1	10.6	102.1	2.5	74.6	18.9	109.8	24	94.5	17.8	124.3	22	74.2
7	13.2	104.1	10.9	105.4	7.7	60.7	19.1	115.8	22.4	94	17.4	127.4	24.6	74.7
8	12.5	103.7	11.2	101.4	4.4	61.8	19.4	109.8	28.2	102.6	19.2	123.2	26.2	72.2
9	12.4	100	10.5	101.5	4.2	67.9	20.1	109	22.6	90	17.9	124.3	23.8	80.8
10	12.7	104.8	11.1	104.4	2.4	70.6	20	113.9	26	95.9	16.3	130.3	24.2	84.4
AVE	**12.34**	**101.59**	**10.75**	**101.42**	**3.96**	**69.13**	**19.21**	**109.79**	**25.18**	**95.14**	**17.65**	**123.97**	**23.8**	**77.15**

Table 14
Minimum and maximum distance between adjacent clusters in the third scenario.

Scenario 3	LEACH		LEACH-EP		LEACH-FL		ASLPR		SIF		ERA		FMSFLA	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
1	14.17	76.87	12.31	77.13	7.52	61.08	19.45	76.75	28.71	51.92	21.34	91.66	42.27	60.21
2	13.88	77.94	12.56	75.27	8.33	81.74	20.14	74.33	28.77	51.10	21.62	88.25	40.74	58.60
3	13.62	75.57	11.14	77.55	3.67	70.37	18.70	75.71	24.25	58.85	21.03	86.22	54.89	52.06
4	14.54	75.74	12.91	74.7	8.43	68.26	20.06	74.75	21.86	58.83	22.23	88.23	42.51	59.61
5	15.88	76.96	13.83	76.73	10.67	66.44	21.27	74.87	25.66	57.65	22.73	90.74	41.24	61.93
6	15.28	76.10	11.89	76.14	4.69	66.82	19.99	75.88	28.72	57.09	23.27	86.24	45.62	57.87
7	12.95	80.89	10.28	79.57	6.96	78.66	17.87	76.43	25.59	59.07	20.85	89.70	46.02	59.11
8	14.42	77.40	12.10	75.99	7.64	58.57	19.11	75.03	25.97	58.85	21.52	90.33	45.74	59.63
9	13.78	80.29	11.59	81.04	5.11	75.82	19.04	79.94	29.98	53.04	21.18	92.19	49.56	58.98
10	13.94	79.74	10.91	80.47	4.41	69.89	18.41	79.12	27.89	53.45	21.55	91.83	44.75	60.32
AVE	**14.24**	**77.75**	**11.95**	**77.45**	**6.74**	**69.76**	**19.40**	**76.28**	**26.74**	**55.98**	**21.73**	**89.53**	**45.33**	**58.83**

(ASLPR and SIF). The protocols are simulated in 100×100 m^2, 200×200 m^2, 300×300 m^2, 400×400 m^2, and 500×500 m^2 spaces with 200 nodes, scattered at random, when the BS is situated at the center of every space. Fig. 27 shows the number of FND round. Accordingly, the proposed protocol shows greater values of FND in all spaces and more rounds than the other protocols. Thus, the FMSFLA is scalable enough in different spaces.

8. Conclusion

This paper presents a fuzzy multi-hop clustering protocol (FMSFLA), which not only is energy-efficient, but it can also manage energy consumptions of nodes for optimization based on FIS rule-base table and five control parameters by using the shuffled frog leaping algorithm (SFLA) for application management. The proposed protocol considers effective parameters including energy, distance to the BS, the number of neighboring nodes, real distance between a node and the BS, mean route load, delay, overlap, and the problem of hot spots with the purpose of achieving the best function. According to the simulation results, the FMSFLA performed better than the other protocols in all scenarios with respect to the goals and applications. In other words, the stable area (until FND) of the first scenario increased by 61%, 13%, 270%, 7%, 56%, and 155% in the proposed protocol compared with the LEACH, LEACH-EP, LEACH-FL, ASLPR, SIF, and ERA protocols, respectively. Regarding the second scenario, the pre-HND area is increased by 115%, 116%, 183%, 97%, 88%, and 160%, respectively, to meet the scenario goals. Given the optimized parameters along with the two optimized fuzzy rule tables, the proper distribution of CHs is also guaranteed in the network based on the application features in addition to the increased network lifetime. It should be noted, however, that despite all the benefits of the proposed protocol, it requires off-line optimization and has a time complexity that is negligible due to the benefits of this protocol.

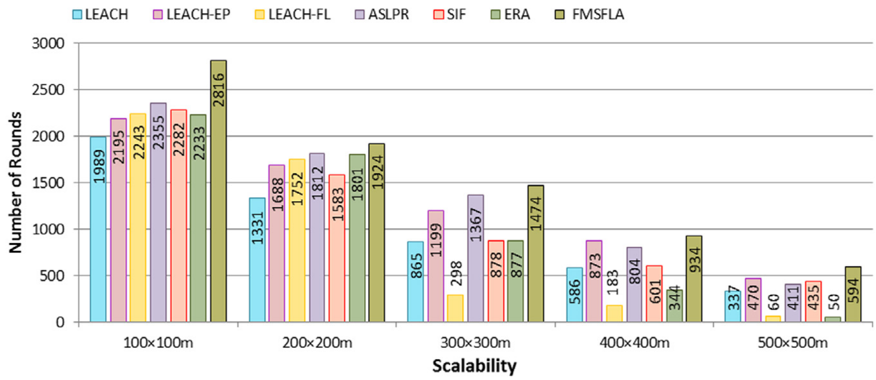


Fig. 27. The effect of changing network size on FND for all protocols.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.asoc.2020.106115.

Acknowledgments

The authors would like to express their thanks to the anonymous referees for their valuable comments and suggestions that improved the paper.

References

[1] A.D. Amis, R. Prakash, T.H. Vuong, D.T. Huynh, Max–min D-cluster formation in wireless ad hoc networks, in: Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM' 01, Tel Aviv, Israel, March 26–30, 2000, pp. 32–41.
[2] X.Y. Kui, J.X. Wang, S.G. Zhang, Energy-balanced clustering protocol for data gathering in wireless sensor networks with unbalanced traffic load, J. Cent. South Univ. 19 (11) (2012) 3180–3187.
[3] S.-H. Moon, S. Park, S.-J. Han, Energy efficient data collection in sink-centric wireless sensor networks: A cluster-ring approach, Comput. Commun. 101 (2017) 12–25.
[4] M.M. Afsar, M.H. Tayarani-N., Clustering in sensor networks: A literature survey, J. Netw. Comput. Appl. 46 (2014) 198–226.
[5] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Trans. Wireless Commun. 1 (4) (2002) 660–670.
[6] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, in: Proceedings of the IEEE International Conference on System Sciences (ICSS), Washington, USA, vol. 2, January 04–07, 2000, pp. 1–10.
[7] D. Chen, P.K. Varshney, QoS support in wireless sensor networks: A Survey, in: Proceedings of the International Conference on Wireless Networks, ICWN '04, Nevada, USA, June 21–24, 2004, pp. 227–233.
[8] J.S. Lee, W.L. Cheng, Fuzzy-logic-based clustering approach for wireless sensor networks using energy predication, IEEE Sens. J. 12 (9) (2012) 2891–2897.
[9] F. Fanian, M. Kuchaki Rafsanjani, Cluster-based routing protocols in wireless sensor networks: A survey based on methodology, J. Netw. Comput. Appl. 142 (2019) 111–142.
[10] D. Wang, L. Lin, L. Xu, A study of subdividing hexagon-clustered WSN for power saving: Analysis and simulation, Ad Hoc Netw. 9 (7) (2011) 1302–1311.
[11] L. Khelladi, D. Djenouri, M. Rossi, N. Badache, Efficient on-demand multi-node charging techniques for wireless sensor networks, Comput. Commun. 101 (2017) 44–56.
[12] B. Rashid, M.H. Rehmani, Applications of wireless sensor networks for urban areas: A survey, J. Netw. Comput. Appl. 60 (2016) 192–219.
[13] B. Avci, G. Trajcevski, R. Tamassia, P. Scheuermann, F. Zhou, Efficient detection of motion-trend predicates in wireless sensor networks, Comput. Commun. 101 (2017) 26–43.
[14] B. Mamalis, D. Gavalas, C. Konstantopoulos, G. Pantziou, Clustering in wireless sensor networks, in: Y. Zhang, L.T. Yang, J. Chen (Eds.), RFID and Sensor Networks, CRC Press, 2010.
[15] L. Frye, L. Cheng, Topology management for wireless sensor networks, in: S. Misra, I. Woungang, S.C. Misra (Eds.), Guide to Wireless Sensor Networks, Springer, 2009.
[16] A.A. Abbasi, M. Younis, A survey on clustering algorithms for wireless sensor networks, Comput. Commun. 30 (14) (2007) 2826–2841.
[17] M. Sarrafzadeh, F. Dabiri, R. Jafari, T. Massey, A. Nahapetian, Low power light-weight embedded systems, in: Proceedings of the International Symposium on Low Power Electronics and Design (ISLPED), Bavaria, Germany, October 04–06, 2006, pp. 207–212.
[18] R. Jafari, A. Encarnacao, A. Zahoory, F. Dabiri, H. Noshadi, M. Sarrafzadeh, Wireless sensor networks for health monitoring, in: Proceedings of the IEEE International Conference on Mobile and Ubiquitous Systems (MobiQuitous), San Diego, USA, July 17-21, 2005, pp. 479–481.
[19] J.A. Stankovic, Wireless sensor networks, Computer 41 (10) (2008).
[20] Su. Wiliam, B.A. Ozgur, E. Cayirci, Communication protocols for sensor networks, in: C.S. Raghavendra, K.M. Sivalingam, T. Znati (Eds.), Wireless Sensor Networks, Springer, 2006.
[21] M.R.M. Kassim, I. Mat, A.N. Harun, Wireless Sensor Network in precision agriculture application, in: Proceedings of the Computer, Information and Telecommunication Systems (CITS), Jeju, South Korea, July 7–9, 2014, pp. 1–5.
[22] J. Hong, J. Kook, S. Lee, D. Kwon, S. Yi, T-LEACH: The method of threshold-based cluster head replacement for wireless sensor networks, Inf. Syst. Front. 11 (5) (2009) 513–521.
[23] J.-G. Jia, Z.-W. He, J.-M. Kuang, Y.-H. Mu, An energy consumption balanced clustering algorithm for wireless sensor network, in: Proceedings of the 6th International Conferences on Wireless Communications Networking and Mobile Computing (WiCOM), Chengdu, China, September 23–25, 2010, pp. 1–4.
[24] S.H. Kang, T. Nguyen, Distance based thresholds for cluster head selection in wireless sensor networks, IEEE Commun. Lett. 16 (9) (2012) 1396–1399.
[25] J. Bagherzadeh, M. Samadzamini, A clustering algorithm for wireless sensor networks based on density of sensors, in: Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia (MoMM), Kuala Lumpur, Malaysia, Dec. 14–16, 2009, pp. 594–598.
[26] A. Wang, D. Yang, D. Sun, A clustering algorithm based on energy information and cluster heads expectation for wireless sensor networks, Comput. Electr. Eng. 38 (3) (2012) 662–671.
[27] G. Ran, H. Zhang, S. Gong, Improving on LEACH protocol of wireless sensor networks using fuzzy logic, J. Inf. Comput. Sci. 7 (3) (2010) 767–775.
[28] M. Mirzaie, S.M. Mazinani, Adaptive MCFL: An adaptive multi-clustering algorithm using fuzzy logic in wireless sensor network, Comput. Commun. 111 (1) (2017) 56–67.
[29] M.O. Oladimeji, M. Turkey, S. Dudley, HACH: Heuristic algorithm for clustering hierarchy protocol in wireless sensor networks, Appl. Soft Comput. 55 (2017) 452–461.
[30] D.C. Hoang, P. Yadav, R. Kumar, S.K. Panda, Real-time implementation of a harmony search algorithm-based clustering protocol for energy-efficient wireless sensor networks, IEEE Trans. Ind. Inform. 10 (1) (2014) 774–783.
[31] E.A. Khalil, A.A. Bara'a, Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks, Swarm Evol. Comput. 1 (2011) 195–203.
[32] M. Shokouhifar, A. Jalali, A new evolutionary based application specific routing protocol for clustered wireless sensor networks, AEU-Int. J. Electron. Commun. 69 (1) (2015) 432–441.
[33] Z.M. Zahedi, R. Akbari, M. Shokouhifar, F. Safaei, A. Jalali, Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks, Expert Syst. Appl. 55 (2016) 313–328.
[34] M. Shokouhifar, A. Jalali, Optimized sugeno fuzzy clustering algorithm for wireless sensor networks, Eng. Appl. Artif. Intell. 60 (2017) 16–25.

[35] F. Fanian, M. Kuchaki Rafsanjani, Memetic fuzzy clustering protocol for wireless sensor networks: Shuffled frog leaping algorithm, Appl. Soft Comput. 71 (2018) 568–590.
[36] D.T. Haia, L.H. Sonb, T. Le Vinh, Novel fuzzy clustering scheme for 3D wireless sensor networks, Appl. Soft Comput. 54 (2017) 141–149.
[37] M. Arioua, Y. El Assari, I. Ez-Zazi, A. El Oualkadi, Multi-hop cluster based routing approach for wireless sensor networks, Procedia Comput. Sci. 83 (2016) 584–591.
[38] A. Shahraki, M. Kuchaki Rafsanjani, A. Borumand Saeid, A new approach for energy and delay trade-off intra-clustering routing in WSNs, Comput. Math. Appl. 62 (4) (2011) 1670–1676.
[39] M. Sabet, H.R. Naji, A decentralized energy efficient hierarchical cluster-based routing algorithm for wireless sensor networks, AEU-Int. J. Electron. Commun. 69 (5) (2015) 790–799.
[40] M. Sabet, H. Naji, An energy efficient multi-level route-aware clustering algorithm for wireless sensor networks: A self-organized approach, Comput. Electr. Eng. 56 (2016) 399–417.
[41] A. Shahraki, M. Kuchaki Rafsanjani, A. Borumand Saeid, Hierarchical distributed management clustering protocol for wireless sensor networks, Telecommun. Syst. 65 (1) (2017) 193–214.
[42] Y. Tao, Y. Zhang, Y. Ji, Flow-balanced routing for multi-hop clustered wireless sensor networks, Ad hoc Netw. 11 (1) (2013) 541–554.
[43] J.-S. Lee, T.-Y. Kao, An improved three-layer low-energy adaptive clustering hierarchy for wireless sensor networks, IEEE Internet Things J. 3 (6) (2016) 951–958.
[44] T. Amgoth, P.K. Jana, Energy-aware routing algorithm for wireless sensor networks, Comput. Electr. Eng. 41 (2015) 357–367.
[45] K. Suganthi, B. Vinayagasundaram, J. Aarthi, Randomized fault-tolerant virtual backbone tree to improve the lifetime of wireless sensor networks, Comput. Electr. Eng. 48 (2015) 286–297.
[46] D. Yi, H. Yang, HEER–A Delay-aware and energy-efficient routing protocol for wireless sensor networks, Comput. Netw. 104 (2016) 155–173.
[47] B. Zhou, A. Marshall, T.-H. Lee, An energy-aware virtual backbone tree for wireless sensor networks, in: Proceedings of the 5th IEEE Global Telecommunications Conference (GLOBECOM), St. Louis, MO, USA, Nov. 28–Dec. 2, 2005, pp. 1–6.
[48] J. Kim, J.-H. Lee, ViTAMin: A Virtual Backbone Tree Algorithm for Minimal energy consumption in wireless sensor network routing, in: Proceedings of the International Conference on Information Network (ICOIN), Bali, Indonesia, Feb. 1–3, 2012, pp. 144–149.
[49] J. Kim, K. Yoon, S. Lee, J.-h. Jung, J.-H. Lee, An m-EVBT algorithm for energy efficient routing in wireless sensor networks, in: Proceedings of the 3th International Conference on Ubiquitous Information Management and Communication (ICUIMC), Suwon, Korea, Jan. 15–16, 2009, pp. 586–591.
[50] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[51] S. Gajjar, A. Talati, M. Sarkar, K. Dasgupta, FUCP: Fuzzy based unequal clustering protocol for wireless sensor networks, in: Proceedings of the National Systems Conference (NSC), Noida, India, Dec 14–16, 2015, pp. 1–6.
[52] R. Logambigai, A. Kannan, Fuzzy logic based unequal clustering for wireless sensor networks, Wirel. Netw. 22 (3) (2016) 945–957.
[53] S.A. Sert, H. Bagci, A. Yazici, MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks, Appl. Soft Comput. 30 (2015) 151–165.
[54] H. Bagci, A. Yazici, An energy aware fuzzy approach to unequal clustering in wireless sensor networks, Appl. Soft Comput. 13 (4) (2013) 1741–1749.
[55] B. Baranidharan, B. Santhi, DUCF: Distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach, Appl. Soft Comput. 40 (2016) 495–506.
[56] I.H. Witten, E. Frank, M.A. Hall, Data Mining Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2011.
[57] E.H. Mamdani, S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, Int. J. Man-Mach. Stud. 7 (1975) 1–13.
[58] T.J. Ross, Fuzzy Logic with Engineering Applications, John Wiley & Sons, 2004.
[59] W.Y. Chiu, B.S. Chen, Multisource prediction under nonlinear dynamics in WSNS using a robust fuzzy approach, IEEE Trans. Circuits Syst. I 58 (2011) 137–149.
[60] H.M. Ammari, On the energy-delay trade-off in geographic forwarding in always-on wireless sensor networks: A multi-objective optimization problem, Comput. Netw. 57 (9) (2013) 1913–1935.
[61] T.-T. Huynh, A.-V. Dinh-Duc, C.-H. Tran, Delay-constrained energy-efficient cluster-based multi-hop routing in wireless sensor networks, J. Commun. Netw. 18 (4) (2016) 580–588.
[62] F. Xunli, D. Feiefi, Shuffled frog leaping algorithm based unequal clustering strategy for wireless sensor networks, Appl. Math. 9 (3) (2015) 1415–1426.
[63] Q.Y. Duan, V.K. Gupta, S. Sorooshian, Shuffled complex evolution approach for effective and efficient global minimization, J. Optim. Theory Appl. 76 (3) (1993) 501–521.
[64] K.M. Passino, P.J. Antsaklis, A metric space approach to the specification of the heuristic function for the A* algorithm, IEEE Trans. Syst. Man Cybern. 24 (1) (1994) 159–166.
[65] I.S. AlShawi, L. Yan, W. Pan, B. Luo, Lifetime enhancement in wireless sensor networks using fuzzy approach and A-star algorithm, IEEE Sens. J. 12 (10) (2012) 3010–3018.