Wireless Netw (2012) 18:847–860
DOI 10.1007/s11276-012-0438-z

Cluster based wireless sensor network routing using artificial bee colony algorithm

Dervis Karaboga · Selcuk Okdem · Celal Ozturk

Published online: 24 April 2012

Abstract Due to recent advances in wireless communication technologies, there has been a rapid growth in wireless sensor networks research during the past few decades. Many novel architectures, protocols, algorithms, and applications have been proposed and implemented. The efficiency of these networks is highly dependent on routing protocols directly affecting the network life-time. Clustering is one of the most popular techniques preferred in routing operations. In this paper, a novel energy efficient clustering mechanism, based on artificial bee colony algorithm, is presented to prolong the network life-time. Artificial bee colony algorithm, simulating the intelligent foraging behavior of honey bee swarms, has been successfully used in clustering techniques. The performance of the proposed approach is compared with protocols based on LEACH and particle swarm optimization, which are studied in several routing applications. The results of the experiments show that the artificial bee colony algorithm based clustering can successfully be applied to WSN routing protocols.

Keywords Wireless sensor networks · Cluster based routing · Artificial bee colony algorithm

A Shorter version of this paper appeared in AIS 2010.

D. Karaboga · S. Okdem · C. Ozturk (✉)
Computer Engineering Department, Engineering Faculty, Erciyes University, 38039 Kayseri, Turkey
e-mail: celal@erciyes.edu.tr

D. Karaboga
e-mail: karaboga@erciyes.edu.tr

S. Okdem
e-mail: okdem@erciyes.edu.tr

1 Introduction

Wireless sensor networks (WSNs) consisting of a number of distributed sensor nodes, have received much attention in recent years [1–5]. They enable a wealth of distributive application areas such as military, environmental monitoring, healthcare services, chemical detection and seismic measurements [6]. In any of these applications, nodes are in sense of potential events which are occurring in their sensing range.

They try to cooperate with each-other to transmit the sensed data to an external base station. Due to their distributed nature, performing a collaborative organization for a robust communication is required.

The sensor nodes are easily configurable to many specific applications [7]. However, they are powered only by irreplaceable batteries with limited energy. Nevertheless, their processors have limited processing power and the communication channels used by the sensor nodes are usually in low bandwidths. By considering these constraints of sensor nodes, innovative techniques are highly required to enable reliable communications. Not only trustable communication is needed, but also the network life-time should be long as much as possible in the applications of WSNs.

Communication is the main factor of energy dissipation in sensor nodes as studied in [8]. The dissipation depends on the distance between communicating elements as source and destination. Electronics such as sensing units and probable Global Positioning System on sensor nodes have also importance on energy consumption [9]. Since GPS systems needs large size of hardware and require additional energy sources, using GPS technology on sensor nodes causes the system to be more expensive and consumer.

Having limited sources necessitate energy-aware routing protocols running on the sensor nodes. The protocols running on these networks should be simple-structured, fast-executable and need low power for successful applications. The basic method to transfer information from sensor nodes to the base station is called flooding in which all sensor nodes are communicating by broadcasting. This method consumes too much energy and bandwidth while disseminating data to all over the network as well as the base station. SPIN family protocols [10] disseminate all the information in the network assuming that all nodes are potential base nodes. However SPIN's data advertisement operation does not guarantee data delivery. Therefore, multi-paths are being used to make reliable communications.

Clustered routing mechanisms are particularly more suitable for wireless sensor networks with continued data flowing. A popular clustering protocol is called LEACH [11] which proposes a two-phase operation based on single-tier network using clusters. The protocol randomly creates cluster heads and performs data aggregation processes in the clusters. LEACH-C [12] is an extension to LEACH protocol to increase the energy efficiency. PEGASIS [13] is another cluster based protocol, which is similar to LEACH but that consumes less energy per round. LEFC [14] is also an effective clustering algorithm developed for WSNs, using local information of the nodes defining their belongings to clusters. Although LEACH-C, PEGASIS and LEFC present better performance over the algorithm of LEACH, the positions of sensor nodes must be known priori to execute these protocols. Since, sensor nodes are usually deployed randomly and hardware of global positioning systems increases the size of nodes and cost of the system, these protocols are not suitable for many WSNs applications.

In recent years, the routing protocols using swarm based optimization algorithms present an alternative way of efficient data gathering techniques on sensor network routings. The algorithms running on the nodes should be in simple forms providing fast execution and little energy consumption. Ant colony optimization (ACO), modeling ants behavior of finding food sources [15], is a candidate method for multi-path routing using a swarm based algorithm to maximize network life time in event based applications as proposed in [16]. However, this proposal is not suitable for monitoring applications which require periodic data transferring [17]. Particle swarm optimization (PSO) which simulates the social behavior of a fock of birds flying to resources [18] is used in LEACH based routing [19, 20]. In these studies, PSO algorithm is used to form clusters and to define cluster-heads in a WSN having the same conditions with simulated network and PSO algorithm produces promising results. However, the network life-time could be prolonged by using other optimization techniques that are successfully applied on clustering. Artificial bee colony (ABC) algorithm a new swarm based optimization technique [21] is applied on clustering of WSN used for monitoring purposes [22].

In this paper, we have adopted a centralized clustering approach for the selection of cluster heads. The selection process of our method is working on the base station and except this process; it is working in a distributed manner. Base station hardware is usually more sophisticated than other network elements such as cluster-heads or sensor nodes. Therefore, it would benefit to take the overhead of the routing process from sensor nodes whose main tasks are sensing the environment and reporting. Our clustering method uses ABC algorithm which is proposed by Karaboga in [21] and the clustering performance of the algorithm is firstly studied in pattern classification in [23]. In [23], a multivariable data clustering is achieved for a given set of classification benchmark data by means of ABC algorithm trying to achieve an optimum data grouping process. In proposed ABC based routing approach, it is aimed to provide a suitable routing scheme for WSN nodes considering the behavior of routing scheme as choosing optimum cluster-head nodes to minimize energy consumption. The remainder of the paper is organized as follows: Sect. 2 outlines the basics of ABC algorithm. Section 3 exhibits the details of cluster based routing and the proposed approach. Section 4 consists of simulation results and evaluates the performance of proposed approach. Finally, the last section concludes the paper by a summary and future works.

2 Artificial bee colony algorithm

Artificial bee colony algorithm, a swarm-based artificial intelligence algorithm, is inspired by intelligent foraging behavior of honey bees [24, 25]. In the ABC algorithm, there are three bee groups in artificial bee colony: onlookers, scouts, and employed bees where each bee represents a position in the search space. When the network consists n cluster-head sensors, the bees fly in the search space with n dimensions. The ABC employs a population of bees to find the cluster-heads. A bee waiting on the dance area to determine to choose a food source is an onlooker and a bee goes to the food source visited by it previously is an employed bee. A bee who carries out random search is called a scout. The position of a food source represents a possible solution to the optimization problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution.

The pseudo-code of the ABC algorithm is:

```
Generate initial population Xi, i = 1…SN
Evaluate the population
Set cycle to 1
Repeat
FOR each employed bee
    Produce new solutions vi by using (1)
    Calculate the fitness
    Apply the greedy selection process
FOR each onlooker bee
    Choose a solution xi depending on pi
    Produce new solutions vi
    Calculate the fitness
    Apply the greedy selection process
If there is an abandoned solution then
    Replace it with a new solution produced by a scout using (3).
Memorize the best solution achieved so far
cycle = cycle + 1
Until cycle = MCN
```

In the ABC algorithm, the first half of the colony consists of employed bees and the second half consists of onlooker bees. The first positions of food sources randomly generated where each employed bee is nominated to a food source. Then, every employed bee determines a new neighboring food source of its currently associated food source by Eq. (1), and computes the nectar amount of the new food source for each iteration. If the nectar amount of the new food source is higher than the previous one, then employed bee moves to the new food source, otherwise it continues with the old one.

$$v_{ij} = x_{ij} + \theta_{ij}\left(x_{ij} - x_{kj}\right) \quad (1)$$

where θ is a random number between [− 1,1], v_i is a candidate solution, x_i is the current solution and x_k is a neighbour solution and $j \in \{1,2,\ldots,D\}$ is randomly chosen index where D is the dimension of the solution vector.

The employed bees share the information about their food sources with onlooker bees after all of them complete the search process. An onlooker bee evaluates the nectar information taken from all employed bees and chooses a food source with a probability related to its nectar amount by Eq. (2), known as roulette wheel selection method which provides better candidates to have a greater chance of being selected.

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n} \quad (2)$$

where fit_i is the fitness value of the solution i proportional to the nectar amount of the food source in the position i and SN is the number of food sources equal to the number of employed bees. Once all onlookers have selected their food sources, each of them determines a new neighboring food source of its selected food source and computes it's nectar amount. The bee memorizes the new position and forgets the old one whether its nectar is higher than that of the previous one; otherwise it keeps that.

The employed bee becomes a scout when a food source is exhausted by the employed and onlooker bees. Any position cannot be improved further through a predetermined number of cycles which is called limit parameter, the food source is assigned as abandoned and employed bee of that source becomes scout. In that position, a new solution is randomly generated by the scout, given in Eq. (3), where abandoned source is represented by x_i and $j \in \{1,2,\ldots,D\}$. The flow chart of the ABC algorithm is given in Fig. 1.

$$x_i^j = x_{\min}^j + rand(0,1)\left(x_{\max}^j - x_{\min}^j\right) \quad (3)$$

3 WSN routing protocols using artificial bee colony algorithm

In the paper, the scenarios of WSNs routing protocols using ABC algorithm are developed for the networks having no global positioning system. The main purpose of the operations of these protocols is to increase the network life-time by maximizing the number of transferred data packages with clustering. The clustering mechanism of the proposed protocols is based on the clustering technique of LEACH protocol where cluster heads perform data aggregation processes of their clusters. Cluster heads use TDMA MAC in intra-cluster communication and CDMA MAC communication with the base station. The main operational difference between the proposed protocols and LEACH is the selection process of cluster heads (CH); clustering head selection is performed by ABC algorithm in proposed protocols while LEACH uses a random selection method. The proposed network clustering protocol is based on a centralized control algorithm that is implemented at the base station. The base station is a node with unlimited energy supply.

For a sensor network with N nodes and k number of clusters, the sensor network can be clustered as follows:

Step 1: Initialize population, each individual containing k randomly selected cluster heads.
Step 2: Evaluate the fitness function of each individual, where:

i. For each node n_i in the network:
   a. Calculate distance(n_i, CH_k) between n_i and all cluster heads CH_k

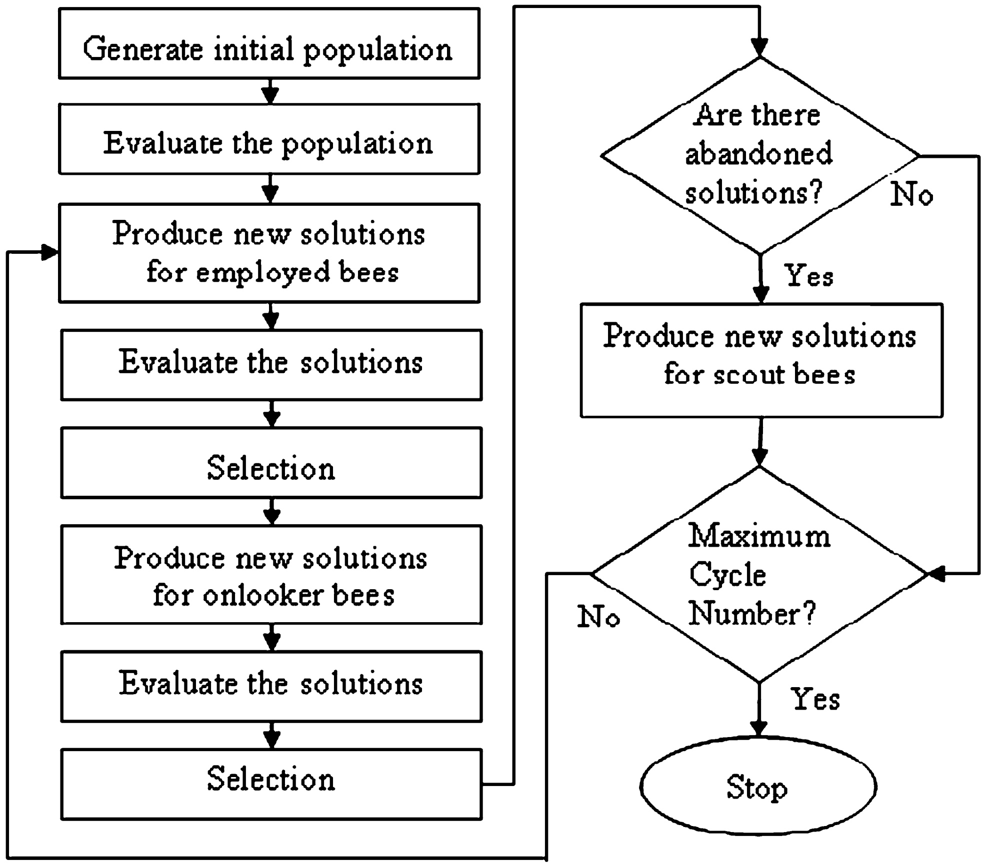

Fig. 1 Main steps of ABC algorithm

b. Assign n_i to CH_k, where distance(n_i, CH_k) = $\min_{\forall k=1,2,\ldots k}$ {distance (n_i, CH_k)}

ii. Compute the fitness function.

Step 3: Perform the position update by the optimization algorithm.

Step 4: Repeat steps 2 to 4 until the maximum number of cycle is reached.

3.1 Cluster based routing strategy

In a typical WSN application, sensor nodes collect data nearby and send it to the destination which is a neighbor node or the base station. In a clustered approach, cluster heads gather data about the common phenomena from sensor nodes and then aggregate this raw data to form the final abstract data. The main idea in the aggregation of data is to combine the data obtained from different sensor nodes in a neighborhood, eliminate redundancies by performing simple processing (such as using max. operator) and minimize the total amount of data transmission before transferring data to the base [26].

LEACH protocol, one of the most popular clustering methods used for WSNs, improves the network life time significantly. It assumes that every node has a capability of reaching to the base in a single hop. Nodes in the network elect themselves as cluster heads in a probabilistic way. The process of the election of cluster heads is carried out in every tour.

In time, nodes take turns of being cluster heads using Eq. (4). In the probabilistic method of LEACH, sensor nodes generate a random number distributed in [0,1]. If the random value is less then threshold *T(n)*, the sensor node advertises itself as a cluster head. *T(n)* is calculated using the values of parameters *P*, *r*, *n* and *G* which present desired percentage to become a cluster head, current round index, sensor node index and set of the nodes that have not being selected in the last *1/P* rounds, respectively.

$$T(n) = \begin{cases} \frac{P}{1-P\cdot[r \bmod (1/P)]} & \text{if } n \in G, \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The communication in LEACH algorithm is achieved in two hops. In the first hop, a member node of a cluster communicates with its cluster head directly. And in the second hop, the cluster head forward the aggregated data to the base. Sensor nodes carried out the election of cluster heads by themselves. Each sensor decides to be a cluster head or a member using a certain probability. The election process is carried out at every tour which means all the sensor nodes in the network transferred their data through their cluster heads to the base. After election process of cluster heads in a tour, cluster heads broadcast their status to other sensors in the network. Then, each sensor node determines to which cluster it wants to join by choosing the cluster head that requires the minimum communication energy. In [12], optimum number of clusters is analytically determined 5 for a LEACH system having 100 sensor nodes as well as used in [19, 20]. After the organization of clusters in the network, each cluster head creates a schedule to gather the data from sensor nodes in its cluster using TDMA MAC protocol. Then, cluster heads send the aggregated data to the base using CDMA codes to prevent the interference.

We consider a wireless sensor network model that consists of a base station and a number of stationary sensor nodes that are grouped into clusters dynamically in each tour as used in LEACH. The proposed approach uses a centralized control mechanism implemented at the base station at two stages (Fig. 2). At the first stage, initialization of the network is made when sensor nodes are deployed to the area. At this stage, information about the distances between all nodes and energy status are gathered. To obtain the values of distances, nodes send advertisement messages to the network. Each node receives these advertisement messages from other nodes at various signal strengths, and then calculates distances using Eq. (5). In the equation, d_{ij} is the distance between node i and node j, f is the communication frequency, c is the speed of light, P^r is the received signal strength, and P^s is the sender signal strength. If the variables of c, P^s, and f are taken as constants then s becomes a constant value used for the calculation of the distance meaning the communication range.

$$d_{ij} = s \cdot (P^r)^{-1/2}, \text{ where } s = c \cdot (P^s)^{1/2}/4\pi f \tag{5}$$

The main structural difference between advanced versions of LEACH and the proposed strategy relies on eliminating additional global positioning system (GPS) hardware to get information about node positions. In the proposed cluster scheme, the distance values are obtained by using signal strengths (Eq. (5)) to achieve routing operations without GPS hardware which would cause extra cost.

Flow chart of the process of gathering advertisement (adv) messages is given in Fig. 3. After receiving and calculating cross-distance values, nodes send these values to the base to be used in the selection process. This process is accomplished by running a centralized program using ABC algorithm. The selection process of cluster-heads needs information of cross distance values between the nodes each-other, including the base. These values are stored in a table as its frame is shown in Table 1. After the first stage is achieved, data gathering steps are performed periodically. In the first step (setup), cluster organization is made by selecting of the cluster-heads for the current tour. Then, sensor nodes are joined to the nearest cluster-heads. After this selection process, periodical data from the network is gathered via the cluster-heads as the second step.

In the proposed approach based on ABC algorithm, selection process of cluster heads is achieved using fitness function obtained analytically in which the communication energy is considered as the significant factor. The distance between the communicating elements is the main concern of energy consumption.

A sender node transmits data signals ensuring that a certain level of signal strength is posed on the receiver antenna to perform a successful communication. The value of the signal strength is decreased by exponentially in transmission of the signal for higher values of distance d. Equation (6) gives the value of signal strength received by

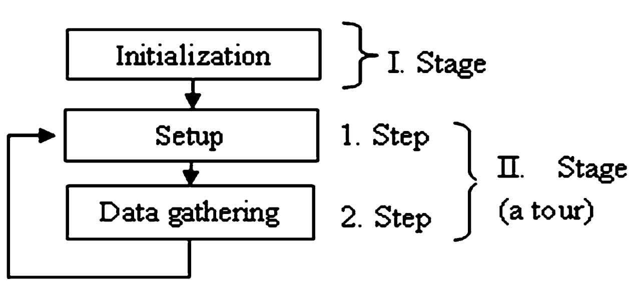

Fig. 2 Main steps of the proposed protocol

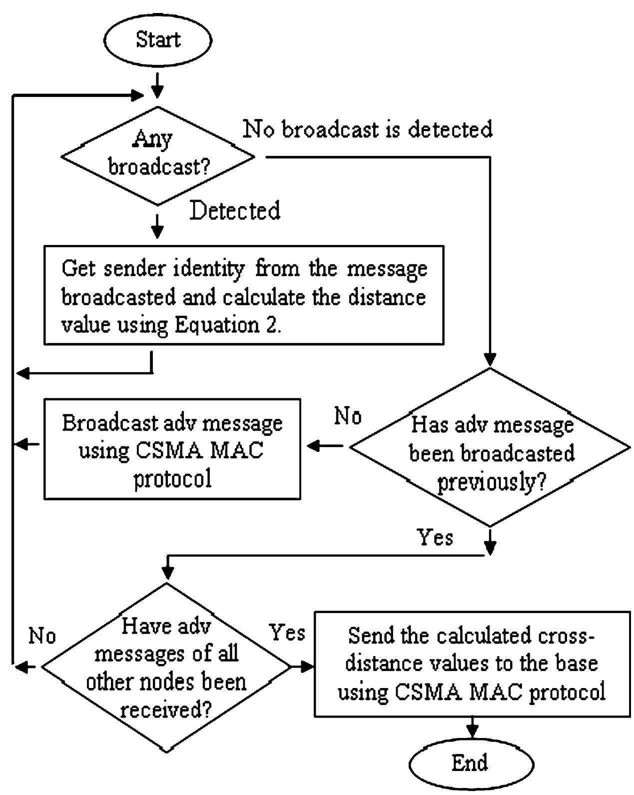

Fig. 3 Gathering of advertisement messages and the calculation of distances

Table 1 Fields of cross distance values

Receiver identity (2-byte/16-bit)	Transmitter identity (2-byte/16-bit)	Distance value (2-byte/16-bit)
–	–	–
	.	.
	.	.
	.	.
	–	–

radio antennas of sensor devices where the communication distance has smaller values than d_{limit} which determines the usability of free-space model [27]. The abbreviations used in the equation are A^s, A^r, λ, d, L, P^s, and P^r representing that antenna gains of the sender and receiver, wavelength of the transmitting signal, distance between the communicating elements, antenna length, transmitting signal strength, and received signal strength, respectively. It is assumed that a receiver device can successfully interpret the signal when received signal strength is equal to or higher than the value of S as given in Eq. (7). This value is a hardware parameter defining the receiving sensitivity of the sensor radios.

$$P^r = \frac{P^s \cdot A^s \cdot A^r \cdot \lambda^2}{(4 \cdot \pi \cdot d)^2 \cdot L} \tag{6}$$

$$S \le P^r \tag{7}$$

when the extension of P^r (Eq. (6)) is used, Eq. (7) becomes:

$$S \le \frac{P^s \cdot A^s \cdot A^r \cdot \lambda^2}{(4\pi)^2 \cdot d^2 \cdot L} \tag{8}$$

when radio sensitivity (S), antenna gains (A^s, A^r), wavelength of the transmitting signal (λ) and antenna length (L) are assumed as constant values represented by α in the Eqs (8), (9) is obtained as:

$$\alpha = \frac{S \cdot (4\pi)^2 \cdot L}{A^s \cdot A^r \cdot \lambda^2}, \quad P^s \ge \alpha \cdot d^2 \tag{9}$$

Equation (9) gives information about the required transmission power value to communicate successfully with the destination element at the distance of d. From the equation, it is seen that the energy consumption is in square ratio of the distance in free-space model. In Fig. 4, communicating elements as sensor nodes and a base station in WSNs using direct communication is illustrated for distances of vector d. Another WSN example is illustrated in Fig. 5 with data gathering by clustering method. In this example, multiple communications occur via cluster heads to transfer a single packet to the base

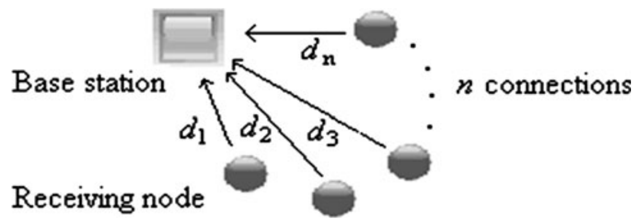

Fig. 4 Data gathering by a direct communication method

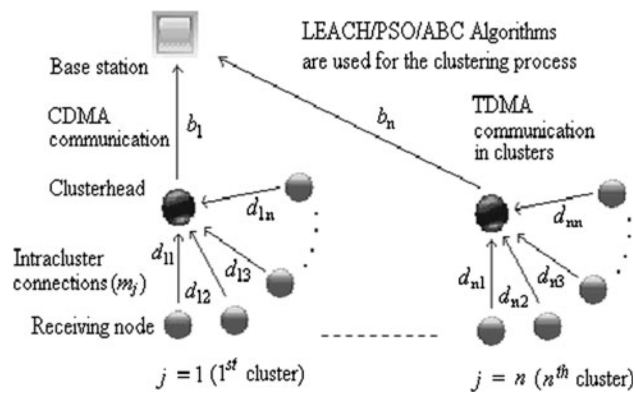

Fig. 5 Data gathering using a clustering method

where n is the number of clusters and m_j is the number of connections of the jth cluster. Vectors b and d represent the distance between the base and cluster heads, and the distance between cluster heads and sensor nodes, respectively.

In the proposed clustering method, a stable base station is assumed with multiple clusters which are grouped dynamically at each tour. A centralized station (base) operates the routing algorithm producing the optimum cluster-heads for the current tour. Then, it assigns sensor nodes to the heads providing the shortest distances. If a node is in the same distance to multiple heads, then the algorithm picks the one prior in its neighbor list. After these operations, base announces the identities of cluster-heads and their members to the network. Then, those cluster-heads serve for a tour. At each tour, data gathering is accomplished in two steps. At the first step, nodes send data sensed nearby to their cluster heads. At the second step, cluster heads fuse collected data and forward it to the base. Since the proposed method uses more complicated process and more calculations, its overhead is expected to be considerably much more than LEACH algorithm. LEACH is a distributive algorithm operating on the sensor nodes having weak hardware capabilities. However, the routing algorithm of the proposed method is operated in the base such as a workstation with a powerful hardware providing the overhead would not be a crucial problem.

Table 2 A solution array

1	2	…	…	n
CH_1	CH_2	…	…	CH_k

3.2 Proposed fitness function for ABC

ABC is used to determine the cluster heads where each solution represents an array having k items in which every item consists a sensor node. A sample of solution array is shown in Table 2.

The ABC employs a population of bees to find the cluster-heads where the bees fly in the search space with k dimensions. Each employed bee corresponds to cluster-heads of sensor nodes.

The fitness of cluster heads selection is stated as a fitness value, which is in inverse proportion to the amount of energy consumption for a tour. If we mention that a certain transfer time is required for a data package, energy consumption is calculated by multiplying transmitting power (P^s) and the time (t). Equations (10) and (11) gives the minimum required energy for a cluster, derived from Eq. (9). In the equations, m is the number of nodes, i is the node index, d_i is the distance between ith node and cluster-head, b is the distance between cluster-head and the base, and E is the transfer energy of the cluster. Considering multiple clusters, the calculation of minimum energy consumption emphasizing the effect of distances will be as in Eq. (12) expressing sum of the energy consumptions of clusters. In the equation, j is the cluster index, d_{ij} is the distance between ith node and jth cluster-head, and b_j is the distance between jth cluster-head and the base.

$$E = \sum_{i}^{m} \left(P_i^s \cdot t\right) \geq \alpha \cdot \left(\sum_{i}^{m} d_i^2 + b^2\right) \cdot t, \tag{10}$$

If w is taken as the multiplication of α ant t, Eq. (11) is obtained.

$$E \geq w \cdot \sum_{i}^{m} d_i^2 + b^2 \tag{11}$$

$$\sum_{j=1}^{n} E \geq w \cdot \sum_{j}^{n} \left(\sum_{i}^{m_j} d_{ij}^2 + b_j^2\right) \tag{12}$$

In order to minimize energy consumption, distances between nodes and cluster-heads, and distances between cluster-heads and the base are considered in the selection process. Since a cluster-head should have enough energy to feed the communication in the current tour as a managing element, energy levels of the candidate nodes have also importance on selection. A candidate cluster head should provide enough energy of receiving messages (E^{RX}) from the nodes in the cluster and transmitting the fused message (E^{TX}) to the base. According to these considerations, fitness function (f^{dist}) is expressed by Eq. (13) (simply inverse of the energy consumption) and the constraints given in Eq. (14) where energy consumptions of E^{TX} and E^{RX} are modeled as in [11]. The abbreviations in the equations are i, j, E_j, m_j, E_{TX}, E_{RX}, k, E^{elec}, E^{amp}, and b are node index, cluster index, energy level of the jth node, number of nodes in the jth cluster, transmit energy, receive energy, number of bits of the transmitting message, parameter of radio electronics, parameter of transmit amplifier, and distance value between jth cluster-head and the base, respectively.

$$f^{CWA} = f^{dist} = \left[w \cdot \sum_{j}^{n} \left(\sum_{i}^{m_j} d_{ij}^2 + b_j^2\right)\right]^{-1}. \tag{13}$$

$$\begin{aligned} E_j &\geq \left(m_j \cdot E^{RX} + E^{TX}\right), \\ E^{RX} &= E^{elec} \cdot k, \\ E_j^{TX} &= E^{elec} \cdot k + E^{amp} \cdot k \cdot b_j^2. \end{aligned} \tag{14}$$

The proposed protocol using fitness function given in Eqs. (13) and (14) simply describes the gateways defined as cluster-heads. The protocol is named as CWA (Cluster based Wireless sensor network routings using Artificial bee colony algorithm) preliminary performance results of which were firstly studied in [22]. To provide a simple scheduled data gathering by cluster heads, TDMA protocol is preferred for data aggregation process in intra-cluster communication. CDMA protocol is preferred for the communication through base station which has more complex hardware than sensor nodes in order to achieve simultaneous data receptions from cluster heads. These protocols are preferred to be able to make a reliable comparison by using the same MAC layer standards as in [12, 19, 20].

3.3 Improvement on the fitness function

Cluster-heads are responsible for fusing the data gathered in the clusters and forwarding it to the base. Since more computation and transferring messages to a distant base are required at the cluster-head nodes, more energy is consumed in these nodes than the others. Batteries of the nodes often selected as cluster-heads can no longer be sufficient to supply the required energy after breaking a critical level. Therefore, the problem of rapid energy depletions in cluster-head batteries is a matter of concern. Fitness calculation is improved by taking into consideration of battery levels of the sensor nodes. It is noticed that when a node serves as a cluster-head frequently, its energy level decreases sharply. In order to increase the network lifetime, it is more appropriate to operate these nodes (having weak batteries) as sensing devices instead of operating as cluster-heads. To be selected as a cluster-head, energy level of each node in the set of candidate cluster-heads should be above than a threshold level in a given solution. This approach defines energy criterion of the fitness function (C), expressed as in Eq. (15) where σ and θ are threshold values, $E_j^{current}$ is current energy level of jth cluster-head (in a set of m elements), and f^{energy} is the fitness value. In simulations, σ is set to the value of 0.1 and the value of θ is set to 0.9 to provide fitness value being between 100 and 90 % until the battery level decreases to a critical level (σ).

$$C = \begin{cases} f^{energy} \geq \theta & \text{if } \min_{j=1}^{m}\left(E_j^{current}\right) \geq \sigma \\ f^{energy} < \theta & \text{otherwise} \end{cases} \tag{15}$$

In the equation, σ is set to 0.1 (10 %) meaning that 90 % of the node battery is discharged. According to the criterion C, the fitness function is aimed to produce a value between 100 and 90 % until 90 % of the energy supply is consumed by setting θ to 0.9. Equation (16) states a mathematical model of this criterion and its curve is plotted in Fig. 6.

$$\tau = -\varphi \cdot \frac{\min_{j=1}^{m}\left(E_j^{current}\right)}{\min_{i=1}^{n}\left(E_i^{current}\right) - \min_{j=1}^{m}\left(E_j^{current}\right)}, \qquad f^{energy} = 1 - e^{\tau} \tag{16}$$

where i is the node index, n is the number of the nodes in the network, j is the index of candidate cluster-head in the current tour, m is the number of cluster-heads, $E^{current}$ is normalized current energy level of the node. The parameter φ, used to adjust the degree of convexity of the fitness function curve, is set to 20.72 to obtain the desired value of f at $\sigma = 10$ % and $\theta = 90$ %.

The routing algorithm which uses fitness function of energy levels besides the fitness function of energy consumption is named as "an Improved version of Cluster based Wireless sensor network routings using Artificial bee colony algorithm (ICWA)".

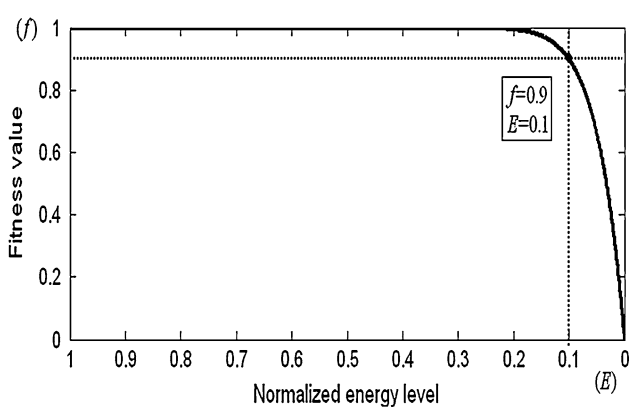

Fig. 6 Fitness function curve of the energy levels

3.4 Proposed quality of service criterion for fitness function

Quality of service (QoS), an important criterion for data communication to be considered, includes parameters such as delay, jitter and bandwidth [28]. Packet delivery delay, one of these QoS parameters, is affected directly from the operation of routing protocol. To provide service quality in the proposed fitness function, delays between delivered packets of clusters are concerned. In the simulations, the influence of delay parameter is observed by total number of received packets in certain times, meaning that smaller delays provide transferring greater number of packets.

In our WSN model, a tour is accomplished by all data is gathered from clusters via cluster-heads. Since TDMA (Time Division Multiple Access) protocol is used for intra-cluster communication, the largest cluster will have the maximum scheduling time for data gathering process. The cluster having the maximum number of elements defines the tour time by producing maximum delay. According to this fact, proposed fitness function is defined as in Eq. (17) as:

$$f^{qos} = \left[\max_{i=1,\dots,n}(m_i + 1)\right]^{-1} \tag{17}$$

where n is number of clusters and m_i member number of ith cluster. Equation 7 is simply inversion of the number of transfer time of the cluster which consists of m_i transfers from nodes to cluster-head and one transfer from cluster-head to the base.

The routing algorithm which uses fitness functions of energy levels, energy consumption and QoS is named as "an Improved version of Cluster based Wireless sensor network routings using Artificial bee colony algorithm considering Quality of service (ICWAQ)".

3.5 Unifying the fitness functions

In the ABC algorithm, the quality of suitability of candidate cluster-heads is defined by a fitness value in the range of [0,1]. The fitness functions derived so far are unified by assigning weights to each of them to get a single fitness value. Equation (18) includes two functions f^{dist} (Eq. (13)) and f^{energy} (Eq. (16)) weighted by parameter β.

$$f^{ICWA} = \beta \cdot f^{dist} + (1 - \beta) \cdot f^{energy} \tag{18}$$

Equation (19) unifies fitness functions f^{qos} (Eq. (17)) and f^{ICWA} (Eq. (18)) by weighting parameter δ. Fitness values in both equations are normalized before the unification.

$$f^{ICWAQ} = \delta \cdot f^{ICWA} + (1 - \delta) \cdot f^{qos} \tag{19}$$

4 Simulation results

The performances of proposed protocols (CWA, ICWA and ICWAQ) are tested with various parameter settings using a parallel discrete model providing periodical data transferring, developed in Matlab. In the simulations, the network consists of 100 nodes randomly placed in a fixed area of 500 m × 500 m and a single base station located at (250, 575 m) near middle edge of the field. It is assumed that every node has a capability of communicating with other nodes in the field as well as the base station. Free space radio model with isotropic antennas is used as discussed in [27] with parameters of radio electronics E^{elec}, power amplifier of transmitter E^{amp} as taken in [11], packet size K and communication frequency f for receiving and transmitting units, where the values of these parameters are taken as 50nJ/bit, 100pJ/bit/m^2, 512Kbit, and 250 Kbit/s, respectively. In the experiments, all sensors are sensing the environment at a fixed rate as assumed in [11], which provides transferring all data. In the model, energy loss due to channel transmission is assumed to be in proportion to d^2 (square of distance).

To evaluate the performances of the proposed protocols, simulation results are compared to that of LEACH protocol [11] and PSO based protocols [20] by obtaining mean values of 10 runs. The probability value of a node to be selected as a cluster head is taken as 0.05 for the LEACH algorithm, and the number of cluster-heads is set to 5 for ABC and PSO algorithms. E^{TX} and E^{RX} energy dissipations are calculated by using distance d and the number of messages (k bit) as in Eqs. (20) and (21):

$$E^{RX} = E^{elec} \cdot k, \tag{20}$$

$$E^{TX} = E^{elec} \cdot k + E^{amp} \cdot k \cdot d^2 \tag{21}$$

We used a small sized lithium battery model suitable for sensor nodes which can supply about 100 Joules (J) [29]. Initially, nodes have energy levels of 100 J. To provide a more realistic WSN modeling, different energy levels are issued as random 20 % of the nodes are given 100 J and remaining 80 % are given 40 J in the initial settings of the simulations. In the experiments, the network is up until the number of alive nodes is 20 % of all nodes or in the case of having less than 5 candidate heads, the network energy is assumed as depleted and the simulations ended.

In simulation runs, it can be easily observed that distance between transmitter and receiver nodes is the key value determining the energy consumption as well as radio parameters, E^{elec} and E^{amp} which provide the qualities of radio transmitter and receiver modules. To make a reliable comparison, the values of E^{elec} and E^{amp} are chosen 50nJ/bit and 100pJ/bit/m^2 respectively, as indicated in [11]. Figures 7 and 8 illustrate individual effects of the parameters in the neighborhood of chosen values, giving by total energy consumption in joules to send a *k-bit* packet to the receiver node at the distance *d*. From the figures, it is seen that having lower values of E^{elec} and E^{amp} provides lower energy consumption as well as lower radio transmission quality, causing lower distance communication.

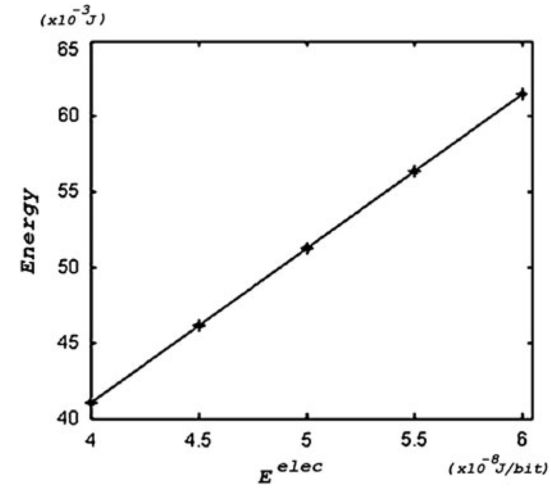

Fig. 7 Total energy dissipation for various E^{elec} values

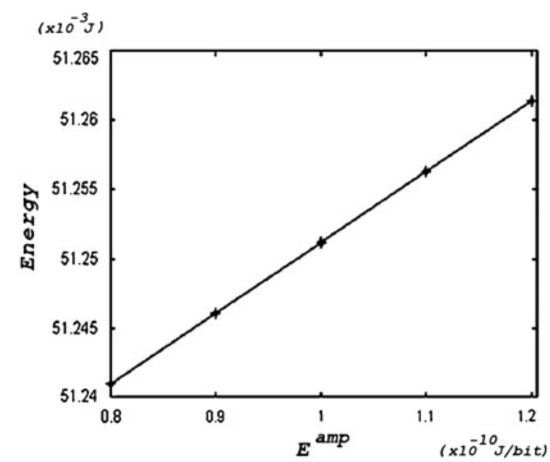

Fig. 8 Total energy dissipation for various E^{amp} values

Distance between the sender and receiver nodes is also another important physical parameter defining the energy consumption. In Fig. 9, it is shown that energy dissipation is increased in square ratio of distance by means of free-space radio model.

Total number of signals gathered from the network during the network life-time is aimed to be maximized. Total energy consumption of the nodes in a tour is an essential parameter defining the network life-time. Since energy depletions of the cluster-heads are much more than the others, energy levels of these nodes should be considered. To take account of these considerations, Eq. (18) depending on the value of β is proposed.

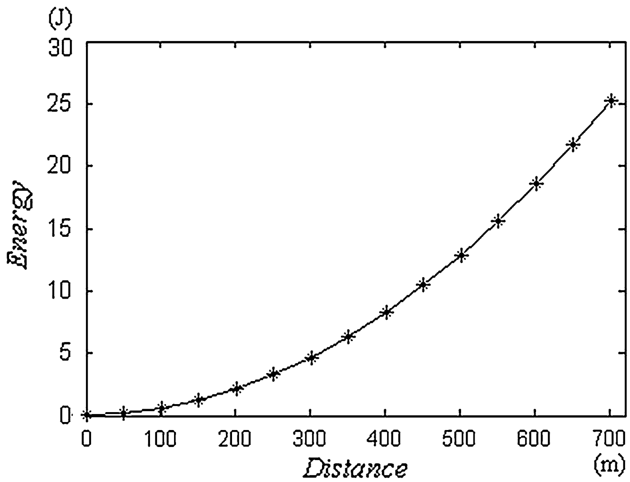

Fig. 9 Total energy dissipation for various distances

Figure 10 shows the performance results of ICWA protocol for 7 different values of β in the range of [0,1], given in Table 3. It is clear that when setting β to 0, ICWA protocol behaves as CWA protocol by discarding energy levels. These simulation runs stopped when 80 % of the nodes are dead (Fig. 10(a)) and 80 % of the network energy is consumed (Fig. 10(b)) to observe the total number of received signals. As seen from the figures, curve of $\beta = 0.3$ remains higher than the others in general, provides greater number of nodes alive in the most of the region. Therefore, it is evaluated as that this parameter value for β provides longer network life-time than the others. It is also seen that the value of β different than 0.3 cause receiving smaller signal numbers during the network life-time.

ICWAQ protocol uses a QoS parameter addition to ICWA metrics which are energy consumption and energy levels. A parameter (δ) is used to weight the QoS value in the unified fitness function, Eq. (19). If the parameter δ is set to 0, ICWAQ protocol behaves as ICWA protocol by

Table 3 The best parameter values of β ve δ

Parameter	Set of values	The best
β	[0, 0.1, 0.3, 0.5, 0.7, 0.9, 1]	0.3
δ	[0, 0.1, 0.3, 0.5, 0.7, 0.9, 1]	0.9

discarding the QoS criterion. Figure 11 shows the performance results of ICWAQ protocol for 7 different values of δ in the range of [0,1]. The simulations run until 80 % of the nodes are dead (Fig. 11(a)) and 80 % of the network energy is consumed (Fig. 11(b)) to observe the total number of received signals, as in the observation of β parameter where the value of it is set to 0.3. In Fig. 11, it is seen that the curve using 0.9 for δ provides more alive nodes than the others during the life-time in general. From the figure, it is also seen that the curve of 0.9 for δ provides an energy efficient communication by transferring greater number of signals.

Figure 12(a) shows the performance curves of ICWAQ protocol for different δ values in terms of the spent time versus number of received signals. Although 0.1, 0.3, and 0.5 valued δ curves provide transferring more signals for a limited period from the beginning, they cause shorter network life-time. Taking account of whole network life-time, the curve using the value of 0.9 causes shorter delays to transfer the same amount of signals. Therefore, the value of 0.9 is concluded as the best parameter value giving good performance for QoS criterion providing the maximum network life-time. When QoS is not considered ($\delta = 1$), energy criterion function (f^{ICWA}) would be more effective on the decision, providing longer network life time instead of considering service quality. In Table 3, the best values of β and δ are shown providing maximum energy saving in the given set. The proposed ICWAQ protocol uses these best values of the parameters.

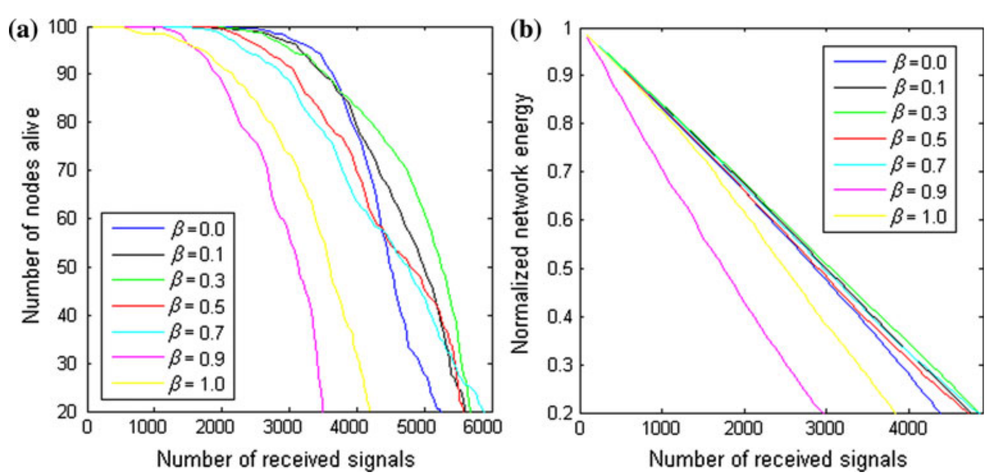

Fig. 10 **a** β values for number of alive nodes **b** β values for residual network energy

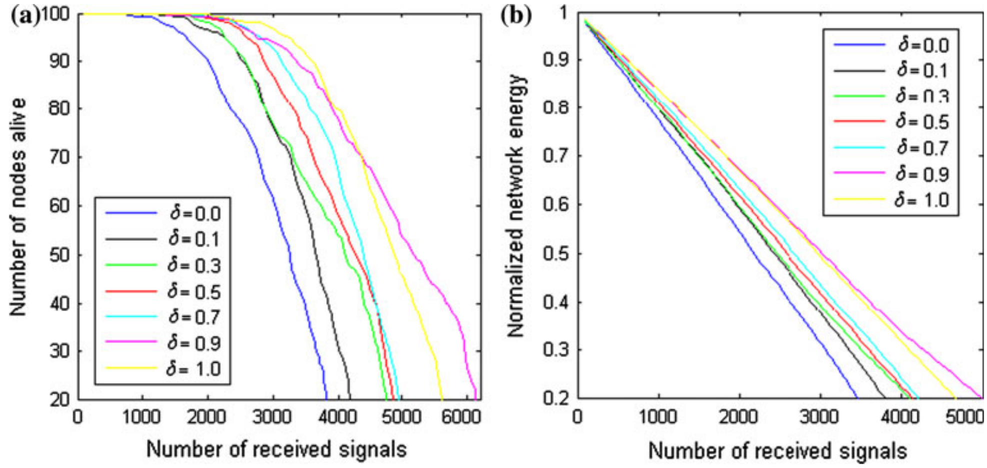

Fig. 11 **a** δ values for number of alive nodes **b** δ values for residual network energy

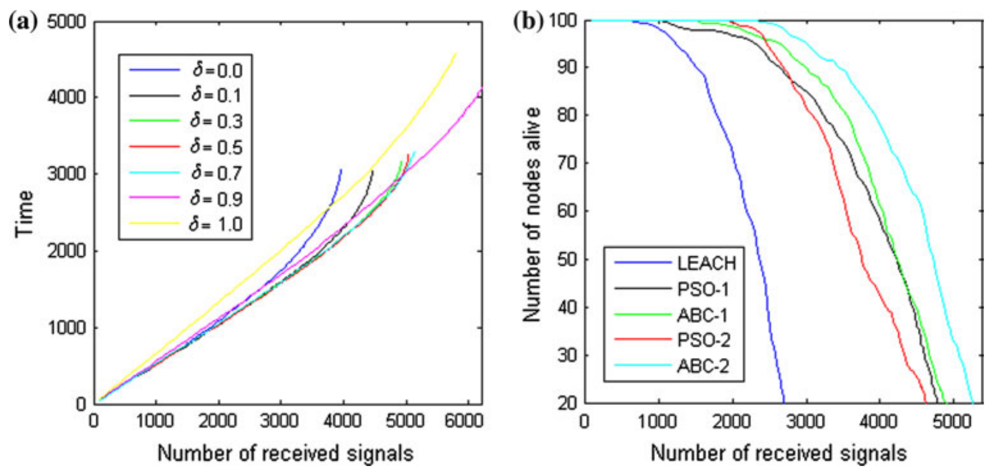

Fig. 12 **a** δ values for transfer times **b** number of alive nodes for different protocols

Performance comparisons of the proposed protocol (ICWAQ) with the protocols PSO-1 and LEACH are given in Figs. 12(b) and 13. The detailed information of the protocols can be found in [11] and in [20] including comparisons with the protocols using direct transmission, MTE routing, static clustering, K-means clustering and GA.

Equations of fitness functions of the protocols (LEACH, PSO-1, ABC-1, PSO-2, ABC-2, ICWAQ) are given briefly in Table 4. In the studies of [11, 15, 20], it is concluded that PSO-1 protocol presents more successful results than the others. In order to make a reliable comparison, ABC-1 protocol is studied using the same fitness function as in PSO-1 [20] as well ABC-2 and PSO-2 protocol use the

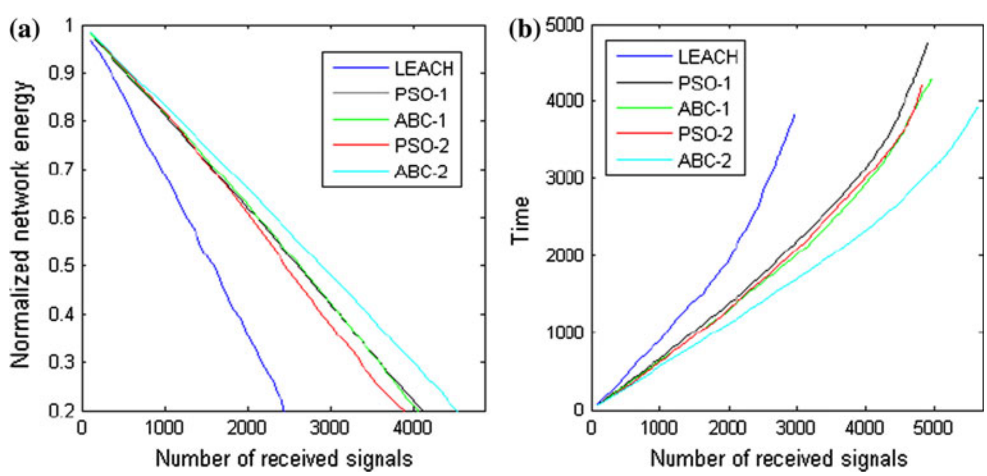

Fig. 13 **a** Normalized network energy **b** transfer times

Table 4 Fitness functions used by protocols

Function	Protocol
$T(n) = \begin{cases} \frac{P}{1-P\cdot[r \bmod (1/P)]} & \text{if } n \in G, \\ 0 & \text{otherwise} \end{cases}$	LEACH
$F = (f_1 + f_2)/2$ $f_1 = \left\{ \max_{j=1,\ldots,n} \left[\sum_{i=1}^{m_j} \left(d_{ij}/m_j \right) \right] \right\}^{-1}$ $f_2 = \sum_{j=1}^{n} E_j \Big/ \sum_{j=1}^{n} \left(\sum_{i=1}^{m_j} E_{ij} \right)$	PSO-1, ABC-1
$f^{ICWAQ} = \delta \cdot f^{ICWA} + (1-\delta) \cdot f^{qos}$ where $f^{qos} = \left[\max_{i=1,\ldots,n} (m_i + 1) \right]^{-1}$	PSO-2, ABC-2 (ICWAQ)

same fitness function. The performance results of mentioned protocols including LEACH are given in the following Figs. 12(b) and 13.

The performances of the protocols are observed by plotting number of nodes alive, normalized residual network energy, and total delivery time versus total number of received signals that can be seen in Figs. 12(b), 13(a) and (b), respectively. Increasing the performance is provided by preserving energy and maintaining the number of nodes alive as much as possible. Fig. 12(b) shows that ABC-2 protocol causes minimum deaths during operation of data gatherings. When the number of alive nodes decreases to half, the protocol receives about 1.5 times more signals than the others. Figure 13(a) gives the number of received signals via the average residual network energy where ABC-2 protocol performs more success than the others by saving more energy consumption. The differences in the performance can be easily seen where the network energy closes to be emptied. In Fig. 13(b), it is seen that ABC-2 protocol performs greater amount of data gathering in shorter times providing a successful result in service quality. On the other hand, it is seen that the protocols PSO-1, ABC-1, and PSO-2 achieve similar performance where three of the protocols provide transferring about 5,000 received signals in similar times. In general, ABC-2 protocol shows the best performance, while performances of others sorted as ABC-1, PSO-1, PSO-2, and LEACH. In the appendix, the positions of the nodes and the numbers of signals received for given configurations based on the used algorithms of the network are demonstrated. Configurations are illustrated by rounds represented as r where j, m, n, s and Σ are the indexes of the cluster-heads, number of members of each cluster-head, number of alive nodes having enough energy to communicate, number of signals received for each cluster-head, and total number of received signals. The value of s for the ith configuration is given by Eq. (22):

$$s_i = z + \sum_{j=1}^{z} n\left(e_{ij} > E\right) \tag{22}$$

where j is the cluster index, z is the number of clusters, e_{ij} is the energy level of the ith member of the jth cluster, and n is a function that produces the number of nodes having minimum energy level (E) to communicate for a tour.

The success of the proposed approach relies on the optimum cluster-heads selection. A sample distribution in simulated network which is clustered by ABC algorithm is shown in Fig. 14 where cluster-heads are represented with squares and members of the clusters are represented with other markers. As can be seen from the figure, cluster heads are uniformly selected by ABC algorithm providing that clusters have approximately equal sized of regions.

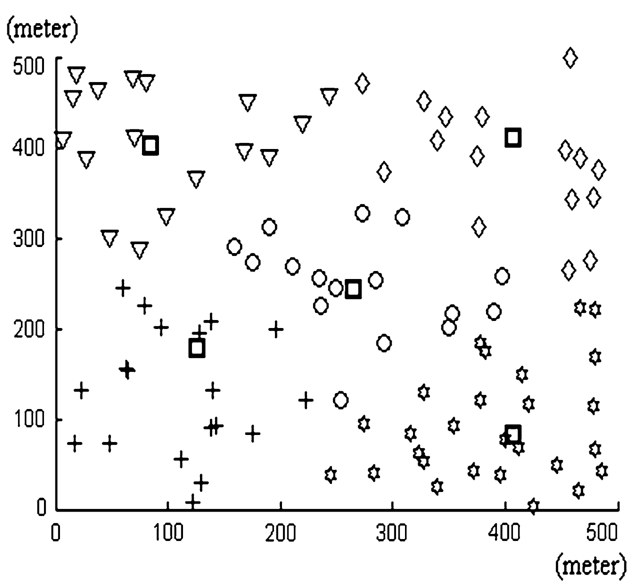

Fig. 14 A sample distribution of sensor nodes

5 Conclusion

The main goal of the monitoring applications of WSNs is to gather information from the field periodically. Increasing the total number of gathered signals during the network life-time is essential to get maximum benefit from the WSNs. In this paper, a novel energy saving routing method providing longer network life time is achieved by gathering greater amount of signals from the field. The proposed protocol ICWAQ uses efficient and fast searching features of the ABC algorithm to optimize clustering of the nodes in the selection process of cluster-heads defining routing gateways. The clustering success of the ABC algorithm is compared with the protocols based on LEACH and PSO. The protocol ICWAQ not only prolongs the network life-time, but also employs a service quality mechanism by considering delays between the signals received from the clusters. Simulation results prove that ICWAQ routing protocol can effectively maximize the network life-time and minimize transfer delays comparing the other techniques. As a future work, we are planning to study clustering with ABC algorithm on routing of networks including mobile nodes as well as comparing the performance of the algorithm against other well-known optimization algorithms and to analyze the effect of the noisy channels and other physical and MAC layer issues on the success of the clustering approaches.

References

1. Akyildiz, I., & Mehmet, C. V. (2010). WSN applications. *Wireless Sensor Networks, 1*, 17–35.
2. Giuseppe, A., Marco, C., Mario, D. F., & Andrea, P. (2009). Energy conservation in wireless sensor networks: A survey. *Ad Hoc Networks, 7*, 537–568.
3. Yang, J., Zhang, C., Li, X., Huang, Y., Fu, S., et al. (2010). Integration of wireless sensor networks in environmental monitoring cyber infrastructure. *Wireless Networks, 16*(4), 1091–1108.
4. Akyildiz, I., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks, 38*, 393–422.
5. Ergen, S. C., & Varaiya, P. (2010). TDMA scheduling algorithms for wireless sensor networks. *Wireless Networks, 16*(4), 985–997.
6. Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks, 52*, 2292–2330.
7. Goldsmith, A. J., & Wicker, S. B. (2002). Design challenges for energy-constrained ad hoc wireless networks. *IEEE Wireless Communications, 9*, 8–27.
8. Anastasi, G., Conti, M., Falchi, A., Gregori, E., & Passarella, A. (2004). Performance measurements of motes sensor networks. In *Proceedings of the 7th ACM International Symposium on modeling, analysis and simulation of wireless and mobile systems* (pp. 174–181).
9. Crossbow Technology, Inc. (2010). MICAz module datasheet. Available at: http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/MICAz_Datasheet.pdf.
10. Al-Karaki, J. N., & Kamal, A. E. (2004). Routing techniques in wireless sensor networks: A Survey. *IEEE Wireless Communications, 11*, 6–28.
11. Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocols for wireless microsensor networks. In *Proc. hawaaian int. conf. on systems science* (pp. 1–10).
12. Heinzelman, W. (2000). *Application specific protocol architectures for wireless networks*. PhD Thesis, MIT.
13. Lindsey, S., & Raghavendra, C. (2002). Pegasis: Power-efficient gathering in sensor networks. In *Proceedings of IEEE aerospace conference* (Vol. 3, pp. 9–16).
14. Huang, Y., Wang, N., & Chen, M. (2008). Performance of a hierarchical cluster-based wireless sensor network. In *IEEE International Conference on ubiquitous and trustworthy computing* (pp. 349–354).
15. Dorigo, M., & Caro, D. G. (1999). Ant colony optimization: A new meta-heuristic. In *Proceedings of CEC99 Congress on Evolutionary Computation* (pp. 1470–1477).
16. Okdem, S., & Karaboga, D. (2009). Routing in wireless sensor networks using an Ant Colony Optimization (ACO) router chip. *Sensors, 9*, 909–921.
17. Qing, L., Zhi, T., Yuejun, Y., & Yue, L. (2009). Monitoring in industrial systems using wireless sensor network with dynamic power management. *IEEE Sensors Journal, 9*, 1596–1604.
18. Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In *Proc. IEEE Int. Conf. on neural networks* (Vol. 4, pp. 1942–1948), Piscataway.
19. Liang, Y., & Yu, H. (2005). PSO-based energy efficient gathering in sensor networks. *Lecture Notes in Computer Science, 3794*, 362–369.
20. Latiff, N. M. A., & Sharif, B. S. (2007). Performance comparison of optimization algorithms for clustering in wireless sensor networks. In *IEEE Int. Conf. on mobile adhoc and sensor systems* (pp. 1–4), Pisa.
21. Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. In *Technical Report-TR06*, Erciyes University, Engineering Faculty, Computer Engineering Department.
22. Karaboga, D., Okdem, S., & Ozturk, C. (2010). Cluster based wireless sensor network routings using artificial bee colony algorithm. In *Int. Conf. on autonomous and intelligent systems, AIS'2010* (pp. 1–5), Portugal.
23. Karaboga, D., & Ozturk, C. (2011). A novel clustering approach: Artificial bee colony (ABC) algorithm. *Applied Soft Computing, 11*, 652–657.
24. Karaboga, D., Ozturk, C., & Gorkemli, B. (2011). Probabilistic dynamic deployment of wireless sensor networks by artificial bee colony algorithm. *Sensors, 11*(6), 6056–6065.
25. Karaboga, D., Gorkemli, B., Ozturk, C., & Karaboga, N., (2012). A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Artificial Intelligence Review*. doi:10.1007/s10462-012-9328-0.
26. Thomas, A. B., Christopher, M., Szymanski, B. K., & Branch, J. W. (2008). Self-selecting reliable paths for wireless sensor network routing. *Computer Communications, 31*, 3799–3809.
27. Bajaj, L., Takai, M., Ahuja, R., Tang, K., Bagrodia, R., & Gerla, M. (1999). GloMoSim: A scalable network simulation environment. In *Technical Report 990027*, Computer Science Department, University of California, Los Angeles.
28. Liu, Z., Kwiatkowska, M. Z, & Constantinou, C. (2005). A biologically inspired qos routing algorithm for mobile ad hoc networks. In *Int. conf. on adv. inf. network applications* (pp. 426–431).
29. GPI Research Group. CR1216 Battery catalog. *GPI International Ltd.*, Available at http://www.gpbatteries.com/pic/CR1216_DS.pdf.

Author Biographies

Dervis Karaboga received the B.Sc. degree in 1983 from the Department of Electronics Engineering, Erciyes University, Turkey and the M.Sc. degree in 1988 from the Department of Electronics and Communication Engineering, Istanbul Technical University, Turkey, and the Ph.D degree in 1994 from Systems Engineering Department, University of Wales, College of Cardiff, UK. He is currently a Professor at the Department of Computer Engineering, Erciyes University, Turkey. His research areas include optimization, fuzzy systems, neural networks, and engineering applications of intelligent methods.

Selcuk Okdem is an Assistant Professor at the Department of Computer Engineering, Engineering Faculty, Erciyes University. He received the degrees of B.Sc., M.S., and Ph.D. from Computer Engineering of Erciyes University, Turkey. His Ph.D. thesis focused on the wireless sensor network routing protocols. His researches are based on wireless communication, hardware description languages, and optimization on computer architecture.

Celal Ozturk received the doctorate in Computer Engineering from the Erciyes University. He received his M.S. degree form Electrical and Computer Engineering, Rutgers University and his B.Sc. degree form Computer Engineering, Erciyes University. From 2007 to 2011, he was a lecturer at the Department of Computer Engineering, Erciyes University and he is currently working as an Assistant Professor. His research interests include training neural networks, clustering, and optimization in wireless sensor networks.