Soft Computing (2021) 25:5663–5682
https://doi.org/10.1007/s00500-020-05563-7

METHODOLOGIES AND APPLICATION

Clustering based on whale optimization algorithm for IoT over wireless nodes

Seyed Mostafa Bozorgi[1] · Mahdi Rohani Hajiabadi[2] · Ali Asghar Rahmani Hosseinabadi[3] · Arun Kumar Sangaiah[4]

Published online: 15 January 2021

Abstract

IoT or Internet of Things can improve the possibility of interaction between various smart components in real time. In the infrastructure of IoT, wireless sensors can be used in order to reduce communication costs. Despite having positive effects, using wireless nodes add some challenges to the system. Limited resources, such as energy, CPU power and memory, are the main concerns in this technology. Energy consumption is the most challenging one. Designing an optimized routing pattern through heuristic algorithms is a common way to tackle this problem. Therefore, in the proposed algorithm, a WOA-based method has been proposed to expand the life span of the system. Also, a novel fitness function is defined for reducing the energy consumption of the network, load balancing and node coverage. Clustering is done unequally; it means that cluster heads (CHs) nearer to the base station (BS) have more energy for data relay. In this paper, for reducing the number of messages, a clustering stage is added at the beginning of each metaround. The number of rounds in a metaround is variable. The status of each node is analyzed by BS before each round. Low energy level causes a new metaround. Moreover, the CH–BS interaction is implemented through multi-hop pattern. Results suggest that there is an enhancement instability, energy-saving, throughput and lifespan.

Keywords Internet of Things · Wireless sensor networks · Whale optimization algorithm · Unequal clustering · Network lifetime

1 Introduction

In the Internet of Things (IoT) domain, each object or thing can use wireless communication to communicate with each other (Machado et al. 2013). Today, IoT has attracted the attention of societies, governments and industries for a wide range of applications, including smart homes, healthcare services, environmental monitoring, smart transportation, smart networks, security, fire detection, finance tracking, smart lighting, etc. (Abdul-Qawy and Srinivasulu 2018). In this context, wireless sensor networks (WSNs) play an important role in widening networks with low-cost smart devices that can be easily installed (Shah et al. 2018). In recent years, WSNs have attracted attention in many applications including environmental monitoring, predicting natural disasters, health monitoring and military applications (Heinzelman et al. 2002; Bozorgi et al. 2017). In these networks, nodes have limited battery and processing strength. Energy is a serious concern in these

✉ Arun Kumar Sangaiah
Arunkumarsangaiah@gmail.com

Seyed Mostafa Bozorgi
M.Bozorgi@iau-tnb.ac.ir

Mahdi Rohani Hajiabadi
Master.Rohani@gmail.com

Ali Asghar Rahmani Hosseinabadi
ark838@uregina.ca

1 Department of Computer Engineering, Tehran North Branch, Islamic Azad University, Tehran, Iran

2 Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

3 Department of Computer Science, University of Regina, Regina, Canada

4 School of Computing Science and Engineering, Vellore Institute of Technology (VIT), Vellore 632014, India

structures because, in many applications, either network's nodes are not accessible or cannot be replaced (Bozorgi et al. 2016; Kuila et al. 2013). Since data transmission in wireless communications consumes more energy compared to processing, data routing and data transfer in these networks are of great importance (Kumar and Kumar 2016).

Clustering is one of the effective methods for saving energy (Kuila et al. 2013). Clustering is a process that divides nodes into clusters. Each cluster has a cluster head (CH) and several cluster members (CMs). A CM senses the environment's information and transmits them to the CH. CH collects and integrates the information transmitted by CMs and then transmits information to the base station (BS) (Kuila and Jana 2014; Afsar and Tayarani-N 2014). The clustering problem in IoT/WSN is an NP-hard problem. Computational intelligence has been widely employed for improving these challenges. Different computational intelligence methods including evolutionary algorithms (EAs) were used for routing in WSN (Khalil and Attea 2011). EAs have been used for optimization in many problems. In recent years, algorithms like a genetic algorithm (GA) (Deb 2000), differential evolution (DE) (Brest et al. 2006) and particle swarm optimization (PSO) (Clerc and Kennedy 2002; Eberhart and Shi 2004) were proposed as optimization algorithms (Ma and Simon 2011; Hosseinabadi et al. 2019; Sangaiah et al. 2019). WOA or Whale Optimization Algorithm is a new evolutionary method to solve optimizing problems. This algorithm drew inspiration from the hunting process in whales. Generating a random whale population is the first step of the algorithm. In the next step, WOA functions are used to lead the whales toward prey (the optimum solution).

Existing clustering protocols are mainly time-oriented. These protocols are categorized as static, dynamic and hybrid. In the static method, clustering is done once and then CH rotation is performed. As an example,virtual concentric circle band-based clustering (VCCBC) (Kumar et al. 2011) and an energy-efficient protocol with static clustering (EEPSC) (Chaurasiya et al. 2011) are static clustering protocols. In static performance, overhead is low and the stability of the network lasts for a short time. A shortcoming of a static method is because energy is discharged at several nodes (Malathi et al. 2015). In the dynamic method, clustering is done at each round. One of the most well-known dynamic protocols is low-energy adaptive clustering hierarchy (LEACH) (Heinzelman et al. 2002). In a dynamic performance, the lifetime of the network can be improved but its overhead is high (Malathi et al. 2015). The hybrid method, not only improves stability and lifetime but also reduces overhead. Recently, hybrid static–dynamic methods have been proposed for clustering; among them, hybrid unequal clustering with layering protocol (HUCL) (Malathi et al. 2015) can be named. These methods employ both static and dynamic methods for clustering. After a specific number of rounds, clustering and formation of new clusters are performed. Also, CH rotation between clustering rounds and new cluster formation is performed. Against these time-oriented methods, the existence of those that are energy-oriented. Unlike, time-oriented approaches like LEACH (Heinzelman et al. 2002), which are selected at the beginning of each round of CH, in the proposed approach, CHs only change when their energy is not suitable. This way overhead is considerably reduced. Table 1 compares the existing methods. In the following, a different method has been investigated.

Assuming that BS knows the network down to details, the clustering process in carried out by base station (Zanjireh and Larijani 2015). The clustering process is implemented through WOA. Saving more energy and load balancing are the main objectives in the fitness function. The performance of the algorithm has been analyzed in terms of energy consumption. This way overhead is considerably reduced. Moreover, for transmitting data from CH to BS, energy-aware multi-hop routing is used. Also, a novel mechanism is used which prevents a node from sending conventional control messages like the head message and join message. Therefore, except for one control message at the beginning of the metaround with name status message, nodes do not broadcast any other control messages.

The structure of the paper is as follows: literature review in Sect. 2, the model of the studied network in Sect. 3, the proposed method in Sect. 4, implementation and simulation results in Sect. 5 and conclusion in Sect. 6.

2 Literature review

Much effort has been done on improving energy consumption using heuristic and meta-heuristic clustering methods. One of the most important clustering methods proposed was LEACH, (Heinzelman et al. 2002; Rostami et al. 2016, 2018). This algorithm constructs clusters randomly and distributed. The role of CH changes from one node to another so that battery is not discharged at one node. Other nodes are connected to CHs that require minimum energy. This method does not consider the remained energy of nodes for selecting the CH.

Other protocols like hybrid energy-efficient distributed clustering (HEED) are also proposed (Younis et al. 2004; Han et al. 2019; Saemi et al. 2016). This algorithm combines remained energy and communication cost as a criterion for selecting CH. A node with higher energy is more probable to be CH. Communication cost might be the node degree of inverse node degree depending on the

Table 1 Comparison of advantages and disadvantages of different clustering methods

Methods	Operation	Advantages	Disadvantages
Static	Clustering is only performed once and then intra-CH rotation is performed	Reducing overhead of control messages	Reducing stability of the network
Dynamic	Clustering is performed at the beginning of each round	Increasing stability of a network	Increasing overhead of control messages
Hybrid	Clustering is performed once every few rounds and intra-CH rotation is performed	Increasing stability of a network Reducing overhead of control messages Reliability against unpredicted events	Difficulty in setting time of clustering and performing intra-CH rotation The complexity of the protocol
Energy oriented	Clustering is performed when the energy of the current cluster heads is not proper.	Increasing stability of a network Reducing overhead of control messages The simplicity of the protocol Reliability against unpredicted events Not requiring time-adjustment for specific operations	Energy status of the nodes should be known

requirement. Unlike LEACH, the distribution of CHs in HEED is performed very well.

In 2015, a combination of static and dynamic clustering methods named "HUCL" was proposed (Malathi et al. 2015). Clusters nearer to BS are smaller. In this method, CHs are selected according to the energy status, distance from BS and the number of neighbors. In addition, multi-hop routing is used to transmit data to BS. In this algorithm, a CH without any member becomes a CM and connects to the closest CH. Data transmission is performed at several time intervals. HUCL has reduced network overhead, optimized energy consumption and increased network lifetime. In this algorithm, the energy of nodes is not considered for calculating the radius of nodes.

An improved energy-aware distributed unequal clustering (EADUC-II) is proposed in 2016 (Gupta and Pandey 2016) as an improvement to the EADUC (Yu et al. 2011) and HUCL (Malathi et al. 2015) protocols. In this method, clusters with unequal sizes were constructed where clusters close to BS were smaller. Another parameter including the energy of the node was also considered in determining the competition radius of clusters. In this method, the energy of a node was considered for selecting the next hop for routing in inter-cluster transmission. This method did not consider the density of nodes and their distance from BS for calculating the delay time of a node. Although this method is performed a s hybrid method, it reduces overhead less than HUCL method.

In recent years, several meta-heuristic methods were proposed for increasing lifetime of WSNs. In 2011, a method entitled as energy-aware evolutionary routing protocol (EAERP) was proposed (Khalil and Attea 2011). The protocol's aim was to optimize stability time or first node death (FND) time and lifetime of the network or last node death (LND) simultaneously. Fitness function was defined as the summation of energy consumed for transmitting nodes to CHs or receiving them from CHs and energy required for transmitting data to BS. Simulation results showed that this method has achieved this goal by proposing a suitable fitness function. However, these methods involve overload. Routing in this method is performed in single-hop mode and since it is centralized, it was not suitable for the high-dimensional environment.

An application specific low-power routing protocol (ASLPR) based on LEACH architecture was proposed in 2014 (Shokouhifar and Jalali 2015). This method considered the distance from BS, remained energy and distance between clusters for determining the proper CH. Considering the application of the network, in order to achieve the best performance, adjusting parameters of a routing protocol is very important. In order to solve this challenge, a hybrid of GA and simulated annealing (SA) is proposed to optimize lifetime of the network and best performance based on application. Simulation results showed that this method could increase the lifetime of the network by saving energy. Despite aforementioned advantages, the proposed method requires additional computations in the central processor in the BS for selecting the optimum CH.

Another algorithm was a stable-aware evolutionary routing protocol (SAERP), proposed in 2013 (Khalil and Attea 2013). This protocol was proposed to improve the stability of the network. Its fitness function was the same as

EAERP algorithm. Its difference with the previous method was in generating the initial population. Only those nodes whose energy was higher than an average energy value could be considered as a candidate selected as CH.

In another study, a differential evolution (DECA) based clustering algorithm was proposed for WSN (Kuila and Jana 2014). In which an efficient vector encoding scheme and an extra phase called local improvement is suggested for improving the performance of the clustering algorithm. Moreover, the proposed method has offered an efficient fitness function for increasing the lifetime of the network, where this fitness function considered energy consumption in sensor nodes and gateway nodes. This method prevents the quick death of CHs. However, despite all improvements, this method is a centralized method with high overhead.

Another protocol was proposed in 2016 (Azharuddin and Jana 2016) in which a clustering algorithm is performed based on particle swarm optimization (CAPSO) and looks to find a solution for hotspot problem caused by multi-hop communications. In the routing stage, traffic load was distributed equally among CHs and in the clustering phase, CHs that did not have proper energy would have less CMs. Moreover, a technique was proposed to prevent and remove traffic load of gateways, which did not have proper energy to increase the lifetime of the network. However, among improvements resulted from this method, the transient failure of CHs could be considered as a disadvantage.

In 2016, an unequal multi-hop balanced immune clustering protocol (UMBIC) was proposed (Sabor et al. 2016). This method was proposed to resolve hot spot for CHs close to BS in multi-hop routing and improving the lifetime of the network for networks with different sizes and different homogeneous and heterogeneous nodes. This protocol has employed unequal clustering mechanism for optimizing energy consumption in intra-cluster and inter-cluster transmission in order to construct clusters with unequal sizes according to the distance of a node to a BS and energy level of a node and also used a multi-objective immune algorithm in order to construct a routing tree for minimizing communication costs. This protocol has changed the role of CH among nodes to reduce the overhead of a network only when the remained energy of one CH was less than a threshold. This method has improved lifetime of the network and reduced overhead. Despite the overhead reduction, this method used a traditional control message in the cluster building phase, which in some way increased the cost and overhead of this phase.

A hybrid HAS-PSO algorithm for energy-efficient CH selection was proposed in 2016 (Shankar et al. 2016). In this method, two harmony search algorithms, HSA and PSO, were combined for faster convergence to select proper CHs in WSN clustering. This algorithm has employed the high search efficiency of HSA method and dynamic capability of PSO to reach a trade-off between exploration and exploitation. Simulation results indicated that throughput has improved. However, in this method, overhead is considerably high.

Another protocol was integrated clustering and routing protocol for wireless sensor networks using cuckoo and harmony search (iCSHS), which was proposed in 2018 (Gupta and Jha 2018). In this paper, the CH selection was performed using the cuckoo algorithm and routing between CHs and BS on large scale using the harmony search algorithm. The clustering algorithm used a new encoding mechanism and a new fitness function. CH selection was done in such a way that the CHs were uniformly distributed in covered areas. This protocol has had a good improvement over previous protocols. However, the overhead continued to be a challenge.

Despite all the methods that have been proposed in recent years, clustering is still a challenge. In this section, clustering protocols including algorithms based on EAs which were used to increase the lifetime of the network are investigated. In these methods, overhead is still the main challenge. Although new clustering methods were proposed, there are still some challenges in improving the lifetime of the network and increasing throughput when energy is limited. A brief comparison of clustering methods is shown in Table 2.

In this study, our main idea is to propose a clustering method with high accuracy such that the throughput of a network is increased with minimum overhead. In order to reduce overhead, the network operates in energy-oriented mode and new CHs are selected when the energy of current CHs is less than a threshold. In fact, a novel mechanism is used which prevents a node from broadcasting control messages. Therefore, except for control messages of a node's energy status that are appended to the main packets and routed to BS with minimum energy consumption, nodes do not transmit any other control messages. In the proposed method, an unequal clustering method based on WOA is proposed. BS performs clustering based on the position of nodes. Clustering is centralized. The BS determines routing for inter-cluster transmission. In the following, the proposed method is presented.

3 The network model

In this study, a wireless network-based IoT and a BS with unlimited supply connected to a network are considered. In a network of size $M \times M$, N nodes are distributed randomly. All nodes and the BS are fixed but they can adjust their transmission power according to the distance. At each

Table 2 Comparison of some of the clustering algorithms

Protocol	Years	Cluster size	Intra com.	Inter com.	Method	CH election	Dynamism
LEACH	2000	Equal	1-hop	1-hop	Distributed	Random	Dynamic
LEACH-C	2002	Equal	1-hop	1-hop	Centralized	Deterministic by BS	Dynamic
HEED	2004	Equal	1-hop	k-hop	Distributed	Hybrid, based on Attribute	Dynamic
VCCBC	2011	Equal	1-hop	k-hop	Distributed	CH rotation	Static
EEPSC	2011	Equal	1-hop	1-hop	Distributed	Selection by previous CH	Static
EAERP	2011	Equal	1-hop	1-hop	Centralized	Based on EA	Dynamic
ASLPR	2013	Equal	1-hop	1-hop	Centralized	Based on a hybrid of GA and SA	Dynamic
SAERP	2013	Equal	1-hop	1-hop	Centralized	Based on EA	Dynamic
DECA	2014	Equal	1-hop	1-hop	Centralized	Based on DE	Dynamic
HUCL	2015	Unequal	1-hop	k-hop	Distributed	Hybrid, based on Attribute	Hybrid
EADUC-II	2016	Unequal	1-hop	k-hop	Distributed	Hybrid, based on Attribute	Hybrid
iCSHS	2018	Equal	1-hop	k-hop	Centralized	Based on cuckoo and harmony search	Dynamic
proposed protocol	2018	Unequal	1-hop	k-hop	Centralized	Based on WOA	Energy oriented

round, data are sampled using sensor nodes and transmitted to BS for routing. Moreover, each node can be a CH or non-CH. CHs can collect and integrate data. The radio model of energy consumption is like LEACH protocol (Heinzelman et al. 2002). Energy consumption of transmitter and energy consumption of receiver are defined using Eqs. (1) and (2), respectively:

$$E_{TX}(i, K, d_{ij}) = \begin{cases} E_{\text{elec}}K + E_{fs}Kd_{ij}^2 & \text{if} \quad d_{ij} \le d_o \\ E_{\text{elec}}K + E_{mp}Kd_{ij}^4 & \text{if} \quad d_{ij} > d_o \end{cases} \tag{1}$$

$$E_{RX}(j, k) = E_{\text{elec}}k, \tag{2}$$

where K is the number of data bits, and d is the distance between two nodes. $E_{\text{elec}}(nj/bit)$ is energy consumption at each bit for transmitting or receiving data. $E_{mp}(pj/bit \times \text{m}^{-4})$ and $E_{fs}(pj/bit \times \text{m}^{-2})$ are energy consumed at each bit for reinforcing transmitter considering transmission distance. In addition, d_o is obtained as $d_o = \sqrt{E_{fs}/E_{mp}}$ Data aggregation model used in simulations assumes that total aggregated information can be compressed to K bits by a set of N nodes in which each node collects K bits of information.

4 The proposed algorithm

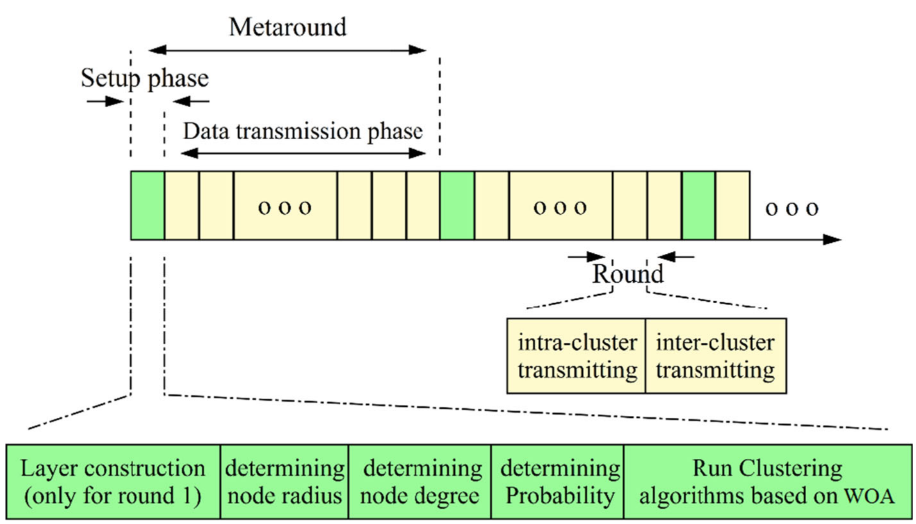
In the proposed algorithm, a novel method, called NUWC (new unequal whale optimization algorithm clustering), has been designed for clustering by means of WOA (Mirjalili and Lewis 2016). Execution of NUWC is concentrated on one single node. This algorithm aims to reduce the consumption of energy. The distance between nodes can be determined through RSSI (received signal strength indicator). In order to reduce messaging overhead, a new message called status message has been designed and used instead of traditional controlling messages. Inter-cluster interactions have been managed by means of multi-hop approach which has energy saving considerations.

The network operations are split into metarounds to decrease the clustering overhead. In a metaround, NUWC operations include a setup phase and data transmission phase. In the setup phase, BS runs WOA algorithm for clustering the nodes in the network. Data transmission phase is divided into several data transmission rounds. Each round consists of inter-cluster transmission and intra-cluster transmission. Inter-cluster routing to the BS is formulated in multi-hop form. In this method, the setup phase is only performed at the beginning of each metaround. In other words, when the energy of current CHs is less than, E_{th}, an energy threshold ($E_{th} = E_{\text{avg}} - E_{\text{std}}$); otherwise, current CHs are used for the new round. This threshold is adopted from Sabor et al. (2016). E_{avg} is the average energy of alive nodes, and E_{std} is the standard deviation of residual energy of all alive nodes. This reduces overhead and energy consumed for transmitting control packets and reduces computation time in the network. The performance of NUWC is shown in Fig. 1. The flowchart of the proposed algorithm is shown in Fig. 2.

4.1 Setup phase

At the beginning of the first setup phase, the BS broadcasts request state messages in the environment. Each node calculates its distance from the BS considering RSSI. Then they transmit messages including location and energy information to the BS. BS performs 1-layering

Fig. 1 NUWC operation



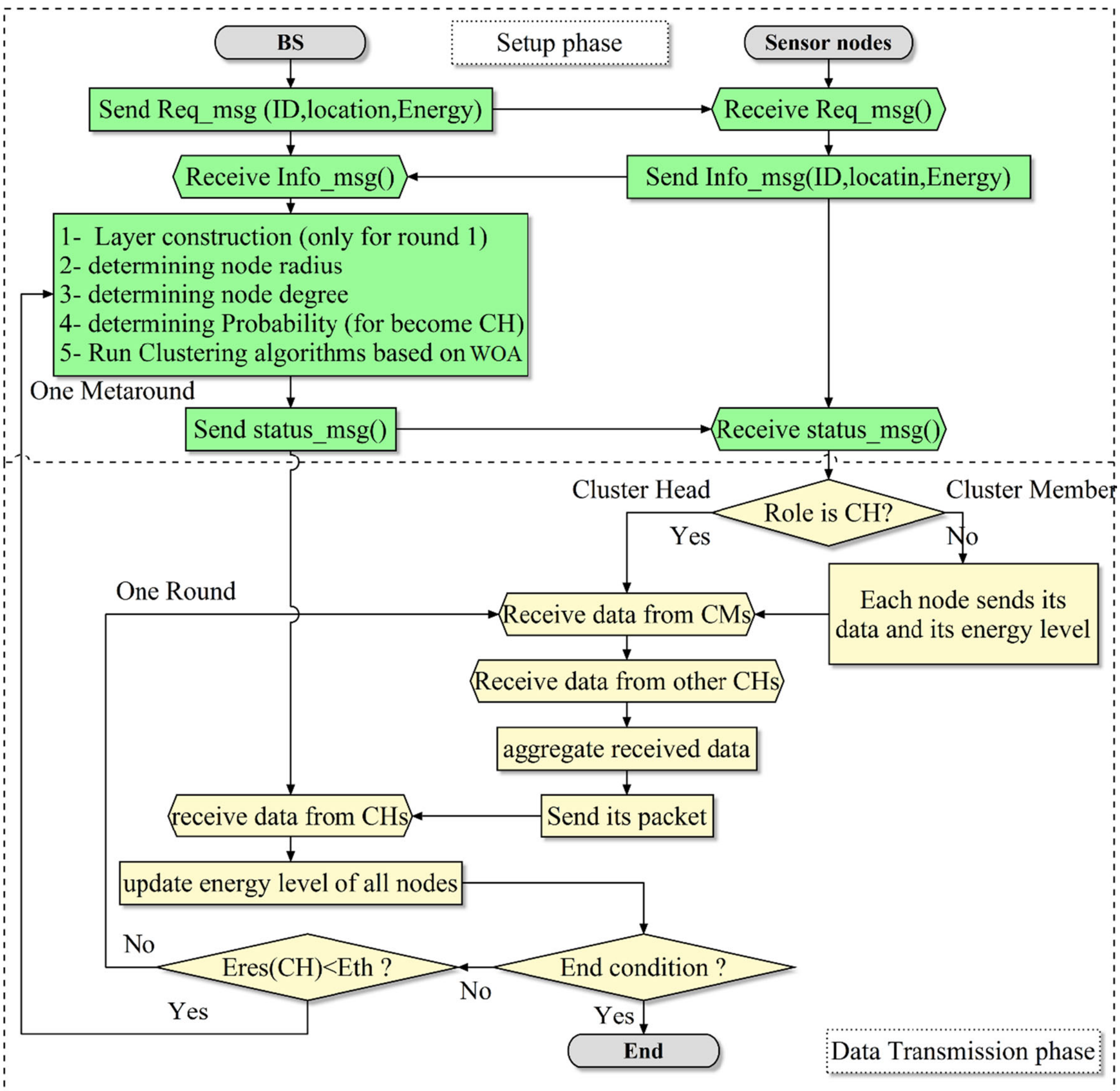


Fig. 2 Flowchart of NUWC protocol

2-determination of a node's radius 3-determination of a node's degree 4-probability calculation and 5-clustering based on WOA algorithm and then transmits a status message to nodes.

Nodes are layered based on their distance from BS. This layering is performed only once for the determination of the node radius coefficient. Layering is independent of clustering. Empirically, and based on papers like reference Malathi et al. (2015), BS defines nodes at 4 levels.

In the second stage, the BS obtained the radius of each node from Eq. (3):

$$R_c(j) = \left[1 - \alpha\left(\frac{d_{\max} - d_{j,BS}}{d_{\max} - d_{\min}}\right) - \beta\left(1 - \frac{E_{\mathrm{rem}}(j,r)}{E_{\mathrm{Max}}}\right)\right] Rl_{\max} \times \lambda l, \tag{3}$$

where $R_c(j)$ is the radius of node j. $Rl_{\max}$ is the maximum radius of the node, which is predetermined. $E_{\mathrm{rem}}(j,r)$ is the remained energy of node j at round r. E_{Max} is the maximum energy capacity of a node. α and β are weight factors that are between 0 and 1. Moreover, in higher layers, $Rl_{\max}$ is multiplied by a coefficient like λl so that nodes which are closer to the BS have a smaller radius and those that are further have a larger radius. For this purpose, λl in the first layer is 1; and in the second layer, it is 1.25 and in the third and fourth layers, they are equal to 1.75. Equation (3) helps having clusters with unequal sizes.

In the third stage, the BS determines the neighbor of each node. Neighbor of node i is the node whose distance from node i is shorter than its radius.

The next stage determines the probability of becoming CH. At first, the BS calculates the average energy of neighbor nodes using Eq. (4):

$$E_{\mathrm{Ave}N}(j,r) = \frac{\sum_{k \in ngh(j,r)} E_{\mathrm{rem}}(k,r)}{\max(|ngh(j,r)|, \varepsilon)}, \tag{4}$$

where $E_{\mathrm{Ave}N}$ (j, r) is the average energy of neighboring nodes of node j. $ngh(j,r)$ is the set of neighboring nodes of node j at round r. $|ngh(j,r)|$ is the number of neighbors of node j at round r. Nodes whose energy is less than the average energy of neighbors have no chance to become CH. Nodes with suitable energy level have the chance to become CH and this chance is calculated as Eq. (5):

$$P(j,r) = P_{\mathrm{Opt}} \times \frac{E_{\mathrm{rem}}(j,r)}{E_{\mathrm{Ave}N(j,r)}}, \tag{5}$$

where $P(j,r)$ is the probability of node j to become CH at round r. P_{Opt} is the optimal probability of becoming CH in the network which has been analyzed in reference to Heinzelman et al. (2002) and equal to 5% of the total number of nodes in the network. According to Eq. (5), the higher the energy level, the higher the chance to become CH.

4.1.1 Clustering based on the WOA algorithm

The last stage in the setup phase is clustering based on WOA by the BS. In the proposed algorithm based on WOA, the fitness function is presented to reduce energy consumption, balance energy consumption and improve network coverage.

4.1.1.1 Initial population The first generation of the population is a set of random solutions for the problem. The competency of each solution can be calculated by the fitness function. In the next stage, random solutions could be chosen and improved through specific functions. This procedure will continue until meeting stop criteria. In the framework of the WOA algorithm, a clustering solution is considered as an individual (I). In a WSN with N sensor nodes, non-CH nodes are considered as 0, CHs are considered as 1 and dead nodes are considered as -1. A population of M individual solution can be specified as Eq. (6):

$$\forall i \in \{1,\ldots,M\} \quad \text{and} \quad \forall j \in \{1,\ldots,N\}$$
$$I_i(j) = \begin{cases} 1 & \text{if} \quad (E_{\mathrm{rem}}(j,r) > 0 \quad \text{and} \quad E_{\mathrm{rem}}(j,r) \geq E_{\mathrm{Ave}N}(j,r) \quad \text{and} \quad rand \leq P(j,r)) \\ 0 & \text{elseif} \quad (E_{\mathrm{rem}}(j,r) > 0 \quad \text{and} \quad (E_{\mathrm{rem}}(j,r) < E_{\mathrm{Ave}N}(j,r) \quad \text{or} \quad rand > P(j,r))) \\ -1 & \text{otherwise} \end{cases} \tag{6}$$

Here, M is a number of individuals. N is a total number of sensor nodes. $I_i(j)$ is the status of node j at a solution i. $E_{\mathrm{rem}}(j.r)$ is the remained energy of node j at round r. $E_{\mathrm{Ave}N}(j.r)$ is the average energy of neighbors of node j at round r. This representation provides a dynamic number of CHs during a round. In solution i, nodes whose $I_i(j)$ value is 1 are selected as CH. Nodes whose $I_i(j)$ value is zero join the closest CH in coverage radius of that CH as CM, where the distance between the node to CH is less than the radius of the CH computed in Eq. (3). Thus, clusters with unequal sizes are established. Clusters with unequal sizes help in a way that CHs which are located close to BS have a smaller number of members. When a CH has a smaller number of members, it means that it has more energy to receive and relay data of CHs which are farther from the BS.

4.1.1.2 Fitness function For each individual, the fitness function is calculated using Eq. (7). The proposed objective function is defined as maximization. For NUWC, the proposed fitness function is used to minimize total energy consumption in the network (Eq. (8)) and balance nodes' consumption (Eq. (9)) and provide suitable coverage of network nodes by the CH (Eq. (10)):

$$F_{\mathrm{BOCA}}(I_i) = w_1 f_1(I_i) + w_2 f_2(I_i) + w_3 f_3(I_i) \tag{7}$$

$$f_1(I_i) = 1 - \frac{E_{\mathrm{cns}-intraC} + E_{\mathrm{cns}-interC} + E_{\mathrm{cns}-unCv} + E_{\mathrm{cns}-ctrlmsg}}{E_{t-\mathrm{Max}}} \tag{8}$$

$$f_2(I_i) = 1 - \left(\frac{1}{N_{\mathrm{alive}}} \times \sum_{j=1}^{N_{\mathrm{alive}}} \left(\frac{E_{\mathrm{cns-avg}} - E_{\mathrm{cns}}(j)}{E_{\mathrm{Max}}(j)}\right)^2\right) \tag{9}$$

$$f_3(I_i) = 1 - \left(\frac{N_{unCv}}{N_{\mathrm{alive}}}\right) \tag{10}$$

In Eq. (7), w_1, w_2 and w_3 are weight factors and $w_1 + w_2 + w_3 = 1$ is satisfied. Interval of Eqs. (8), (9) and (10) is between [0,1].

4.1.1.3 Minimize energy consumption (f_1) Equation (8) is proposed to reduce energy consumption. In Eq. (8), $E_{t\text{–Max}}$ is the sum of the total battery capacity of alive nodes which is obtained using Eq. (11). In Eq. (8), $E_{\mathrm{cns}-intraC}$ is intra-cluster energy consumption and $E_{\mathrm{cns}-interC}$ is inter-cluster energy consumption based on inter-cluster multi-hop routing which is calculated using Eqs. (12) and (13). In addition, $E_{\mathrm{cns}-unCv}$ is the energy of nodes that have not been covered by the CH and should transmit their data to the BS directly. It is calculated using Eq. (14). $E_{\mathrm{cns}-ctrlmsg}$ is the total energy consumed by control messages.

$$E_{t-\mathrm{Max}} = \sum_{j=1}^{N_{\mathrm{alive}}} E_{\mathrm{Max}}(j) \tag{11}$$

$$E_{\mathrm{cns}-intraC} = \sum_{c=1}^{nc} \sum_{j \in c_l} E_{TX}(j, k, d_{j,c}) + E_{RX}(c, k) + E_{DA}(c, k) \tag{12}$$

$$E_{\mathrm{cns-interC}} = \sum_{c=1}^{nc} \begin{cases} E_{TX}(c, k, d_{c,a}) + E_{RX}(a, k) + E_{DA}(a, k) & \text{if}(\text{relay}(c) = a) \\ E_{TX}(c, k, d_{c,BS}) & \text{otherwise} \end{cases} \tag{13}$$

$$E_{\mathrm{cns}-unCv} = \sum_{j \notin c_l} E_{TX}(j, k, d_{j,BS}) \tag{14}$$

In Eq. (11), $E_{\mathrm{Max}}(j)$ is the maximum capacity of node j. N_{alive} is the number of alive nodes. In Eq. (12), nc is the number of clusters. $E_{TX}(j,k,d_{j,c})$ is the energy consumed by the node j which is CM and transmit data (k bit) to their own CH ($d_{j,c}$ is the distance between node j and CH_c). $E_{RX}(c,k)$ is the energy consumed at CH c for receiving data from CM or other CH in inter-cluster multi-hop routing. $E_{DA}(c,k)$ is the energy consumed for aggregating k bit data by CH c. In Eq. (13), $E_{TX}(c,k,d_{c,BS})$ is energy consumed by CH c which transmits data to the BS directly. $E_{TX}(c,k,d_{c.a})$ is the energy consumed for transmitting data in multi-hop format from CH c to CH a. In Eq. (14), $E_{TX}(j,k,d_{j,\ BS})$ is the energy consumed for node j which transmit data to BS directly (nodes which are not covered by the CH).

4.1.1.4 Balance energy consumption (f_2) In Eq. (9), f_2 has been proposed to balance energy consumption and load balancing. This measure helps to select a state of the network in which nodes consume energy simultaneously in a balanced manner which increases stability. In this equation, $E_{\mathrm{cns\text{–}avg}}$ is the average energy consumed in the network which is calculated using Eq. (15). In Eq. (9), $E_{\mathrm{cns}}(j)$ is the energy consumed by node j and $E_{\mathrm{Max}}(j)$ is the maximum capacity of node j.

$$E_{\mathrm{cns-avg}} = \frac{E_{\mathrm{cns}-intraC} + E_{\mathrm{cns}-interC} + E_{\mathrm{cns}-unCv} + E_{\mathrm{cns}-ctrlmsg}}{N_{\mathrm{alive}}} \tag{15}$$

4.1.1.5 Coverage of network (f_3) As discussed before, a normal node joins the closest CH; where the distance between a node and CH is less than the radius of the CH calculated using Eq. (4). Therefore, there might be a node in the network which has not been covered by the CH. Improper coverage of network nodes is adverse because the node should transmit its data to the BS directly. Therefore, the fitness function should be designed such that all network nodes are covered by the CHs. Equation (10) is proposed to cover all nodes of the network. N_{unCv} is the number of nodes that are not in the radius of any CH. In other words, it is the number of nodes that have not been covered. N_{alive} is the number of alive nodes.

4.1.1.6 Inter-cluster multi-hop routing In Eq. (13), selecting the next hop in routing for inter-cluster transmission at each individual is performed such that minimum energy is required for transmission and transmission quality is also guaranteed. BS at each individual calculates the cost of transmission to BS according to Eq. (16):

$$\mathrm{relayCost}(j) = E_{TX}(j, k, d_{j,BS}) \tag{16}$$

In this equation, $E_{TX}(j,k,d_{j,\ BS})$ is the energy required for transmitting k data bits from CH j to BS which its distance is d. Then, BS calculates the cost of transmitting data to the middle CH and next hop. The next hop is selected with minimum cost for multi-hop transmission such that conditions of Eq. (17) are satisfied. Evaluation parameter for transmission to BS or middle CH is calculated using Eq. (17):

After clustering, the BS transmits information regarding the status of each node of the network considering its role. In this section, a new control message called status message is used. This message is divided into CH status message and CM status message. Considering the role of each node in the cluster, the BS transmits a control message associated with its role for that node. CM status message includes information regarding the ID of the node, status of the node (whether it is a CH or not), ID and distance from CH, and its turn to transmit node based on TDMA in intra-cluster transmission phase. Figure 3a shows the structure of this control message. Moreover, CH status message includes information regarding node's ID, node status, ID and distance of next hop of this CH for inter-cluster transmission, number of nodes which are CMs and number of CHs which transmit their data for this CH, Fig. 3b shows the structure of this control message.

$$\mathrm{relayCost}(j) = \begin{cases} E_{TX}(j,k,d_{j,l}) \;\; + E_{TX}(l,k,d_{l,\mathrm{nexthop}_l}) + E_{RX}(l,k) & \text{if } (E_{rem}(j,r) \geq E_{TX}(j,k,d_{j,l}) \text{ and } E_{\mathrm{rem}}(l,r) \geq E_{TX}(l,k,d_{l,\mathrm{nexthop}_l}) \;\; + E_{RX}(l,k)) \\ \mathrm{Inf} & \text{Otherwise} \end{cases} \tag{17}$$

where $E_{TX}(j,k,d_{j,l})$ is the energy required for transmitting data from CH j to CH l and $E_{TX}(l,k,d_{l,\mathrm{nexthop}l})$ is the energy required for transmitting data from CH l to the next hop. Thus, for transmitting data of CH j, BS considers two subsequent hops. $E_{RX}(l,k)$ is the energy required for receiving data by CH l. The BS selects CH of next hop such that it has enough energy for supporting reception of relay packet considering the number of CMs, data receiving cost and distance. After selecting CH with minimum relay cost, the next hop for CH j is selected. In addition, closer CHs will also transmit directly to the BS.

4.1.1.7 Best solution Each single round is composed of generating new solutions and evaluating fitness value. Until meeting quit criteria, this cycle will continue. Finally, best person, *best_I* among k individuals is considered as a clustering solution.

4.2 Data transmission phase

Data transmission includes some rounds and each round has two intra-cluster and inter-cluster transmission phases.

4.2.1 Round

At the first phase of each round, according to time division multiple accesses (TDMA) scheduling performed by the BS, nodes transmit their data and their amount of energy to the CH. When a CH receives a packet from a node which is a CM, integrates and aggregates it. After transmitting information of nodes to CH, the CHs transmit their packets to the BS in multi-hop format. Routing is formulated such that energy consumed for transmitting packets is

ID	status	CH ID	distance to CH	turn of send data

a. status message structure for CMs

ID	status	Nexthop ID	distance to Nexthop	num.of.CMs to receive data	num.of.CHs to receive data

b. status message structure for CHs

Fig. 3 Status messages structure: **a** CM status message and **b** CH status message

Table 3 Simulation scenarios

Scenario	Network area	BS location	Number of nodes	Initial Energy
Scenario#1	200 × 200	(100, 250)	100	0.2–0.5 J
Scenario#2	100 × 100	(50, 50)	100	0.2 J
Scenario#3	400 × 400	(200, 400)	200	0.2–0.5 J
Scenario#4	300 × 300	(300, 300)	150	0.2 J

Table 4 Parameters used for simulation

Parameters	The amount
E_{elec}	50 nJ/bit
E_{fs}	10 pJ/bit/m^2
E_{mp}	0.0013 pJ/bit/m^4
E_{DA}	5 nJ/bit/message
Data packet size	500 bytes
Packet header size	25 bytes
Control message size	25 bytes
Status message size	25 bytes
$Rl_{\max}$	110 m
P_{Mutation}	0.04
Population size	20
Max iteration	20
α, β	0.333
W_1, W_2, W_3	0.333

minimized. At this phase, the carrier sense multiple access (CSMA) method is used for transmitting data. Each CH transmits data to the BS according to the route constructed by the BS.

5 Simulation results and evaluation

In this study, MATLAB 2014 on an Intel Core i7 processor, 2.60 GHz CPU and 16 GB RAM on Microsoft Windows 8.1 is used to evaluate the proposed protocol. In addition, the proposed NUWC is compared with HUCL (Malathi et al. 2015), EADUC-II (Gupta and Pandey 2016) and SAERP (Khalil and Attea 2013) for evaluation. Simulations performed for four different scenarios can be seen in Table 3. The number of alive nodes and average network energy during simulation, first node death (FND), half node death (HND), last node death (LND) and efficiency of the network are evaluated. Simulations are performed 50 times and average results are considered.

5.1 Simulation parameters

Simulation parameters are set as in Table 4. Moreover, some important parameters are also determined through several simulations. It should be noted that evolutionary optimization parameters like the size of the population, number of generations and value of operators like mutation are the same as SAERP method. $Rl_{\max}$ is the same in all methods and equal to 110 m in the first, second and third scenarios. Another point that should be noted is that the size of packets is 500 bytes. Considering the performance of the proposed protocol, the size of the packets in the proposed protocol is 525 bytes.

In order to determine parameters $w1$, $w2$ and $w3$ in Eq. (7) (fitness function), it should be important, because it includes the orientation of the algorithm. The objective of f_1 is to reduce energy consumption and increase network lifetime. The objective of f_2 is to improve load balancing and increase stability. The objective of f_3 is to cover the network, properly. However, some objectives oppose each other. In other words, improving one causes degrading another objective. For instance, although improving load balancing increases FND and stability, but LND and network lifetime are reduced (Khalil and Attea 2011). Parameters can be adjusted depending on the application of the network. This paper tries to find a proper balance among different objectives like increasing lifetime, stability and proper coverage of the network. Therefore, the values of these parameters for $w1$, $w2$, and $w3$ are considered to be 0.333.

Additionally, in the proposed method, instead of sending conventional control messages by the nodes, new control messages are used to form clusters.

As mentioned, clustering is unequal. Figure 4 shows the routing and clustering graph generated by NUWC in one of the simulations for scenarios 1, 2, 3 and 4. Green nodes indicate CM and blue nodes are CH. As seen in these figures, clusters close to the BS are smaller than others.

5.2 Stability and lifetime of the network

Tables 5, 6, 7 and 8 show stability and the lifetime of the network including FND, HND, and LND of the protocols for the first, second, third and fourth scenarios. The proposed protocol shows better performance compared to EADUC-II, HUCL and SAERP in four scenarios. In the first scenario, the BS is out of the network and network dimensions are larger compared to the second scenario. Centralized protocols in large dimensions perform weaker

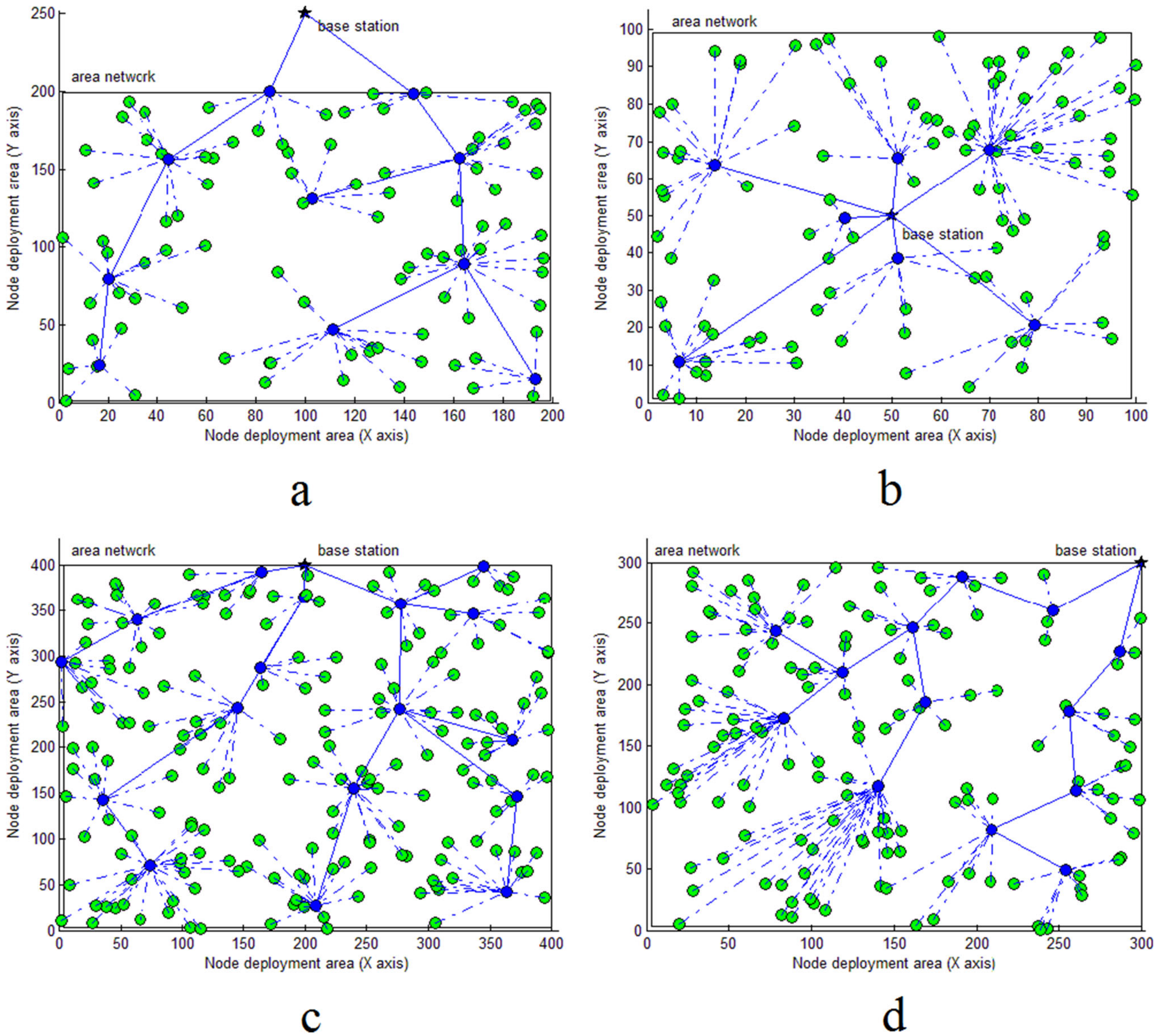


Fig. 4 Sample of clustering graph formed by NUWC. **a** scenario#1, **b** scenario#2, **c** scenario#3 and **d** scenario#4

Table 5 Simulation results of scenario1

Protocol	FND (100 Nodes)		HND (50 Nodes)		LND (0 Nodes)	
	Time	Packets	Time	Packets	Time	Packets
HUCL	130.3	12936	469.7	42093	557.1	44231
EADUC-II	230.3	22936	460.0	44167	483.3	44792
SAERP	35.2	3426	379.9	29834	477.6	32574
NUWC	514	51308	600.3	58596	696.8	60311

Table 6 Simulation results of scenario2

Protocol	FND (100 Nodes)		HND (50 Nodes)		LND (0 Nodes)	
	Time	Packets	Time	Packets	Time	Packets
HUCL	158.2	15720	311	29497	387.5	31383
EADUC-II	227.2	22625	280.9	27182	298.6	27613
SAERP	321.3	32030	332.8	32998	385	33243
NUWC	386.1	38515	426.7	41914	567	42612

compared to distributed protocols (Zanjireh and Larijani 2015). Thus, SAERP which is also based on single-hop transmission performs weaker compared to two distributed protocols, multi-hop HUCL and EADUC-II. In the second scenario, BS is at the center of the network and the dimensions of the network are smaller than the first

Table 7 Simulation results of scenario3

Protocol	FND (200 Nodes)		HND (100 Nodes)		LND (0 Nodes)	
	Time	Packets	Time	Packets	Time	Packets
HUCL	23.4	4480	350.4	55828	728	70030
EADUC-II	50.2	9762	363.8	65107	473	70913
SAERP	3.2	440	89	11706	557	34084
NUWC	256.2	51040	428	78281	891.6	90025

Table 8 Simulation results of scenario4

Protocol	FND (150 Nodes)		HND (75 Nodes)		LND (0 Nodes)	
	Time	Packets	Time	Packets	Time	Packets
HUCL	23.1	3459	237.6	29900	401.4	33551
EADUC-II	24.8	3477	221.8	30467	266.5	31850
SAERP	3.7	408	58.5	6124.9	309.9	14218
NUWC	36.4	5019	240.7	31553	486.2	36402

scenario. The maximum transmission distance from one node to the BS is about 70 m. That is, the distance between nodes is shorter, thus less energy is consumed for reinforcing data transmission. In such conditions, SAERP using a clustering method based on an evolutionary algorithm performs better. In scenario 2, multi-hop routing is not applicable. However, the method proposed in this environment has obtained better performance without using multi-hop routing. In the third scenario, the condition is a bit different. Network dimensions have increased. SAERP protocol is not efficient because it is single-hop. NUWC has preserved its stability well and performs better than HUCL and EADUC which are multi-hop protocols. Our proposed method has obtained high stability in three different environments and its performance in the three scenarios is better. NUWC uses fitness function for clustering and multi-hop routing for inter-cluster transmission.

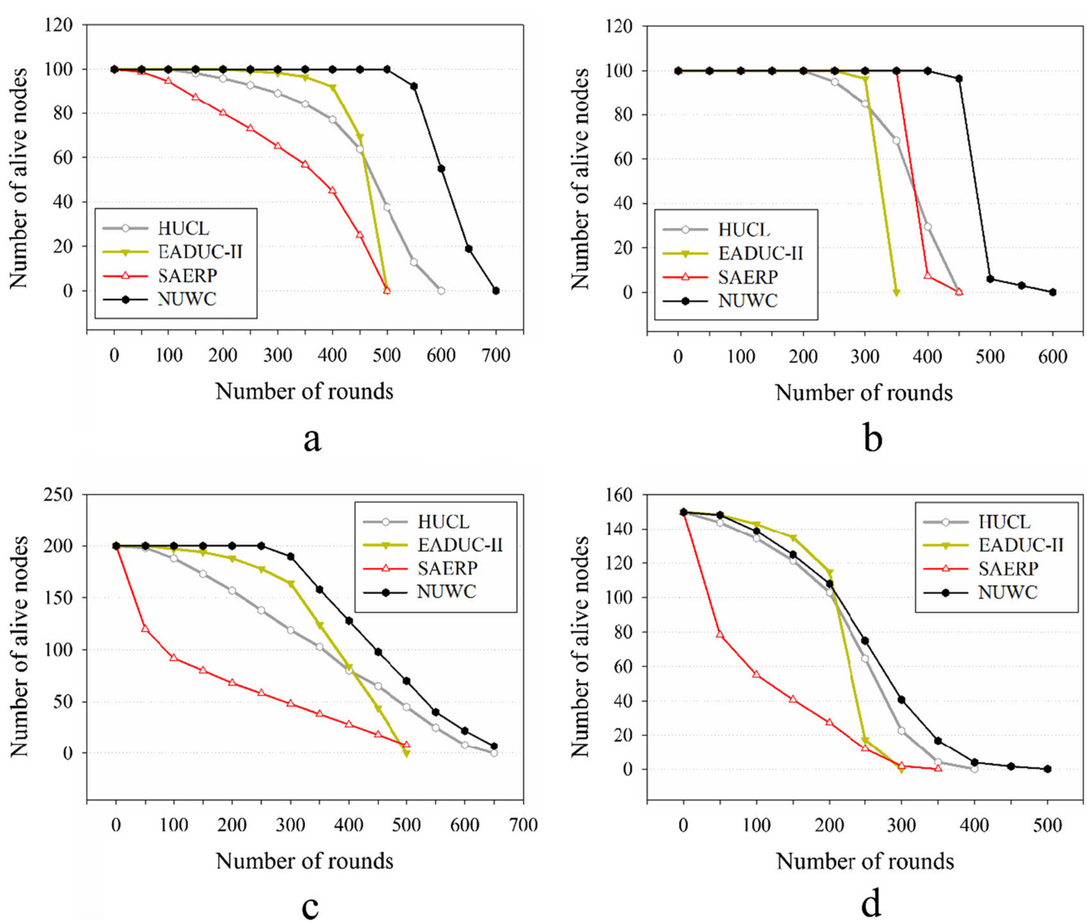


Fig. 5 Average number of available nodes in a round. **a** scenario#1, **b** scenario#2, **c** scenario#3 and **d** scenario#4

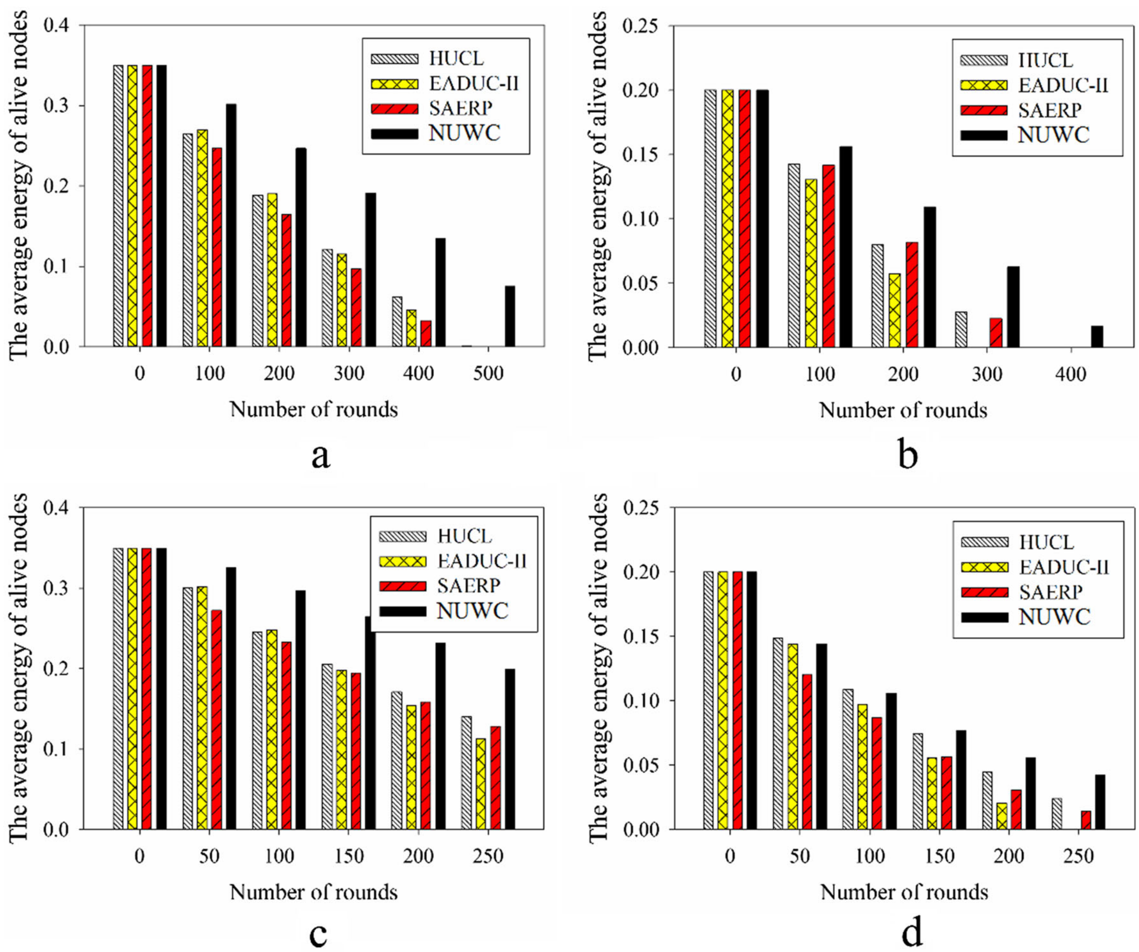


Fig. 6 The average energy of alive nodes in round. **a** scenario#1, **b** scenario#2, **c** scenario#3 and **d** scenario#4

Tables 5, 6, 7 and 8 show the obtained results, it can be seen that the proposed protocol contains delayed events and transmits more packets between occurrences of these events. This increase instability is due to reducing energy consumption and load balancing by calculating the variance of energy consumption. Since CH consumes more energy, focusing on this idea results in a balanced distribution of energy among nodes during the simulation. Moreover, the proposed protocol improves the lifetime of the network by eliminating unnecessary control messages and reducing overhead which results in energy consumption reduction.

5.3 Number of alive nodes

Figure 5 shows the number of alive nodes during the simulation. Results show that the performance of NUWC is better than other protocols and it could increase the number of alive nodes, because the proposed protocol reduces the energy consumption of nodes and improves network coverage, thus it prevents the death of nodes. What should be considered about the number of alive nodes during simulation is that in the first and third scenarios, the initial energy of the nodes is between 0.2 to 0.5 J and in the second and fourth scenario, the initial energy of the nodes is 0.2 J.

5.4 Energy of nodes

Figure 6 shows the energy of alive nodes during the simulation. In the proposed protocol, energy is divided among nodes by determining CH. In addition, using multi-hop transmission has lead to energy consumption reduction. In this type of transmission, since transmission distance is shortened, signals are less attenuated, thus less energy is required for transmitting them. The proposed protocol prevents a sudden reduction in energy by proper selection of CH and preventing direct transmissions for long distances. This strategy balances energy consumption among near and far nodes which increases stability and improves efficiency. On the other hand, the proposed protocol selects a new CH only when current CHs do not have proper

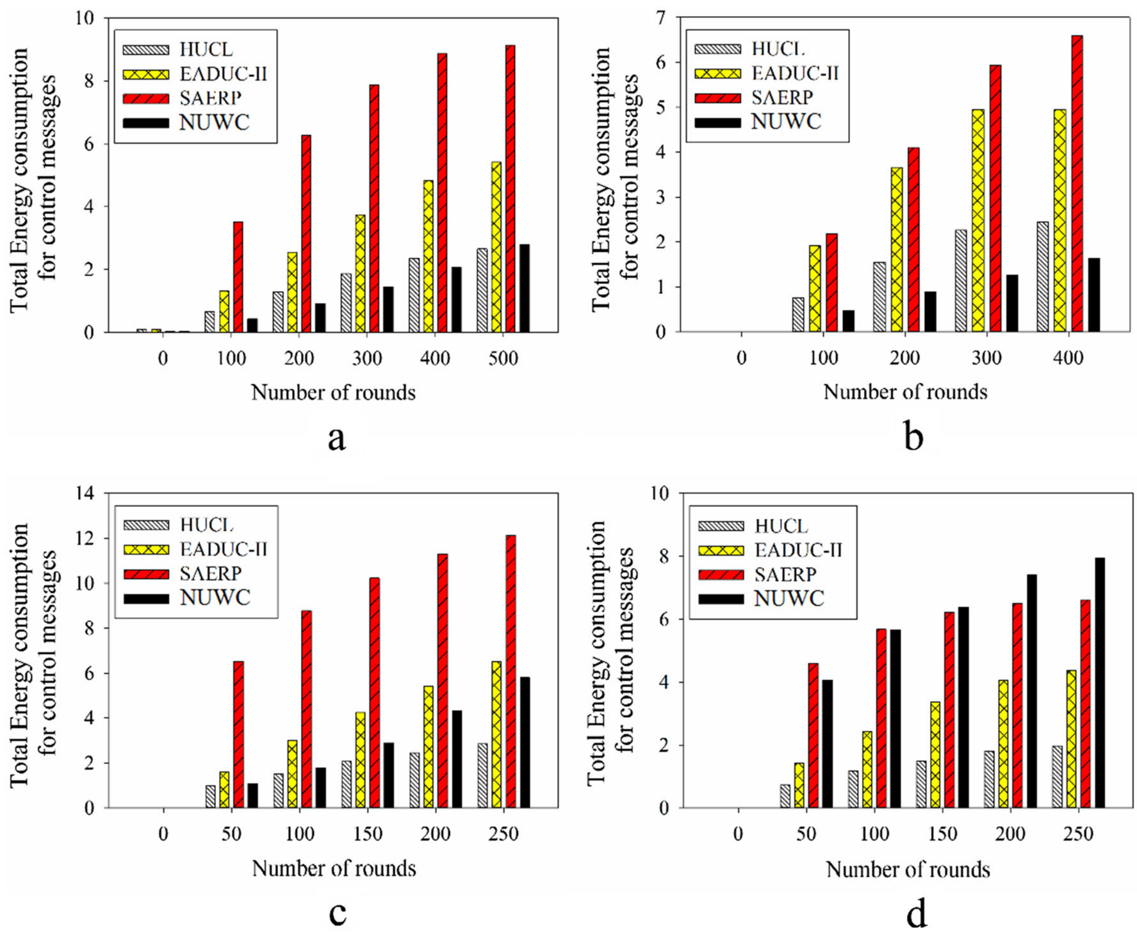


Fig. 7 Total energy consumption for control message in round. **a** scenario#1, **b** scenario#2, **c** scenario#3 and **d** scenario#4

energy; thus, reducing overhead improves the average energy of the nodes.

Figure 7 shows the total energy consumed for control messages. As it is seen, the proposed protocol has consumed less energy for control messages so that the lifetime of the network is increased.

The proposed protocol reduces overhead by reducing control messages. Eliminating conventional control messages and replacing them with status messages reduces the energy consumption of nodes for transmitting control messages. However, this is not enough, thus in the proposed protocol, CHs only change when a CH does not have suitable energy. In the second and fourth scenario, since nodes are homogeneous, their energy has a smaller variance and energy consumption of CHs is higher, the considered threshold for performing setup phase is satisfied. In the third scenario, the proposed protocol has reduced energy consumption through clustering and energy-aware routing, but it has not reduced overhead of control messages compared to hybrid methods like HUCL and EADUC. The most setup phase is performed in the final rounds. Since dimensions of the network are higher and the energy of the nodes is reduced gradually, the considered condition for performing setup phase is satisfied.

Figure 8, 9, 10 and 11 show the energy of nodes during simulation of first, second, third and fourth scenarios. In the first and third scenarios, nodes are heterogeneous. Thus, the energy difference among nodes is obvious. In the second and fourth scenario, nodes are homogeneous.

The proposed protocol has improved load balancing. Although there is a relative uniformity among the energy of most nodes, there are nodes with different energy compared to most nodes. The main reason for this issue is determining weight in Eq. (7) and fitness function. As the weight of f_2 which is associated to load balancing is increased, the load balancing of the protocol would be stronger which increases energy consumption and reduces network coverage. Since the purpose of this study is to balance FND and LND increase, associated weights are considered the same.

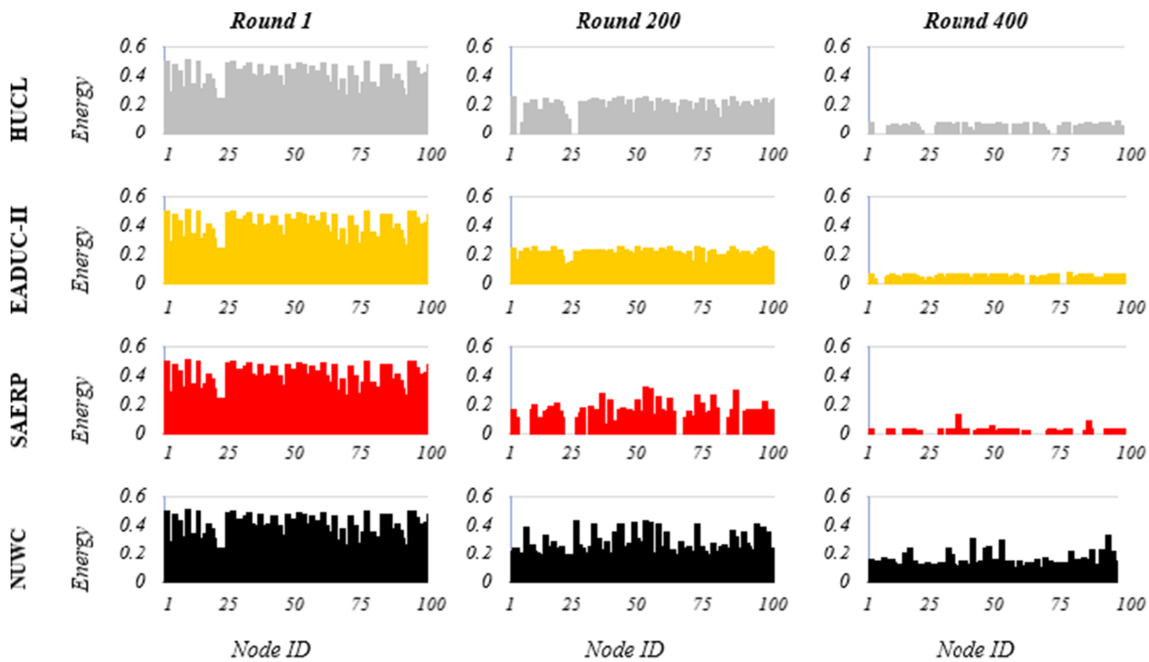


Fig. 8 Load balancing for scenario1

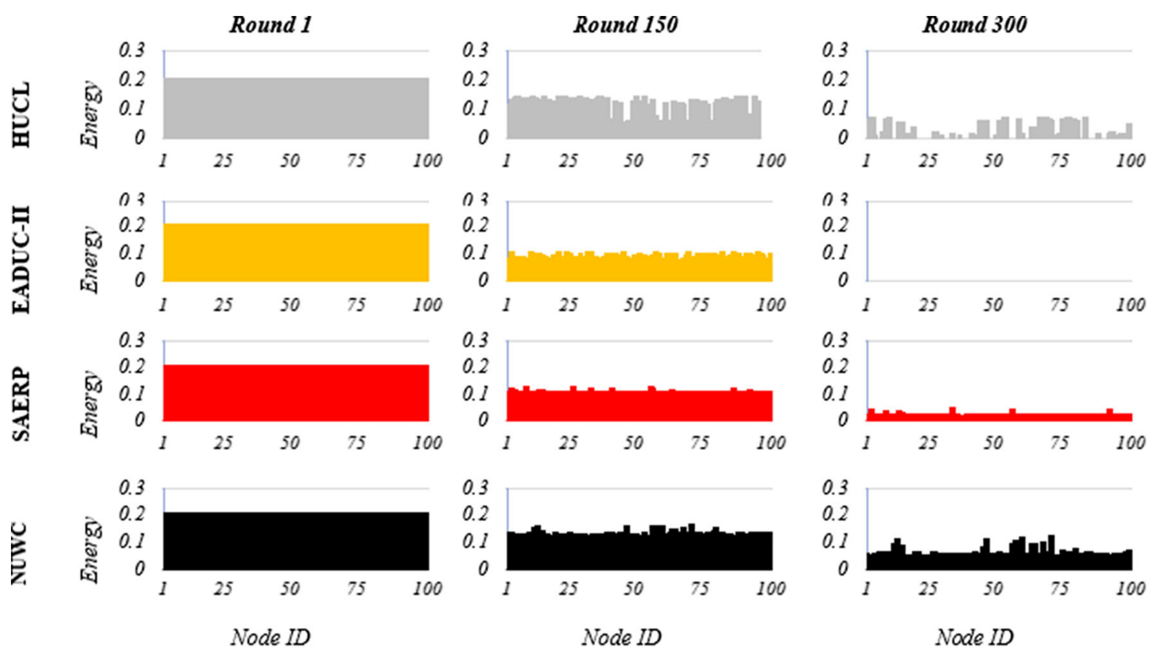


Fig. 9 Load balancing for scenario2

5.5 Throughput

Figure 12 the show number of generated packets per round which indicates the efficiency of the proposed protocol. This method has generated and transmitted more packets during the simulation. This efficiency is due to energy balancing which increases the stability and improves the number of accessible nodes.

Lost packets are the imperfect ones that have not been received at the destination due to various reasons. Calculating lost packets is important for measuring the efficiency of a parameter. Tables 9, 10, 11 and 12 show average generated packets, average lost packets and percentage of lost packets at different times. According to Table 8, in the first scenario, the proposed protocol performs better than other protocols. When the lifetime of other protocols is

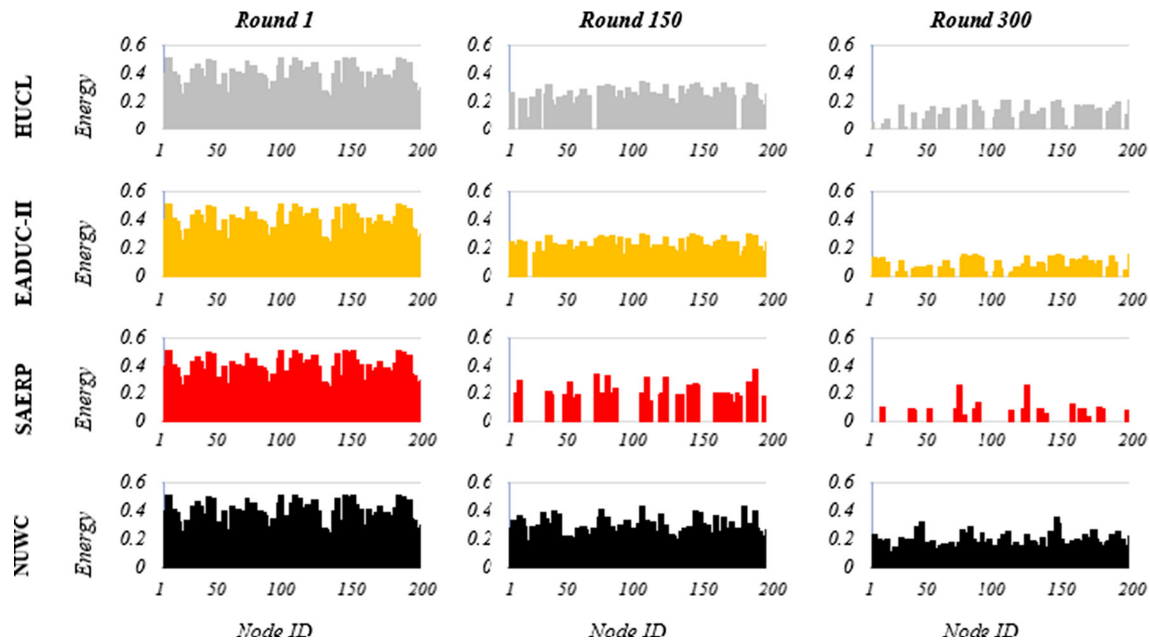


Fig. 10 Load balancing for scenario3

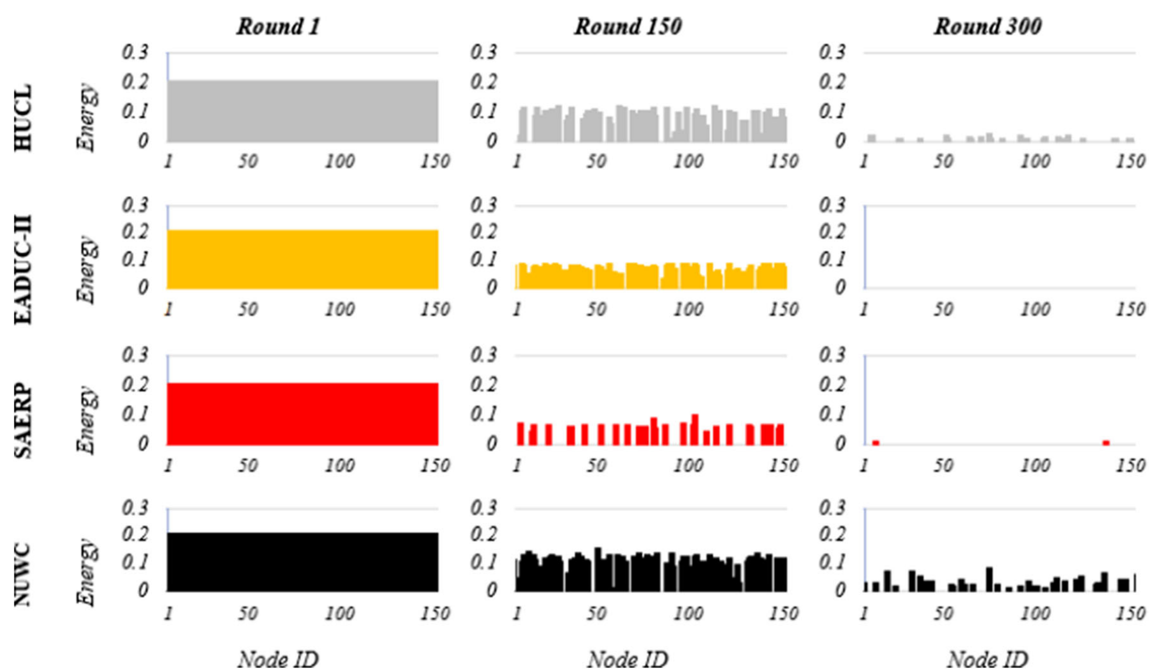


Fig. 11 Load balancing for scenario4

finished, the proposed protocol continues operation. In the second scenario, the proposed protocol has better efficiency and a minimum number of lost packets at different simulation times. In the third scenario, the proposed protocol has performed better. In addition to increasing network lifetime, this protocol has increased data transmission reliability also.

5.6 Average running time

Table 13 represents the average execution time of the protocol at each round for all scenarios. Since the NUWC is an energy-oriented algorithm, the time complexity of it has relatively decreased but, the computational complexity of cluster formation in the algorithm has increased comparatively. According to Table 13, the execution time of

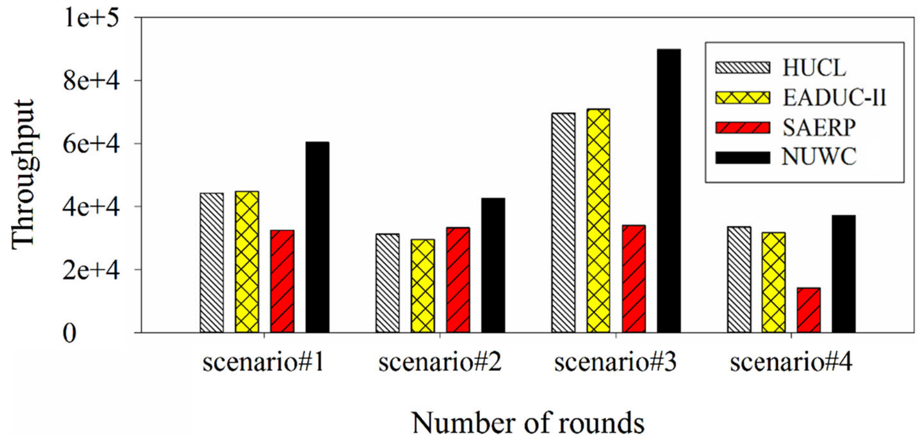


Fig. 12 A number of produced packages in a round. **a** scenario#1, **b** scenario#2, **c** scenario#3 and **d** scenario#4

Table 9 Average packet loss and average throughput for scenario1 (0.0% means percentage of lost packets is negligible)

Time	HUCL			EADUC-II			SAERP			NUWC		
	Made	Loss	%	Made	Loss	%	Made	Loss	%	Made	Loss	%
1	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%
100	100	0.02	0.02%	100	0.0	0.0%	94.6	0.16	0.17%	100	0.0	0.0%
200	95.9	0.04	0.04%	100	0.0	0.0%	78.2	0.1	0.13%	100	0.0	0.0%
300	89.2	0.06	0.07%	98.4	0.02	0.02%	63.3	0.1	0.14%	100	0.0	0.0%
400	77.1	1.88	2.81%	92.1	0.42	0.47%	47.8	0.34	1.01%	100	0.0	0.0%
500	37.6	4.62	12.2%	0	–	–	0	–	–	100	0.0	0.0%
600	0	–	–	0	–	–	0	–	–	50.2	0.9	1.7%
700	0	–	–	0	–	–	0	–	–	0	–	–

Table 10 Average packet loss and average throughput for scenario2 (0.0% means the percentage of lost packets is negligible)

Time	HUCL			EADUC-II			SAERP			NUWC		
	Made	Loss	%	Made	Loss	%	Made	Loss	%	Made	Loss	%
1	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%
100	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%
150	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%
200	97.3	0.1	0.1%	100	0.0	0.0%	100	0.0	0.0%	100	0.0	0.0%
250	92.2	2.5	2.7%	94.3	2.7	2.8%	100	0.0	0.0%	100	0.0	0.0%
300	59.3	3.7	6.2%	0	–	–	100	0.0	0.0%	100	0.0	0.0%
350	29.5	11.6	39.3	0	–	–	7.3	1.3	17.8%	100	0.0	0.0%
400	0	–	–	0	–	–	0	–	–	97.5	0.5	0.5%
500	0	–	–	0	–	–	0	–	–	3.1	0.4	12.9%
600	0	–	–	0	–	–	0	–	–	0	–	–

the NUWC is faster than SAERP as seen from scenarios 1 and 2. However, in scenarios 3 and 4, the computation time of SAERP has decreased compared to NUWC, significantly; because, at the beginning of simulating SAERP, the number of alive nodes decreases significantly and network dimensions is bigger as shown in Fig. 5c, d. It should be considered that the computational complexity of NUWC is directly proportional to a number of setup phases. The proposed protocol deploys the WOA algorithm for clustering only when the energy of the current CH is not suitable, that is why computational complexity has decreased. For evaluation that is more accurate, Table 14 shows that the average computational cost of NUWC, when it uses WOA algorithm, is 1.5 times that of SAERP when it uses EA with the same number of nodes. This computational complexity is due to more computations of fitness function (Eq. (7)). Although the computation complexity of cluster formation in the proposed algorithm has increased comparatively, but it can be executed in real time without any drawback.

Table 11 Average packet loss and average throughput for scenario3 (0.0% means the percentage of lost packets is negligible)

Time	HUCL			EADUC-II			SAERP			NUWC		
	Made	Loss	%	Made	Loss	%	Made	Loss	%	Made	Loss	%
1	200	0.0	0.0%	200	0.0	0.0%	200	0.0	0.0%	**200**	**0.0**	**0.0%**
50	198.8	0.7	0.3%	200	0.1	0.05%	123.8	1.6	1.2%	**200**	**0.0**	**0.0%**
100	188.2	1.6	0.8%	197	0.2	0.1%	93.2	1.4	1.5%	**200**	**0.0**	**0.0%**
200	143.8	2	1.3%	188.8	0.6	0.3%	71.4	0.8	1.1%	**200**	**0.0**	**0.0%**
300	119.8	8.2	6.8%	164.6	4	2.4%	60.6	0.7	1.1%	**189.8**	**0.2**	**0.1%**
400	80.8	6.2	7.6%	78.1	2.6	3.3%	28.8	0.4	1.3%	**128.2**	**0.2**	**0.1%**
500	45.8	4	8.7%	0	–	–	8.2	0.1	1.2%	**71**	**0.2**	**0.2%**
600	9.8	2	20.4%	0	–	–	0	–	–	**25.2**	**0.1**	**0.3%**
700	0.4	0.3	75%	0	–	–	0	–	–	**7.8**	**0.01**	**0.1%**
800	0	–	–	0	–	–	0	–	–	**3**	**0.01**	**0.3%**

Table 12 Average packet loss and average throughput for scenario4 (0.0% means the percentage of lost packets is negligible)

Time	HUCL			EADUC-II			SAERP			NUWC		
	Made	Loss	%	Made	Loss	%	Made	Loss	%	Made	Loss	%
1	150	0.0	0.0%	150	0.0	0.0%	150	0.0	0.0%	150	0.0	0.0%
50	143.9	0.9	0.6%	148.3	0.08	0.05%	78.4	0.72	0.9%	148.9	0.08	0.05%
100	134.4	1.86	1.38%	143.1	0.12	0.08%	55.1	0.22	0.39%	138.8	0.32	0.23%
200	102.9	5.54	5.38%	114.9	4.34	3.77%	27.2	0.18	0.66%	108.6	0.66	0.6%
300	22.6	2.98	4.42%	0	–	–	1.7	0.04	2.35%	40.6	0.38	0.9%
400	0.38	0.06	15.7%	0	–	–	0	–	–	3.8	0.06	1.5%
500	0	–	–	0	–	–	0	–	–	1.5	0.02	1.33%

Table 13 Average running time (microseconds) in each round

Scenario	HUCL	EADUC-II	SAERP	NUWC
scenario#1 (with 100 nodes)	5115	14,918	593,216	135,564
scenario#2 (with 100 nodes)	6607	20,127	623,119	463,001
scenario#3 (with 200 nodes)	9691	37,618	471,921	688,457
scenario#4 (with 150 nodes)	6327	24,915	375,524	562,916

Table 14 The average running time (microsecond) when EA running in SAERP and WOA in NUWC at the beginning rounds and with an equal number of nodes

Scenario	SAERP	NUWC
scenario#1 (with 100 nodes)	580,611	924,903
scenario#2 (with 100 nodes)	962,038	1,021,032
scenario#3 (with 200 nodes)	2,093,840	3,211,813
scenario#4 (with 150 nodes)	1,445,984	2,179,914

6 Conclusion

Using IoT technology which is implemented through wireless nodes is a new trend in the world. Limited energy is the most important challenge in these networks. Sending and receiving messages is the main reason for consuming energy. Hence, designing an efficient method for routing and managing messages could have a significant impact on the network's lifespan. In the proposed method, a concentrated clustering method, which has been implemented by WOA, is designed to tackle the problem. Inter-cluster interactions have been managed by means of multi-hop approach which has energy saving considerations. Saving more energy, load balancing and CH coverage at different points of the network are the main objectives in fitness function control messages are replaced by status messages in order to reduce messaging overhead. The performance of the algorithm has been analyzed in terms of energy consumption. Results suggest that NUWC managed to improve the level of stability and lifespan. Moreover, it can balance energy consumption and throughput of the network. In future works, it is intended to develop the proposed method for IoT over energy harvesting wireless nodes.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflicts of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Research involving human participants and/or animals This article does not contain any studies with human participants performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Abdul-Qawy ASH, Srinivasulu T (2018) SEES: a scalable and energy-efficient scheme for green IoT-based heterogeneous wireless nodes. J Ambient Intell Humaniz Comput 10:1571–1596

Afsar MM, Tayarani-N MH (2014) Clustering in sensor networks: a literature survey. J Netw Comput Appl 46:198–226

Azharuddin M, Jana PK (2016) Particle swarm optimization for maximizing lifetime of wireless sensor networks. Comput Electr Eng 51:26–42

Bozorgi SM, Amiri MG, Rostami AS, Mohanna F (2016) A novel dynamic multi-hop clustering protocol based on renewable energy for energy harvesting wireless sensor networks. In: Conference proceedings of 2015 2nd international conference on knowledge-based engineering and innovation, KBEI 2015

Bozorgi SM, Shokouhi Rostami A, Hosseinabadi AAR, Balas VE (2017) A new clustering protocol for energy harvesting-wireless sensor networks. Comput Electr Eng 64:233–247

Brest J, Zumer V, Maucec MS (2006) Self-adaptive differential evolution algorithm in constrained real-parameter optimization. IEEE Int Conf Evol Comput 2006:215–222

Chaurasiya SK, Pal T, Bit SD (2011) An enhanced energy-efficient protocol with static clustering for WSN. In: International conference on information networking 2011, ICOIN 2011, pp 58–63

Clerc M, Kennedy J (2002) The particle swarm—explosion, stability, and convergence in a\nmultidimensional complex space. IEEE Trans Evol Comput 6(1):58–73

Deb K (2000) An efficient constraint handling method for genetic algorithms. Comput Methods Appl Mech Eng 186:311–338

Eberhart R, Shi Y (2004) Guest editorial special issue on particle swarm optimization. IEEE Trans Evol Comput 8(3):201–203

Gupta GP, Jha S (2018) Integrated clustering and routing protocol for wireless sensor networks using Cuckoo and Harmony Search based metaheuristic techniques. Eng Appl Artif Intell 68:101–109

Gupta V, Pandey R (2016) An improved energy aware distributed unequal clustering protocol for heterogeneous wireless sensor networks. Eng Sci Technol Int J 19(2):1050–1058

Han T, Bozorgi SM, Orang AV, Hosseinabadi AR, Sangaiah AK, Chen MY (2019) A hybrid unequal clustering based on density with energy conservation in wireless nodes. Sustainability 11:1–26

Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun 1(4):660–670

Hosseinabadi AR, Slowik A, Sadeghilalimi M, Farokhzad M, Babazadeh M, Sangaiah AK (2019) An ameliorative hybrid meta-heuristic algorithm for solving the capacitated vehicle routing problem. IEEE Access 7:175454–175465

Khalil EA, Attea BA (2011) Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks. Swarm Evol Comput 1(4):195–203

Khalil EA, Attea BA (2013) Stable-aware evolutionary routing protocol for wireless sensor networks. Wirel Pers Commun 69:1799–1817

Kuila P, Jana PK (2014a) Energy efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach. Eng Appl Artif Intell 33:127–140

Kuila P, Jana PK (2014b) A novel differential evolution based clustering algorithm for wireless sensor networks. Appl Soft Comput J. 25:414–425

Kuila P, Gupta SK, Jana PK (2013) A novel evolutionary approach for load balanced clustering problem for wireless sensor networks. Swarm Evol Comput 12:48–56

Kumar V, Kumar S (2016) Energy balanced position-based routing for lifetime maximization of wireless sensor networks. Ad Hoc Netw 52:117–129

Kumar A, Kumar V, Narottam C (2011) Energy efficient clustering and cluster head rotation scheme for wireless sensor networks. Int J Adv Comput Sci Appl 3(5):129–136

Ma H, Simon D (2011) Blended biogeography-based optimization for constrained optimization. Eng Appl Artif Intell 24(3):517–525

Machado K, Rosário D, Cerqueira E, Loureiro A, Neto A, de Souza J (2013) A routing protocol based on energy and link quality for internet of things applications. Sensors 13(2):1942–1964

Malathi L, Gnanamurthy RK, Chandrasekaran K (2015) Energy efficient data collection through hybrid unequal clustering for wireless sensor networks. Comput Electr Eng 48:358–370

Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67

Rostami AS, Badkoobe M, Mohanna F, Hosseinabadi AR, Kardgar M, Balas VE (2016) Imperialist competition based clustering algorithm to improve the lifetime of wireless sensor network. In: 7th International workshop in soft computing applications (SOFA 2016), Springer, vol. 633. pp 189–202

Rostami AS, Badkoobe M, Mohanna F, Keshavarz H, Hosseinabadi AR, Kumar Sangaiah A (2018) Survey on clustering in heterogeneous and homogeneous wireless sensor networks. J Supercomput 74:277–323

Sabor N, Abo-Zahhad M, Sasaki S, Ahmed SM (2016) An unequal multi-hop balanced immune clustering protocol for wireless sensor networks. Appl Soft Comput J 43:372–389

Saemi B, Hosseinabadi AR, Kardgar M, Balas VE (2016) Nature inspired partitioning clustering algorithms: a review and analysis. In: 7th International workshop in soft computing applications (SOFA 2016), vol. 634. Springer, pp 96–116

Sangaiah AK, Sadeghilalimi M, Hosseinabadi AR, Zhang W (2019) Energy consumption in point-coverage wireless sensor networks via bat algorithm. IEEE Access 7:180258–180269

Shah SB, Chen Z, Yin F, Khan IU, Ahmad N (2018) Energy and interoperable aware routing for throughput optimization in clustered IoT-wireless sensor networks. Future Gener Comput Syst 81:372–381

Shankar T, Shanmugavel S, Rajesh A (2016) Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks. Swarm Evol Comput 30:1–10

Shokouhifar M, Jalali A (2015) A new evolutionary based application specific routing protocol for clustered wireless sensor networks. AEU Int J Electron Commun 69(1):432–441

Yu J, Qi Y, Wang G, Guo Q, Gu X (2011) An energy-aware distributed unequal clustering protocol for wireless sensor networks. Int J Distrib Sens Netw 7(1):2021–2045

Younis O, Member S, Fahmy S (2004) HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor. Networks 3(4):366–379

Zanjireh MM, Larijani H (2015) A survey on centralised and distributed clustering routing algorithms for WSNs. 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), pp 1–6