

Improved metaheuristic-based energy-efficient clustering protocol with optimal base station location in wireless sensor networks

Palvinder Singh Mann¹ · Satvir Singh¹

Published online: 13 September 2017
© Springer-Verlag GmbH Germany 2017

Abstract Efficient clustering is a well-documented NP-hard optimization problem in wireless sensor networks (WSNs). Variety of computational intelligence techniques including evolutionary algorithms, reinforcement learning, artificial immune systems and recently, artificial bee colony (ABC) metaheuristic have been applied for efficient clustering in WSNs. Due to ease of use and adaptive nature, ABC arose much interest over other population-based metaheuristics for solving optimization problems in WSNs. However, its search equation contributes to its insufficiency due to comparably poor exploitation cycle and requirement of certain control parameters. Thus, we propose an improved artificial bee colony (iABC) metaheuristic with an improved solution search equation to improve exploitation capabilities of existing metaheuristic. Further, to enhance the global convergence of the proposed metaheuristic, an improved population sampling technique is introduced through Student's t-distribution, which require only one control parameter to compute and store and therefore increase efficiency of proposed metaheuristic. The proposed metaheuristic maintain a good balance between exploration and exploitation search abilities with least memory requirements; moreover, the use of first-of-its-kind compact Student's t-distribution makes it suitable for limited hardware requirements of WSNs. Additionally, an energy-efficient clustering protocol based on iABC metaheuristic is presented, which inherits the capabilities of the proposed metaheuristic to obtain optimal cluster heads along with

an optimal base station location to improve energy efficiency in WSNs. Simulation results show that the proposed clustering protocol outperforms other well-known protocols on the basis of packet delivery, throughput, energy consumption, network lifetime and latency as performance metric.

Keywords Wireless sensor networks · Energy-efficient clustering · Improved artificial bee colony (iABC) metaheuristic

1 Introduction

Developments in wireless communication and electronics have led to the development of low-cost, low-power, multi-functional WSNs that can operate in wide variety of diverse and complex environments, even where human activity is not possible at all. WSNs contain self-configured, distributed and autonomous sensor nodes (SNs) that monitor physical or environmental activities like humidity, temperature or sound in a specific area of deployment (Yick et al. 2008; Zhu et al. 2012). SNs can have more than one sensor to capture data from the physical environment, wherever deployed. A sensor with limited storage and computation capabilities receives the sensed data through analogue to digital converter (ADC) and processes it further for transmission to a main location, known as base station (BS), where the data can be analyzed for decision making in variety of applications (Al-Karaki and Kamal 2004). Every node also acts as a repeater for passing information of other sensor nodes to the sink. The most important part of the sensor node is its power supply, which caters to the energy requirements of sensors, processors and transceiver; however, its limited battery life can lead to premature exhaust of the net-

Communicated by V. Loia.

✉ Palvinder Singh Mann
psmaan@hotmail.com

¹ IKG Punjab Technical University, Kapurthala, Punjab 144603, India

work due to excessive usage (Akkaya and Younis 2005). As manual recharging of batteries is not possible in complex deployments, efficient use of the energy becomes a tough challenge in applications where a prolonged life of the network is required (Gaura 2010). Researchers are heavily involved in designing of energy-efficient solutions; however, on the other hand network life can also be extended by planning energy-efficient approaches. It is well accepted that cluster-based hierarchical approach is an efficient way to save energy for distributed WSNs (Abbasi and Younis 2007; Tyagi and Kumar 2012), which increase network life by effectively utilizing the node energy and support dynamic WSNs environment. In a cluster-based WSN, SNs are divided into several groups known as clusters with a group leader known as cluster head (CH). All the SNs sense data and send it to their corresponding CH, which finally sends it to the BS for further processing. Clustering has various significant advantages over classical schemes (Abbasi and Younis 2007). First, data aggregation is applied on data, received from various SNs within a cluster, to reduce the amount of data to be transmitted to BS thus energy requirements decrease sharply. Secondly, rotation of CHs helps to ensure a balanced energy consumption within the network, which prevent getting specific nodes starved due to lack of energy (Chamam and Pierre 2010). However, selection of appropriate CH with optimal capabilities while balancing energy efficiency ratio of the network is a well-defined NP-hard optimization problem in WSNs (Khalil and Attea 2011). Thus, computational intelligence (CI) (Kulkarni et al. 2011)-based approaches such as evolutionary algorithms (EAs), reinforcement learning (RL), artificial immune systems (AIS), and more recently, artificial bee colony (ABC) have been used extensively as population-based metaheuristic for energy-efficient clustering protocols in WSNs (Das et al. 2009). Results prove that the performance of the ABC metaheuristic is competitive to other population-based algorithms with the advantage of employing fewer control parameters, simplicity of use and ease of implementation (Samrat and Udgata 2010). However, similar to other population-based algorithms, the standard ABC metaheuristic also face some challenges, as it is considered to have poor exploitation phase than exploration; moreover, convergence rate is typically slower, especially while handling multimodal optimization problems (Karaboga and Akay 2009). Therefore, we propose an improved artificial bee colony (iABC) metaheuristic, with an improved solution search equation, which will be able to search an optimal solution to improve its exploitation capabilities and an improved technique for population sampling through the use of first-of-its-kind compact Student's t-distribution to enhance the global convergence of the proposed metaheuristic. Further, to utilize the capabilities of the proposed metaheuristic, an improved artificial bee colony-based clustering protocol, Becluster is introduced, which

selects optimal cluster heads (CHs) with energy-efficient approach in WSNs.

2 Related work

A large variety of clustering protocols have been developed so far for WSNs. Here, we present only the vital contributions of the researchers based on Classical as well as CI-based metaheuristic approach; however, we underline the significance of CI, as our proposed protocol is part of this approach. Low-energy adaptive clustering hierarchy (LEACH) (Heinzelman et al. 2002) is a classical clustering protocol which combines energy-efficient cluster-based routing to application-oriented data aggregation and achieve better lifetime for a WSN. LEACH introduces algorithm for adapting clusters and rotating CHs positions to evenly distribute the energy load among all the SNs and thus enables self-organization in WSNs. LEACH remain a paradigm architect for designing clustering protocols for WSNs till date. HEED (hybrid energy-efficient distributed clustering) (Younis and Fahmy 2004) is another classical clustering protocol that selects CHs based on hybridization of node residual energy and node proximity to its neighbors or node degree thus achieves uniform CH distribution across the network. HEED approach can be useful to design WSNs protocols that require scalability, prolonged network lifetime, fault tolerance, and load balancing, but it only provided algorithms for building a two-level hierarchy and no idea is presented for designing protocol to multilevel hierarchies. Power-efficient and adaptive clustering hierarchy (PEACH) (Yi et al. 2007) selects CHs without additional overhead of wireless communication and supports adaptive multilevel clustering for both location-unaware and location-aware WSNs but with high latency and low scalability, thus making it suitable only for small networks. T-ANT (Selvakennedy et al. 2007) is a swarm-inspired clustering protocol which exploits two swarm principles, namely separation and alignment, through pheromone control to obtain a stable and near-uniform distribution for selection of CHs. Energy-efficient multilevel clustering (EEMC) (Jin et al. 2008) achieves less energy consumption and minimum latency in WSNs by forming multilevel clustering with minimum algorithm overhead. However, the authors ignored the issue of channel collision which happens frequently in wireless networks. Energy-efficient heterogeneous clustered scheme (EEHC) (Kumar et al. 2009) selects CHs based on weighted election probabilities of each node, which is a function of their residual energy and further support node heterogeneity in WSNs. Multipath routing protocol (MRP) (Yang et al. 2009) is based on dynamic clustering with ant colony optimization (ACO) metaheuristic. A CH is selected based on residual energy of nodes, and an improved ACO algorithm is

applied to search multiple paths that exist between the CH and BS. MRP prolonged the network lifetime and reduces the average energy consumption effectively using proposed metaheuristic. Energy-efficient cluster formation protocol (EECF) (Chamam and Pierre 2010) presents a distributed clustering algorithm where CHs are selected based on a three-way message exchange between each sensor and its neighbors while possessing maximum residual energy and degree. However, the protocol does not support multilevel clustering and consider small transmission ranges. Mobility-based clustering (MBC) protocol (Deng et al. 2011) supports node mobility; hence, CHs will be selected based on nodes residual energy and mobility, whereas a non-CH node maintains link stability with its CH during setup phase. UCFIA (Mao and Zhao 2011) is a novel energy-efficient unequal clustering algorithm for large scale WSNs, which use fuzzy logic to determine node's chance to become CH based on local information such as residual energy, distance to BS and local density of nodes. In addition, an adaptive max–min ACO metaheuristic is used to construct energy-aware inter-cluster routing between CHs and BS, thus balancing the energy consumption of CHs. Distributed energy-efficient clustering with improved coverage (DEECIC) (Liu et al. 2012) selects minimum number of CHs to cover the whole network based on nodes local information and periodically updates CHs according to nodes residual energy and distribution. By reducing overheads of time synchronization and geographic location information, it prolongs network lifetime and improves network coverage. Energy-aware evolutionary routing protocol (ERP) (Attea and Khalil 2012) is based on evolutionary algorithms (EAs) and ensures better trade-off between lifetime and node stability period of a network with efficient energy utilization in complex WSNs environment. Harmony search algorithm-based clustering protocol (HSACP) (Hoang et al. 2014) is a centralized clustering protocols based on harmony search algorithm (HSA), a music-inspired metaheuristic, which is designed and implemented in real time for WSNs. It is designed to minimize the intra-cluster distances between the cluster members and their CHs, thus optimizing the energy distribution for WSNs. BeeSensor (Saleem and Farooq 2012) is an energy-aware, event-driven, reactive and on-demand routing protocol for WSNs. Inspired from biological system of bees and based on a typical bee agent model, which works with four type of agents, namely packers, scouts, foragers and swarms, BeeSensor demonstrates good performance over other CI-based protocols with least communication and processing cost. One major drawback of the protocol is its flat nature or non-cluster-based approach, which affects its performance on various fronts. Kula and Jana (2014) presents a linear/nonlinear programming (LP/NLP) formulations of energy-efficient clustering and routing problems in WSNs, followed by two algorithms for the same based on a par-

ticle swarm optimization (PSO). Their proposed algorithms demonstrate their proficiency in terms of network life, energy consumption and delivery of data packets to the BS. Ozturk and Hancer (2015) presents an improved binary artificial bee colony algorithm (IDisABC), which can be used for real-world optimization problems of dynamic clustering with hands-on experiments conducted on data and image clustering benchmark problems. Ari and Yenke (2016) proposes ABC-SD, a novel centralized cluster-based routing protocol based on artificial bee colony metaheuristic with a multi-objective fitness function to select CHs along with a cost-based function for routing, which makes a trade-off between the energy efficiency and the number of hops in the routing path. Extensive evaluation demonstrate the effectiveness of the proposed protocol in terms of network lifetime, network coverage and packet delivery ratio. Ding et al. (2016) presents a rule-driven multipath routing algorithm with dynamic immune clustering, which applies biological immune system to the event-driven dynamic clustering algorithm for WSNs. Below in Table 1, we present a relative comparison of these protocols, highlighting their features and limitations for a better insight:

It is very much clear from the comparison that classical as well as CI-based approaches have their own features as well as limitations. Classical approaches are better in self-organization, load balancing with minimum overhead but average in energy efficiency, whereas CI-based metaheuristic showed to be good in energy efficiency with prolonged network life. Therefore, CI-based metaheuristic approaches need to be further explored and improved for energy-efficient solutions in WSNs.

3 Standard artificial bee colony (ABC) metaheuristic

Original artificial bee colony (ABC) metaheuristic is proposed by Karaboga and Akay (2009) for optimizing multivariable and multimodal continuous functions, which has aroused much interest in research community due to less computational complexity and use of less number of control parameters. Moreover, optimization performance of ABC is competitive to well-known state-of-the-art metaheuristics (Karaboga and Basturk 2008). In ABC, there are three type of bees: employed bees, onlookers and scout bees (Zhang and Wu 2011). The employed bee carries exploitation of a food source and share information like direction and richness of food source with the onlooker bee, through a waggle dance; thereafter, onlooker bee will select a food source based on a probability function related to the richness of that food source, whereas scout bee explores new food sources randomly around the hive. When a scout or onlooker bee finds a new food source, it becomes employed again; on the other

Table 1 Relative comparison of protocols in WSNs

Protocol	Classification	Energy efficiency	Features	Limitations
LEACH (Heinzelman et al. 2002)	Classical	Average	Self-organization	High communication cost
HEED (Younis and Fahmy 2004)	Classical	Average	Low communication cost	High latency
PEACH (Yi et al. 2007)	Classical	Average	Load balancing	High latency, low scalability
T-ANT (Selvakennedy et al. 2007)	Computational intelligence	Good	Fast convergence, low overhead	Low coverage
EEMC (Jin et al. 2008)	Classical	Average	Minimum overhead, low latency	Only uniform node distribution
EEHC (Kumar et al. 2009)	Classical	Good	Support node heterogeneity	Low scalability
MRP (Yang et al. 2009)	Computational intelligence	Good	Prolong network lifetime	Need parameters adjustment
EECF (Chamam and Pierre 2010)	Classical	Average	Prolong network lifetime	Low transmission range
MBC (Deng et al. 2011)	Classical	Average	High node mobility, low packet loss	High communication cost
UCFIA (Mao and Zhao Cl 2011)	Computational intelligence	Good	Prolong network lifetime	Need parameters adjustment
DEECIC (Liu et al. 2012)	Classical	Average	Better network coverage	Low scalability
ERP (Attea and Khalil 2012)	Computational intelligence	Average	Better network lifetime	Non-cluster-based approach
HSACP (Hoang et al. 2014)	Computational intelligence	Good	Fast convergence	No load balancing
BeeSensor (Saleem and Farooq 2012)	Computational intelligence	Good	Low processing cost	Non-cluster-based approach
PSO (Kuila and Jana 2014)	Computational intelligence	Good	Better packet delivery	Network overhead
IDisABC (Ozturk and Hancer 2015)	Computational intelligence	Good	Better throughput	High latency
ABC-SD (Ari and Yenke 2016)	Computational intelligence	Good	Low energy consumption	Low scalability

hand, when a food source has been fully exploited, all the employed bees will abandon the site and may become scouts again. In ABC metaheuristic, a food source corresponds to a possible solution to the optimization problem and the number of employed bees are equal to the number of food sources.

Below we present the detail procedure of ABC metaheuristic in different phases.

3.1 Initialization phase

ABC metaheuristic starts with initial population number (PN), randomly generated through D-dimensional real set of vectors. Let $x_{ij} = \{x_{i1}, x_{i2}, \dots, x_{iD}\}$ is the i th food source, where $j = 1, 2, \dots, D$ is obtained by:

$$x_{ij} = x_{\min_j} + \text{rand}(0, 1)(x_{\max_j} - x_{\min_j}) \quad (1)$$

where x_{\min_j} and x_{\max_j} denotes lower and upper limits, respectively.

3.2 Employed bee phase

In this phase, each employed bee obtains a new solution v_{ij} from x_{ij} using expression:

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \quad (2)$$

where k is randomly obtained from $\{1, 2, \dots, SN\}$ and ϕ_{ij} is a uniform random number between $[-1, 1]$. The value of v_{ij} is obtained and compared to x_{ij} , further if fitness of v_{ij} comes out better than x_{ij} , then bee will forget the old solution and remember the new one. Otherwise, it will keep exploiting x_{ij} .

3.3 Onlooker bee phase

All employed bees share the nectar information of their food source with the onlookers, through a waggle dance performed at their hive, after which they select a food source depending on a probability p_i as:

$$p_i = \frac{f_i}{\sum_{n=1}^{SN} f_i} \quad (3)$$

where f_i is the fitness of x_i . Onlooker bee chooses a food source with higher fitness and search x_{ij} according to Eq. (2); now if the new solution has a better fitness, it will replace x_{ij} .

3.4 Scout bee phase

After a number of trials, called maximum cycle number (MCN), if a solution cannot be improved further then food source is abandoned, and the corresponding employed bee becomes a scout again. The scout will then produce a new food source randomly by using Eq. (1) again.

4 Variants of ABC

Various modifications has been proposed in the past to inculcate efficiency in the existing original version of the ABC metaheuristic. One of the factors, which affect the outcome of the metaheuristic, is influenced by its solution search equations and thus require many modifications. Advising an improved solution search equation trying to set a balance between the exploitation and exploration capabilities of the metaheuristic, Gao and Liu (2011) introduced search equations which are based on differential evolution (DE) (Ferrante Neri 2001) algorithms, to solve numerous real-world optimization problems.

$$v_{ij} = x_{bj} + \phi_{ij}(x_{ij} - x_{rj}) \quad (4)$$

$$v_{ij} = x_{r1j} + \phi_{ij}(x_{ij} - x_{r2j}) \quad (5)$$

where x_{bj} is the global best solution, $\phi_{i,j}$ is a uniform random number and $x_{r,j}$ is the j th component of a random solution. In addition, Heinzelman et al. (2002) introduced another variant later as

$$v_{ij} = x_{bj} + \phi_{ij}(x_{r1j} - x_{r2j}) \quad (6)$$

Other search equations are proposed by Abro and Mohamad-Saleh (2012) and Gao and Liu (2013)

$$v_{ij} = x_{r1j} + \phi_{ij}(x_{r2j} - x_{r3j}) \quad (7)$$

$$v_{ij} = x_{bj} + \phi_{ij}(x_{r1j} - x_{r2j}) + \psi_{ij}(x_{r3j} - x_{r4j}) \quad (8)$$

$$v_{ij} = x_{bj} + \phi_{ij}(x_{r1j} - x_{r2j}) + \psi_{ij}(x_{sbj} - x_{ij}) \quad (9)$$

where x_{sbj} is the j coefficient of the second best solution. However, the effectiveness of these DE-based equations will critically depend on the appropriate setting of population size and strategy parameters. Therefore, to obtain optimal solution, the parameters setting must be required.

These equations are further modified by Li and Niu (2013) into Eq. (10), where w_{ij} is the relative weight and θ_1, θ_2 are parameters to control step size.

$$v_{ij} = w_{ij}x_{bj} + \phi_{ij}\theta_1(x_{ij} - x_{rj}) + \psi_{ij}\theta_2(x_{bj} - x_{ij}) \quad (10)$$

Although the above-mentioned solution search equation may refine the exploitation process, using two different control parameters sometimes leads to oscillation. The search equations introduced above can be utilized in onlooker bee phase as well; however, in that case neighborhood search will be performed on most anticipating solutions with best fitness.

Scout bee phase also witnessed some improvements (Guo and Cheng 2011) in form of new solution search Eqs. (11), (12) as:

$$v_{ij} = x_{ij} + \phi_{ij}(x_{bj} - x_{ij}) \quad (11)$$

$$v_{ij} = x_{ij} + \phi_{ij}(x_{sbj} - x_{ij}) \quad (12)$$

In addition, some improved versions of the ABC metaheuristic (Akay and Karaboga 2012) include some parameters, like modification rate (MR) and scale factor (SF); MR controls the neighborhood search, whereas SF controls the length of the search. However, most of the above-mentioned works do not control the adaptation of the population and does not specify any means to improve sampling space, which is a significant measure to improve the effectiveness of ABC metaheuristic.

5 Author's contribution

The main contributions of this paper are listed as follows:

1. *Improved artificial bee colony (iABC) metaheuristic*: In an attempt to further improve the convergence rate and attain a perfect balance between exploitation and exploration capabilities of existing ABC metaheuristic, we propose an improved artificial bee colony (iABC) metaheuristic with better sampling technique using Student's t-distribution; a compact probability density function (cPDF), which requires only one control parameter to be stored on memory. Student's t-distribution is being introduced first time from the widely acclaimed family of estimation of distribution algorithms (EDAs) framework. Further, an improved solution search equation named ABC/rand-to-opt/1 is proposed, which is motivated by existing differential evolution (DE) family framework and educe a optimal solution from the current best solutions thus improving convergence rate of the proposed metaheuristic.
2. *Beecluster—an improved artificial bee colony-based clustering protocol*: Utilizing capabilities of the proposed

metaheuristic, we introduce Beecluster, an improved artificial bee colony-based clustering protocol for optimal cluster head (CH) selection, which is a well-identified NP-hard optimization problem in WSNs. Additionally, we determine the optimal position of the BS through analytical evaluation of energy equations, which reduce the energy requirements of the existing network, thus extending network life of WSNs.

6 Improved artificial bee colony (iABC) metaheuristic

Like standard ABC metaheuristic, its variants too face some challenges, like the convergence rate is typically slow since they find difficulty in choosing the most promising search solution, while solving complex multimodal optimization problems. To overcome these limitations, we propose an improved artificial bee colony (iABC) metaheuristic with an improved initialization phase for better sampling and improved solution search equation, namely ABC/rand-to-opt/1 with optimal search abilities. The details of the proposed metaheuristic are as follows:

6.1 Improved initialization phase

Population initialization is an important step in evolutionary algorithms as it can affect the convergence rate and quality of the final solution. Moreover, a large amount of the memory is needed either to store the trial solutions or control parameters of the problem. To reduce the memory requirements, the concept of virtual population has been introduced (Mininno and Cupertino 2008) through family of estimation of distribution algorithms (EDA) (Larranaga and Lozano 2001) framework by considering compact probability density functions (cPDFs). Therefore, we propose Student's t-distribution (Walck 1996); a cPDF, which needs only one vector to be stored on memory, thus reduces storage and step up convergence rate. The proposed distribution can be described by Eq. (13) where $f(x_{ij})$ is the value of the cPDF corresponding to variable x_{ij} , the $(-\infty, \infty)$ domain of the proposed cPDF is truncated to $[-1, 1]$ and B represents a Beta function. By applying this cPDF, only vector κ is needed to be stored on memory. This cPDF is being introduced first time in population-based metaheuristic due to its compact nature.

$$f(x_{ij}) = \frac{\left(1 + \frac{x_{ij}^2}{\kappa}\right)^{-(\kappa+1)/2}}{\sqrt{\kappa} B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \quad (13)$$

Further, we suggested a new alternative with cumulative distribution function (CDF) of the proposed Student's t-distribution, where a pair of cPDFs that share the same parameters are derived through Student's t-CDF by taking integral from $-x$ to x with respect to dx for function $f(x)$ as mentioned below:

$$\begin{aligned} \int_{-x}^x f(x) dx &= \frac{1}{\sqrt{\kappa} B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \int_{-x}^x \left(1 + \frac{x^2}{\kappa}\right)^{-(\kappa+1)/2} dx \\ &= \frac{2}{\sqrt{\kappa} B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \int_0^x \left(1 + \frac{x^2}{\kappa}\right)^{-(\kappa+1)/2} dx \\ &= \frac{-2}{\sqrt{\kappa} B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \int_1^{\kappa/\kappa+x^2} \frac{x^{\kappa+1/2} \kappa \sqrt{x}}{2x^2 \sqrt{\kappa} \sqrt{(1-x)}} dx \\ &= \frac{1}{B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \int_{\kappa/\kappa+x^2}^1 (1-x)^{-1/2} x^{\kappa/2-1} dx \\ &= \frac{1}{B\left(\frac{1}{2}, \frac{\kappa}{2}\right)} \left(B\left(\frac{\kappa}{2}, \frac{1}{2}\right) - B\left(\frac{\kappa}{\kappa+x^2}, \frac{1}{2}\right) \right) \\ &= 1 - I\left(\frac{\kappa}{\kappa+x^2}, \frac{1}{2}\right) = I\left(\frac{1}{\kappa+x^2}, \frac{\kappa}{2}\right) \end{aligned} \quad (14)$$

where I corresponds to incomplete Beta function. Therefore, the search space corresponding to variable x_{ij} is now divided into $[-1, 0]$ and $[0, 1]$ and instead of applying one cPDF, a pair of cPDFs $P_j(x)$ (15) and $Q_j(x)$ (16) are employed for better sampling based on a parameter ξ that controls the probability of sampling.

$$P_j(x) = \frac{1}{2} - \frac{1}{2} I\left(\frac{x_{ij}^2}{\kappa + x_{ij}^2}, \frac{1}{2}, \frac{\kappa}{2}\right) \quad \text{for } -1 < x < 0 \quad (15)$$

$$Q_j(x) = \frac{1}{2} + \frac{1}{2} I\left(\frac{x_{ij}^2}{\kappa + x_{ij}^2}, \frac{1}{2}, \frac{\kappa}{2}\right) \quad \text{for } 0 \leq x < 1 \quad (16)$$

These equations are employed to refine the sampling process which ultimately enhance the convergence rate of the proposed metaheuristic globally.

6.2 Improved solution search equation

Differential evolution (DE) (Storn R 2010) employs most powerful stochastic real-parameter algorithms to solve multimodal optimization problems with the optimal combination of population size and their associated control parameters. In other words, a well-contrive parameter adaptation approach can effectively solve various optimization problems and convergence rate can improve further if the control parameters are adjusted to appropriate values with improved solution search equations at different evolution stages of a specific problem. There are various DE variants which are different in their mutation strategies, but DE/rand-to-best/1 (Das and Sugantha 2011; Gonuguntla et al. 2015) is one of its kind which explore best solutions to direct the movement of the current population and can effectively maintain population diversity as well.

$$DE/rand - to - best/1 : \quad v_t = x_t + SF_1(x_{bes} - x_t) + SF_2(x_r - x_s) \quad (17)$$

where SF_1 and SF_2 are scaling factors for neighborhood search. Inspired by this DE variant (17) and inculcating properties of the ABC metaheuristic, we propose a new solution search equations ABC/rand-to-opt/1 as follows:

$$ABC/rand - to - opt/1 : \quad v_{ij} = x_{ij} + \phi_{ij}(x_{opt,j} - x_{ij}) + \psi_{ij}(x_{r1j} - x_{r2j}) \quad (18)$$

where r_1, r_2 are random variables from $1, 2, \dots, SN$, x_{opt} is the optimal individual solution with optimal fitness in the current population with ϕ_{ij} and ψ_{ij} are scaling factors, respectively.

The proposed solution search equation ABC/rand-to-opt/1, which utilizes the information of only optimal solutions in the current population, can improve the convergence rate of the proposed metaheuristic.

To increase the multifariousness of the population further, a crossover operation is performed as :

$$u_{ij} = \begin{cases} v_{ij} & \text{if } r[0, 1] \leq CR, \\ x_{opt,j} & \text{otherwise} \end{cases} \quad (19)$$

Then, a selection operation will be performed as:

$$x_{i,j} = \begin{cases} u_{ij} & \text{if } f(u_{ij}) \leq f(x_{ij}), \\ x_{opt,j} & \text{otherwise} \end{cases} \quad (20)$$

where $f(x_{ij})$ is the fitness function, if the new solution seems to have high fitness value, then it replaces the corresponding old solution; otherwise, the old solution is retained in the memory. Therefore, with the proposed improved solution

Table 2 Benchmark functions used in experiment

Function	Search range
$f_1(x) = \sum_{i=1}^n x_i^2$	$[-100, 100]^n$
$f_2(x) = \sum_{i=1}^n ix_i^2$	$[-10, 10]^n$
$f_3(x) = \sum_{i=1}^n x_i ^{i+1}$	$[-10, 10]^n$
$f_4(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	$[-10, 10]^n$
$f_5(x) = [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$[-6.45, 6.45]^n$
$f_6(x) = [y_i^2 - 10 \cos(2\pi y_i) + 10]$	$[-6.45, 6.45]^n$
where $y_i = \begin{cases} x_i & \text{if } x_i < \frac{1}{2} \\ \frac{\text{round}(2x_i)}{2} & \text{if } x_i \geq \frac{1}{2} \end{cases}$	$[-6.45, 6.45]^n$
$f_7(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	$[-500, 500]^n$
$f_8(x) = -\sum_{i=1}^n \sin(x_i) \sin^{20}(\frac{i \times x_i^2}{\pi})$	$[0, \pi]^n$

search equation, optimal solution is obtained with optimal exploration and exploitation ability thus contribute to a better convergence rate.

We have evaluated the convergence rate of our proposed iABC metaheuristic with the standard ABC metaheuristic using set of eight scalable benchmark functions f_1 to f_8 , where functions f_1 to f_4 are unimodal and functions f_5 to f_8 are multimodal functions as shown in Table 2:

Figures 1 and 2 show that the proposed iABC metaheuristic convergence fast with optimal or closer-to-optimal solutions on the unimodal as well as complex multimodal functions over to its standard ABC variant. Therefore, the proposed metaheuristic can improve searching abilities, increase convergence rate and possess more computational efficiency.

7 Beecluster-proposed clustering protocol

We inherit the capabilities of our proposed metaheuristic to solve well-known NP-hard optimization problem of energy-efficient clustering in WSNs by proposing Beecluster, an improved artificial bee colony-based clustering protocol with an optimal CH selection ability. Additionally, we also determine the optimal location of the BS through analytical evaluation of energy equations, which reduce the energy consumption of the network and help to enhance the network life of existing WSN.

7.1 Notations for network model

Following notations are used for network model in our proposed work:

1. S is set of sensor nodes $S = \{s_1, s_2, \dots, s_n\}$, which are randomly distributed over a geographical area of defined dimensions $m \times m$, whereas s_{n+1} denotes the BS. Each sensor node has a communication radius r .
2. L is set of bidirectional wireless links between two sensor nodes, where $l_{i,j} \in L$ represents wireless link between node s_i and s_j .
3. Set of cluster heads (CH's) are denoted by $S_{ch} = \{ch_1, ch_2, \dots, ch_k\}$ where S_{ch} is included in S .
4. $D_{s_i}^{s_j}$ denotes the maximum distance between a sensor node s_i and s_j is calculated by squared Euclidean distance between them as

$$D_{s_i}^{s_j} = \text{dis}(s_i, s_j) \quad | \quad \forall s_i, s_j \in S$$

$$D_{s_i}^{s_j} = \|s_i - s_j\|^2 = \sum (s_i - s_j)^2 \quad | \quad \forall s_i, s_j \in S \quad (21)$$

5. $D_{s_i}^{s_{n+1}}$ denotes the maximum distance between a sensor node s_i and BS, calculated by squared Euclidean distance between them as

$$D_{s_i}^{s_{n+1}} = \text{dis}(s_i, s_{n+1}) \quad | \quad \forall s_i \in S$$

$$D_{s_i}^{s_{n+1}} = \|s_i - s_{n+1}\|^2 = \sum (s_i - s_{n+1})^2 \quad | \quad \forall s_i \in S \quad (22)$$

6. $D_{s_i}^{ch_j}$ denotes the maximum distance between a sensor node s_i and cluster head ch_j , calculated by squared Euclidean distance between them as

$$D_{s_i}^{ch_j} = \text{dis}(s_i, ch_j) \quad | \quad \forall s_i, ch_j \in S$$

$$D_{s_i}^{ch_j} = \|s_i - ch_j\|^2 = \sum (s_i - ch_j)^2 \quad | \quad \forall s_i, ch_j \in S \quad (23)$$

7. $D_{ch_j}^{s_{n+1}}$ represents the maximum distance between a cluster head ch_j and BS, calculated by squared Euclidean distance between them as

$$D_{ch_j}^{s_{n+1}} = \text{dis}(ch_j, s_{n+1}) \quad | \quad \forall j \in S_{ch}$$

$$D_{ch_j}^{s_{n+1}} = \|ch_j - s_{n+1}\|^2 = \sum (ch_j - s_{n+1})^2 \quad | \quad \forall j \in S_{ch} \quad (24)$$

8. Transmission power of a sensor node s_i is calculated as:

$$P_{\text{tran}_i} = \frac{1}{k} \left(\frac{\gamma \cdot T_{\text{delay}}}{dc} \right)^\alpha \quad (25)$$

where T_{delay} is the sum of three delay components as

$$T_{\text{delay}} = T_{\text{que}} + T_{\text{tran}} + T_{\text{ack}} \quad (26)$$

The first component, T_{que} , is the queuing delay, second component, T_{tran} , is the transmission delay and third one,

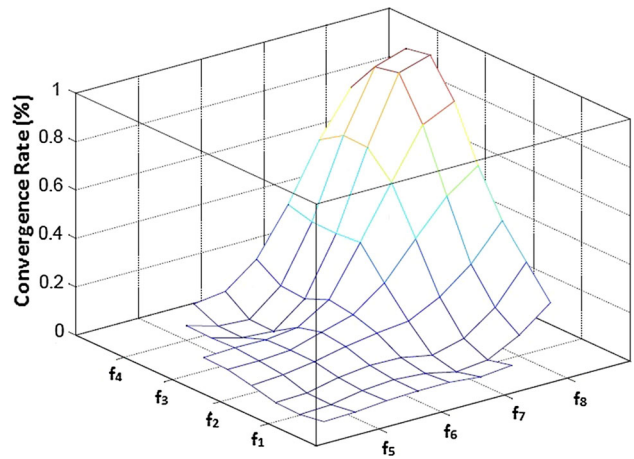


Fig. 1 Convergence rate of iABC

T_{ack} , is delay due to acknowledgement packet. γ , α , dc are signal-to-noise ratio, path loss exponent and delay constraint, respectively, whereas k is a power constant.

7.2 Optimal base station (BS) position

Placement of BS at appropriate location is an all important task in WSNs as it affect its performance by reducing energy consumption and enhancing lifetime of the network. Therefore, we present an analytical evaluation to determine the optimal location of the BS based on critical energy analysis of WSNs. The radio model is assumed to be same which has been used in earlier work (Heinzelman et al. 2002), in which CHs receive data packets from SNs for aggregation but include an additional acknowledgment packet (ACK) in return to the source node after receiving a correct packet. This is first time we incorporate the significance of energy consumption by the exchange of a ACK in WSNs. The radio hardware, which include a transmitter, dissipates energy to run transmitter radio electronics and power amplifier, whereas the receiver dissipates energy to run the receive radio electronics as shown in Fig. 3:

Therefore, the energy consumption for transmission of l bits of data is composed of three parts: the energy consumed by the transmitter E_{trans} , by the receiver E_{rec} and by the ACK packet exchange E_{ack} .

$$E_{\text{total}}(l, d) = E_{\text{trans}}(l, d) + E_{\text{rec}}(l, d) + E_{\text{ack}} \quad (27)$$

Now, energy consumed for transmitting l bits of data is given by:

$$E_{\text{trans}}(l, d) = l \cdot E_{\text{elec}} + E_{\text{amp}}(l, d) \quad (28)$$

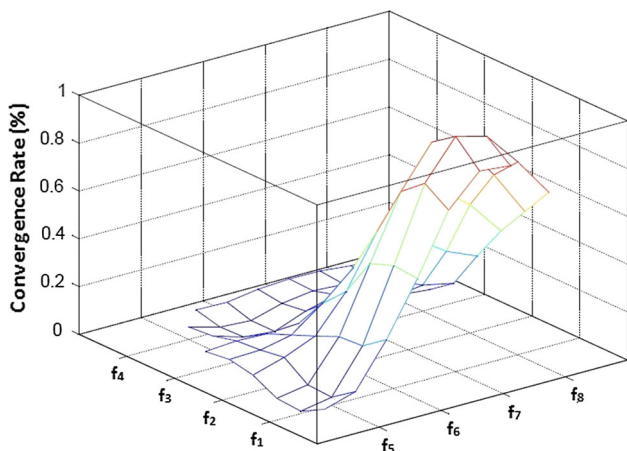


Fig. 2 Convergence rate of ABC

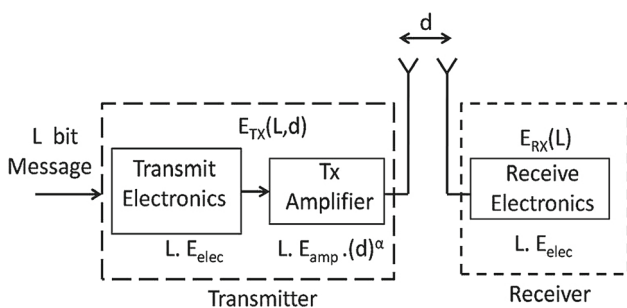


Fig. 3 Radio model for energy analysis

Further, if the distance between transmitter, and receiver is d , then

$$E_{trans}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2 & \text{if } d < d_0, \\ lE_{elec} + l\epsilon_{mp}d^4 & \text{if } d \geq d_0, \end{cases} \tag{29}$$

where d_0 is the threshold distance.

And to receive l bit message, the radio spends $E_{rec}(l, d)$ as

$$E_{rec}(l, d) = l.E_{elec} \tag{30}$$

Energy consumed for ACK packet exchange is given by

$$E_{ack} = \tau_{ack}(E_{trans} + E_{rec}) \tag{31}$$

where $\tau_{ack} = \frac{l_{ack}}{l}$ is the ratio of length of acknowledgement packet to data packet.

Therefore, residual energy of each sensor node is calculated as:

$$E_{res} = E_{total} - (E_{trans} + E_{rec} + E_{ack}) \tag{32}$$

If there will be n nodes uniformly distributed in an $m * m$ field with k clusters, then there will be $\frac{n}{k}$ nodes per cluster. Out of these, there will be one CH node and remaining $\frac{n}{k} - 1$ non-CH nodes.

Now, energy consumed by a non-CH node is given by:

$$E_{non-ch}(l, d) = E_{trans}(l, d) \tag{33}$$

$$= l.E_{elec} + E_{amp}(l, d) \tag{34}$$

and energy consumed by a CH node is given by :

$$E_{ch}(l, d) = E_{trans}(l, d) + \left(\frac{n}{k} - 1\right) .l.E_{elec} + \frac{n}{k}.l.E_{da} + \left(\frac{n}{k} - 1\right) E_{ack} \tag{35}$$

where E_{da} is the energy consumed by CH for data aggregation at its end. Now, the total energy consumed in a cluster is given by :

$$E_{cluster} = E_{ch}(l, d) + \left(\frac{n}{k} - 1\right) E_{non-ch} \tag{36}$$

Therefore, energy consumed in whole network per round

$$E_{round} = \sum_{j=1}^k E_{cluster}(j) \tag{37}$$

Lemma 1 $\forall \{s_1, s_2, \dots, s_n\} \in S$ near to BS, the total energy consumption for transmission of 1 bits of data from nodes to BS will be a function of $\sum_{j=1}^n (ch_j - s_{n+1})^2$ and optimal BS position will be located at centroid of $\{s_1, s_2, \dots, s_n\} \in S$.

Proof Let $\{s_1, s_2, \dots, s_n\} \in S$ are randomly deployed at $M = m \times m$ sensor field, within direct communication range to BS.

From Eqs. (33), (35), (36), Eq. (37) can be written as :

$$E_{round} = l.(2n - k).E_{elec} + n.l.E_{da} + \sum_{i=1}^n E_{amp}(l.D_{s_i}^{ch_j}) + \sum_{j=1}^k E_{amp}(l.D_{ch_j}^{s_{n+1}}) + \left(\frac{n}{k} - 1\right) E_{ack} \tag{38}$$

Signals from all the nodes suffer free space loss while transmitting data; therefore, we have:

$$\sum_{i=1}^n E_{amp}(l.D_{s_i}^{ch_j}) = l.(n - k)\epsilon_{fs} \left(\frac{M^2}{2\pi k}\right) \tag{39}$$

So, Eq. (38) becomes:

$$\begin{aligned}
 E_{\text{round}} &= l \left((2n - k) \cdot E_{\text{elec}} + n \cdot E_{\text{da}} + (n - k) \epsilon_{\text{fs}} \left(\frac{M^2}{2\pi k} \right) \right. \\
 &\quad \left. + \left(\frac{n}{k} - 1 \right) E_{\text{ack}} + \epsilon_{\text{fs}} \sum_{j=1}^k (D_{\text{ch}_j}^{s_{n+1}})^2 \right) \quad (40)
 \end{aligned}$$

After n/k rounds, total energy consumption for transmission of l bits of data from nodes to BS, $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ will be given as:

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &= l \cdot \frac{n}{k} \left((2n - k) \cdot E_{\text{elec}} + n \cdot E_{\text{da}} + (n - k) \epsilon_{\text{fs}} \left(\frac{M^2}{2\pi k} \right) + E_{\text{ack}} \right) \\
 &\quad + l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (D_{\text{ch}_j}^{s_{n+1}})^2 \quad (41)
 \end{aligned}$$

Therefore,

$$E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) \propto l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (D_{\text{ch}_j}^{s_{n+1}})^2 \quad (42)$$

thus, from Eq. (24):

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &\propto \text{Min} \left(l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (\text{ch}_j - s_{n+1})^2 \right) \quad | \quad \forall j \in S_{\text{ch}} \quad (43)
 \end{aligned}$$

Further, to obtain optimal location of BS, we need to minimize $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ and sum of square of Euclidean distances will be minimum at the centroid (R Apostol 2003). Thus, $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ will be minimum if BS will be placed at centroid of $\{s_1, s_2, \dots, s_n\} \in S$.

Lemma 2 $\forall \{s_1, s_2, \dots, s_n\} \in S$ far from BS, the total energy consumption for transmission of 1 bits of data from nodes to BS will be a function of $\sum_{j=1}^n (\text{ch}_j - s_{n+1})^4$ and optimal BS position will be located at centroid of $\{s_1, s_2, \dots, s_n\} \in S$.

Proof As $\{s_1, s_2, \dots, s_n\} \in S$ are deployed at $M = m \times m$ sensor field and are far from the direct communication range to BS, after n/k rounds, total energy consumption for transmission of l bits of data from nodes to BS, $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ will be given by:

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &= l \cdot \frac{n}{k} \left((2n - k) \cdot E_{\text{elec}} + n \cdot E_{\text{da}} + (n - k) \epsilon_{\text{fs}} \left(\frac{M^2}{2\pi k} \right) + E_{\text{ack}} \right) \\
 &\quad + l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (D_{\text{ch}_j}^{s_{n+1}})^4 \quad (44)
 \end{aligned}$$

Therefore,

$$E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) \propto l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (D_{\text{ch}_j}^{s_{n+1}})^4 \quad (45)$$

Thus, from Eq. (24):

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &\propto \text{Min} \left(l \cdot \epsilon_{\text{fs}} \sum_{j=1}^n (\text{ch}_j - s_{n+1})^4 \right) \quad | \quad \forall j \in S_{\text{ch}} \quad (46)
 \end{aligned}$$

Further, to obtain optimal location of BS, we need to minimize $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ and polynomial of Euclidean distances will be minimum at the centroid only (Chen 1984). Thus, $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ will be minimum if BS will be placed at centroid of $\{s_1, s_2, \dots, s_n\} \in S$.

Lemma 3 $\forall \{s_1, s_2, \dots, s_a\} \in S$ near to BS and $\forall \{s_1, s_2, \dots, s_b\} \in S$ far from BS, the total energy consumption for transmission of 1 bits of data from nodes to BS, will be a function of $\left(\sum_{j=1}^n (\text{ch}_j - s_{n+1})^2 + \sum_{j=1}^n (\text{ch}_j - s_{n+1})^4 \right)$ and optimal BS position will be located at centroid of $\{s_1, s_2, \dots, s_a\}$ and $\{s_1, s_2, \dots, s_b\} \in S$, respectively.

Proof As $\{s_1, s_2, \dots, s_a\} \in S$ are deployed near to BS and $\{s_1, s_2, \dots, s_b\} \in S$ are far from BS at $M = m \times m$ sensor field, after n/k rounds, total energy consumption for transmission of l bits of data from nodes to BS, $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$ will be given by:

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &= l \cdot \frac{n}{k} \left((2n - k) \cdot E_{\text{elec}} + n \cdot E_{\text{da}} + (n - k) \epsilon_{\text{fs}} \left(\frac{M^2}{2\pi k} \right) + E_{\text{ack}} \right) \\
 &\quad + l \cdot \epsilon_{\text{fs}} \sum_{i=1}^x (D_{\text{ch}_i}^{s_{n+1}})^2 + l \cdot \epsilon_{\text{fs}} \sum_{j=1}^y (D_{\text{ch}_j}^{s_{n+1}})^4 \quad (47)
 \end{aligned}$$

$$\begin{aligned}
 E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) &= l \cdot \frac{n}{k} \left((2n - k) \cdot E_{\text{elec}} + n \cdot E_{\text{da}} + (n - k) \epsilon_{\text{fs}} \left(\frac{M^2}{2\pi k} \right) + E_{\text{ack}} \right) \\
 &\quad + l \cdot \epsilon_{\text{fs}} \left(\sum_{i=1}^x (D_{\text{ch}_i}^{s_{n+1}})^2 + \sum_{j=1}^y (D_{\text{ch}_j}^{s_{n+1}})^4 \right) \quad (48)
 \end{aligned}$$

Therefore,

$$E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) \propto l \cdot \epsilon_{\text{fs}} \left(\sum_{i=1}^x (D_{\text{ch}_i}^{s_{n+1}})^2 + \sum_{j=1}^y (D_{\text{ch}_j}^{s_{n+1}})^4 \right) \quad (49)$$

Thus, from Eq. (24):

$$E_{\text{total}}(l, D_{s_i}^{s_{n+1}}) \propto \text{Min} \left(l \cdot \epsilon_{fs} \left(\sum_{j=1}^n (\text{ch}_j - s_{n+1})^2 + \sum_{j=1}^n (\text{ch}_j - s_{n+1})^4 \right) \right) \quad | \quad \forall j \in S_{\text{ch}} \quad (50)$$

Now, to obtain optimal location of BS, we need to minimize $E_{\text{total}}(l, D_{s_i}^{s_{n+1}})$, that depends on which type of nodes are dominating in residual energy. If nearby nodes are dominating, i.e., $E_{\text{res}_a} \gg E_{\text{res}_b}$ then centroid of $\{s_1, s_2, \dots, s_a\} \in S$ nodes will be the optimal location for BS, and if $E_{\text{res}_b} \gg E_{\text{res}_a}$, then centroid of $\{s_1, s_2, \dots, s_b\} \in S$ nodes will be the optimal location for BS.

7.3 Optimal cluster head (CH) selection

CH selection is one of the crucial task for cluster formation in WSNs as it effects the overall performance of the network. CH will be responsible for collection of data coming from various SNs and transmission of aggregated data to the BS. Selection of appropriate node as a CH will remain a challenging multimodal optimization problem. Therefore, we propose a optimal CH selection algorithm based on our proposed *iABC* metaheuristic for an improved energy-efficient clustering protocol. The working of proposed algorithm is as follows:

7.3.1 Initialization phase

The population number (PN) and corresponding food sources (SN) are initialized along with control parameters maximum cycle number (MCN), control parameter ξ and crossover rate (CR).

We employ the proposed improved sampling technique of *iABC* metaheuristic to generate *i*th food source x_{ij} , for which we generate $r \in [0, 1]$ according to uniform distribution and obtain x_{ij} as :

$$x_{ij} = \begin{cases} \frac{1}{2} - \frac{1}{2}I \frac{x_{ij}^2}{\kappa + x_{ij}^2} \left(\frac{1}{2}, \frac{\kappa}{2} \right) & \text{if } r \leq \xi, \\ \frac{1}{2} + \frac{1}{2}I \frac{x_{ij}^2}{\kappa + x_{ij}^2} \left(\frac{1}{2}, \frac{\kappa}{2} \right) & r > \xi, \end{cases} \quad (51)$$

7.3.2 Fitness function derivation

Now, we construct a fitness function to evaluate the fitness of individual food source of the population. There are three

objectives in our proposed CH selection algorithm, firstly the node elected as CH will have maximum residual energy, i.e.,

$$f_i \propto \text{Max}(E_{\text{res}}) \quad (52)$$

Secondly, we ensure to minimize the maximum distance between elected node as CH and BS and finally with minimum transmission power to transmit aggregated data from CH to BS:

$$f_i \propto \frac{1}{\text{Min}(D_{\text{ch}_j}^{s_{n+1}}(\text{max}) + P_{\text{tran}_i})} \quad (53)$$

Aggregating Eqs. (52), (53) as

$$f_i \propto \frac{\text{Max}(E_{\text{res}})}{\text{Min}(D_{\text{ch}_j}^{s_{n+1}}(\text{max}) + P_{\text{tran}_i})} \quad (54)$$

$$f_i = K \frac{\text{Max}(E_{\text{res}})}{\text{Min}(D_{\text{ch}_j}^{s_{n+1}}(\text{max}) + P_{\text{tran}_i})} \quad (55)$$

where K is constant of proportionality, assuming $K=1$,

$$f_i = \frac{\text{Max}(E_{\text{res}})}{\text{Min}(D_{\text{ch}_j}^{s_{n+1}}(\text{max}) + P_{\text{tran}_i})} \quad (56)$$

Therefore, Eq. (56) will determine the fitness value of each solution of population.

7.3.3 Employed bee phase

Now, each employed bee select a new solution v_{ij} using proposed improved search Eq. (19) of proposed *iABC* metaheuristic as:

$$v_{ij} = x_{ij} + \phi_{ij}(x_{\text{opt},j} - x_{ij}) + \psi_{ij}(x_{r1j} - x_{r2j}) \quad (57)$$

The obtained value of v_{ij} is compared to x_{ij} and if the fitness of v_{ij} comes out better than x_{ij} , the bee will forget the previous old solution and retain the new optimal solution $x_{\text{opt},j}$ found so far; otherwise, it will keep working on x_{ij} .

7.3.4 Onlooker bee phase

Now, employee bee will share the information of their food source with the onlooker bee, through a waggle dance performed at their hive, each of whom will then generate a food source u_{ij} according to distribution as:

$$u_{ij} = \begin{cases} v_{ij} & \text{if } r[0, 1] \leq \text{CR}, \\ x_{\text{opt},j} & \text{otherwise} \end{cases} \quad (58)$$

where CR is crossover rate, further fitness of generated food source $f(u_{ij})$ is calculated and compared with previous food

source as:

$$x_{i,j} = \begin{cases} u_{ij} & \text{if } f(u_{ij}) \leq f(x_{ij}), \\ x_{opt,j} & \text{otherwise} \end{cases} \quad (59)$$

where $f(x_{ij})$ is the fitness value of x_{ij} . Onlooker bee will then choose a food source with higher fitness and conduct a local search on x_{ij} , if the new solution has a better fitness, then it will replace x_{ij} with optimal solution $x_{opt,j}$ and assigned as a CH; otherwise, the old solution will be retained.

7.3.5 Scout bee phase

Now, if the fitness cannot improve further, after a number of trials then the corresponding employed bee becomes a scout to produce a new food source randomly by using Eq. (51) again.

Details of the cluster head (CH) selection algorithm is discussed as below:

Optimal Cluster Head(CH) Selection Algorithm

Input:

$PN \leftarrow$ Population number,
 $MCN \leftarrow$ Maximum cycle number,
 $D \leftarrow$ Dimension of vector to be optimized,
 $SN \leftarrow$ Food sources,
 $x_{Min} \leftarrow$ Lower bound of each element,
 $x_{max} \leftarrow$ Upper bound of each element,
 $\xi \leftarrow$ Control parameter,
 $CR \leftarrow$ Crossover rate.

Output:

$Ch_j \leftarrow x_{opt,j}$

begin

$round \leftarrow 0$

for $i = 1 \rightarrow SN$ **do**

Generate $r \in [0, 1]$ according to uniform distribution.

if $r \leq \xi$ **then**

Generate $x_{ij} \in [-1, 0]$ according to PDF $P_j(x)$.

else

Generate $x_{ij} \in [0, 1]$ according to PDF $Q_j(x)$.

Evaluate fitness $f_i(x_{ij})$

$trial(s) \leftarrow 0$

$round ++$

end if

end for

repeat

until

for $i = 1 \rightarrow SN$ **do**

Generate v_{ij} according to Eq.(57)

Evaluate fitness $f_i(v_{ij})$

$round ++$

if $f_i(x_{ij}) < f_i(v_{ij})$ **then**

$x_{ij} \leftarrow v_{ij}$

$f_i(x_{ij}) \leftarrow f_i(v_{ij})$

$trial(s) \leftarrow 0$

else

$trial(s) \leftarrow trial(s) + 1$

end if

end for

if $round == MCN$ **then**

Memorize the optimal solution, $x_{opt,j}$ achieved so far and exit repeat.

$Ch_j \leftarrow x_{opt,j}$

end if

repeat

until

for $i = 1 \rightarrow SN$ **do**

$r \leftarrow rand[0, 1]$

if $r \leq CR$ **then**

$u_{ij} \leftarrow v_{ij}$

else

$u_{ij} \leftarrow x_{opt,j}$

end if

Evaluate fitness $f_i(u_{ij})$ and $f_i(x_{opt,j})$

if $f_i(u_{ij}) \leq f_i(x_{opt,j})$ **then**

$x_{i,j} \leftarrow u_{ij}$

if $f_i(u_{ij}) > f_i(x_{opt,j})$ **then**

$x_{opt,j} \leftarrow u_{ij}$

$f_i(x_{opt,j}) \leftarrow f_i(u_{ij})$

$trial(s) \leftarrow trial(s) + 1$

end if

end if

if solution need to be abandoned

replace with a new solution, produced using

Eq. (51)

$round ++$

end for

if $round == MCN$ **then**

Memorize the optimal solution, $x_{opt,j}$ achieved.

$Ch_j \leftarrow x_{opt,j}$

end if

end

To evaluate and compare the performance of our proposed clustering algorithm based on iABC with standard ABC metaheuristic, we performed simulation for 10,000 rounds and obtained number of CHs per round.

Figures 4, 5 show that the proposed metaheuristic clearly supports more number of CHs per round compared to its peer, which in turns is contributed to its optimal search ability and better convergence rate. The proposed improved metaheuristic also supported network for larger number of rounds, compared to standard metaheuristic, which is attributed to Student's t-distribution's compact nature that saved energy to support network for longer duration.

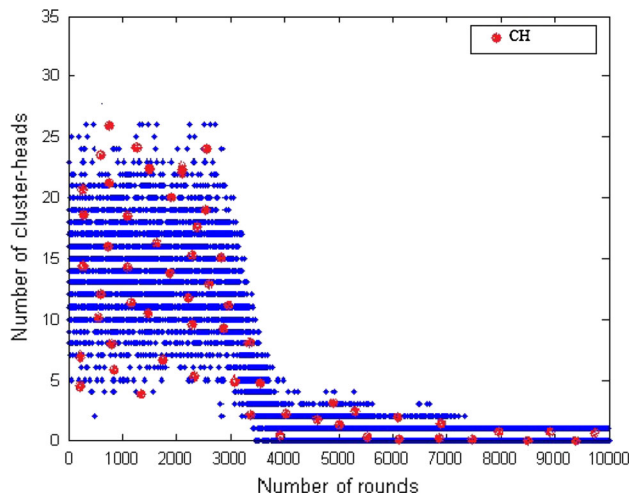


Fig. 4 CHs selected per round in $iABC^2$

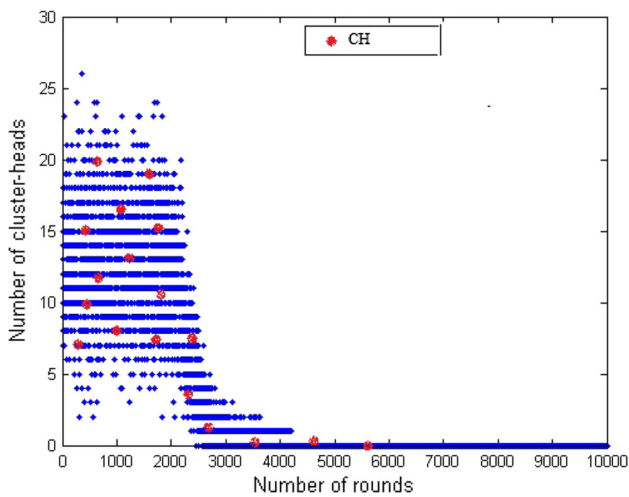


Fig. 5 CHs selected per round in ABC

7.4 Cluster formation phase

After selection of CHs, each CH will advertise a Join-Request (J-REQ) message to all its neighbor nodes for cluster formation. Then, each non-CH node will join the nearest CH node based on squared Euclidean distance between them (Eq. 24), through a Join-Acknowledgment (J-ACK) short message which will be transmitted using a CSMA/CD MAC protocol, to become member of the cluster. During this communication, all CH nodes must keep their receivers on and listen to the channel. If a particular node receives multiple J-REQ message from same CH, then it discards the message to eliminate duplicate frames. After receiving J-ACK messages from all the surrounding nodes, each CH must maintain a cluster member table and create a TDMA schedule for each member node of the cluster for data transmission. During cluster formation, it is ensured that each non-CH node must join a cluster under a CH to avoid node isolation.

Table 3 Simulation Parameters

Parameter	Value
Terrain size	$200 \times 200 \text{ m}^2$
MAC protocol	802.11
Radio propagation	Free space
Traffic type	CBR
ϵ_{fs}	6 pJ/bit/m
ϵ_{mp}	0.0011 pJ/bit/m ⁴
Propagation limit	-111 dBm
Receiver sensitivity	-89
Data rate	3 Mbps
Packet size	4000 bits
Message size	500 bits

7.5 Data transmission phase

After cluster formation, when TDMA schedule is communicated to each member node for data transmission, SNs collect data and transmit it to their CH during their allocated TDMA schedule. The *non-CH* nodes can turn their radio transmitter off during other members transmission turn to save energy consumption. After receiving all the data, CH nodes aggregate it at its end using data aggregation algorithms and route the aggregated data packets to the BS.

8 Simulation results and discussion

Now, we evaluate the performance of proposed Beecluster protocol with existing PSO (Kuila and Jana 2014), MRP (Yang et al. 2009) and ERP (Attea and Khalil 2012) protocols using ns-2 simulator. The protocols are simulated over two different BS position scenarios to assess their behavior toward packet delivery ratio, throughput, energy consumption, network lifetime and average latency. The simulation will be performed over standard MAC protocol with free space radio propagation and CBR traffic type, considering other parameters as shown in Table 3 below:

In the first scenario WSN # 1, a network of sensor nodes ranging from 100 to 700 are deployed randomly over an area of size $200 \times 200 \text{ m}^2$ with a BS, located at optimal position calculated through proposed evaluation of energy Eq. (7.2) within the network field, whereas in second scenario WSN # 2, a BS will be placed at position (100, 300 m) outside the network field. First, we execute the protocols to compare packet delivery ratio (PDR) in the network for both the scenarios.

In scenario WSN # 1, Fig. 6 shows that the proposed protocol delivers highest number of packets among its all peers, even at highest density of nodes. Beecluster deliver approximately 100% packets at 100 nodes with BS located at optimal

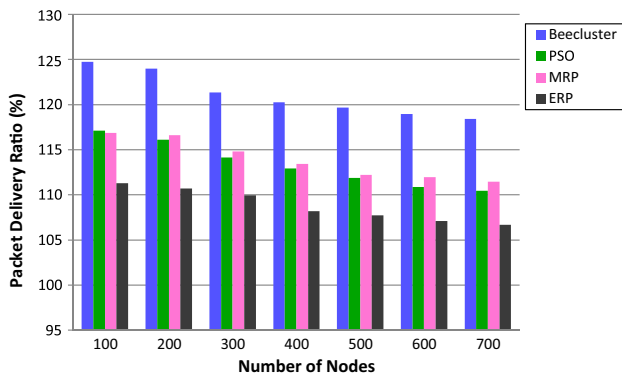


Fig. 6 Packet delivery ratio in WSN # 1

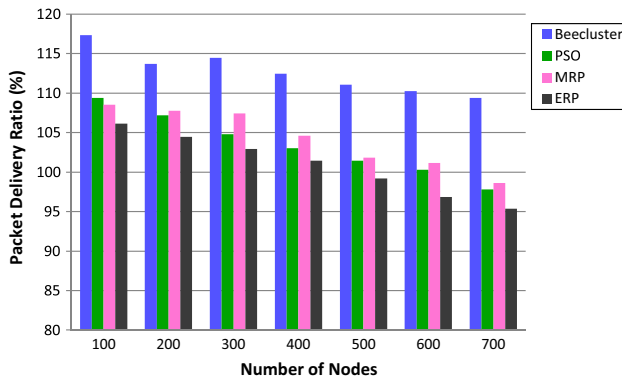


Fig. 7 Packet delivery ratio in WSN # 2

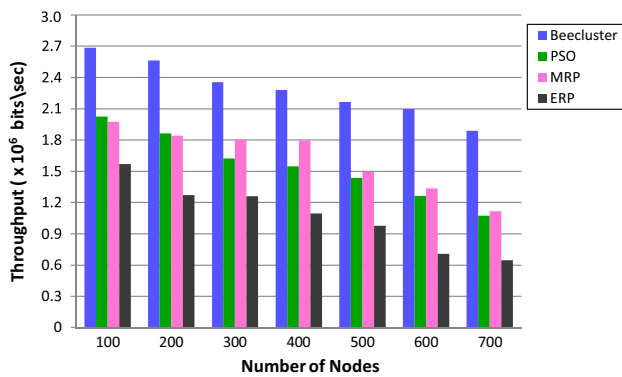


Fig. 8 Throughput in WSN # 1

position in WSN # 1 scenario. Even in WSN # 2 scenario, Fig. 7 shows that Beecluster has highest PDR as compared to PSO, MRP and ERP at nodes ranging from 100 to 700. It is further analyzed from above Figs. 6 and 7 that PDR remain highest, at optimal BS position, i.e., in WSN # 1 scenario which clearly shows the significance of placing BS at optimal position, as manual intervention is possible for BS position while deployment of nodes in a WSN.

Throughput is a measure of robustness for any protocol, and Figs. 8 and 9 show that Beecluster is delivering highest number of packets per second even at 700 nodes deployed in

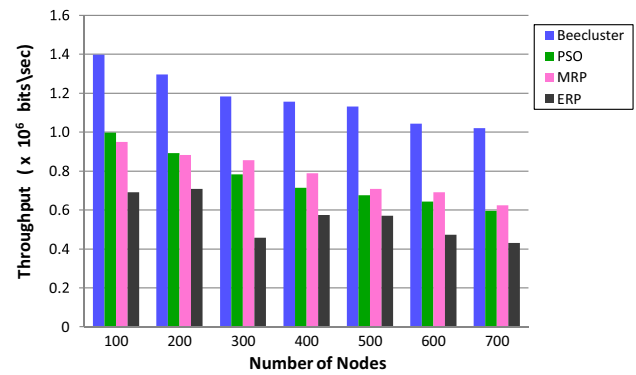


Fig. 9 Throughput in WSN # 2

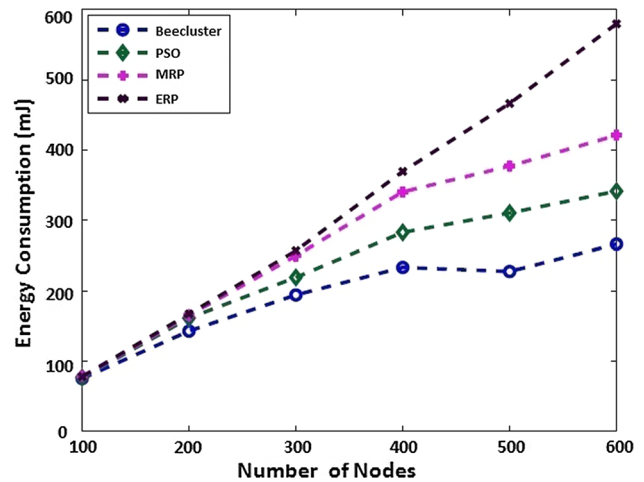


Fig. 10 Energy consumption in WSN # 1

the both scenarios. The low performance of PSO and ERP is due to the fact that they are non-cluster-based protocols and hence lack in performance while delivering data from nodes to BS. Additionally, it is observed that optimal BS position scenario WSN # 1 Fig. 8, witness highest throughput compared to WSN # 2 Fig. 9, thus placing BS at optimal position will increase data delivery per second in WSNs.

Figure 10 shows that in scenario WSN # 1, energy consumption of the proposed protocol is approximately 32, 73, 131% less than PSO, MRP and ERP protocols, respectively, which is attributed to the use of compact Student's t-distribution and improved solution search equation to selects optimal CHs, thus minimizing energy consumption in the network. Even in scenario WSN #2, Fig. 11, Beecluster consumes 41% less energy as compared to its contender PSO, which clearly shows the effectiveness of the proposed meta-heuristic iABC. In Beecluster, optimal CHs are selected not only based on their proximity to BS but with the condition of minimum power consumption in data transmission; moreover, the SNs are assigned to their nearest CH, thus consuming less energy and as a result the overall energy consumption of the network becomes lesser than other protocols.

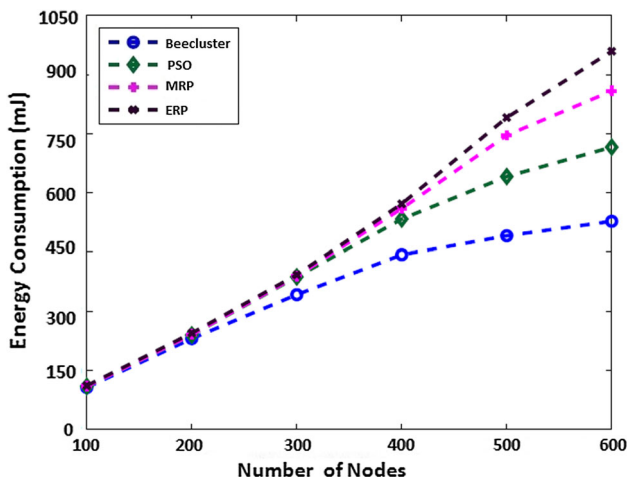


Fig. 11 Energy consumption in WSN #2

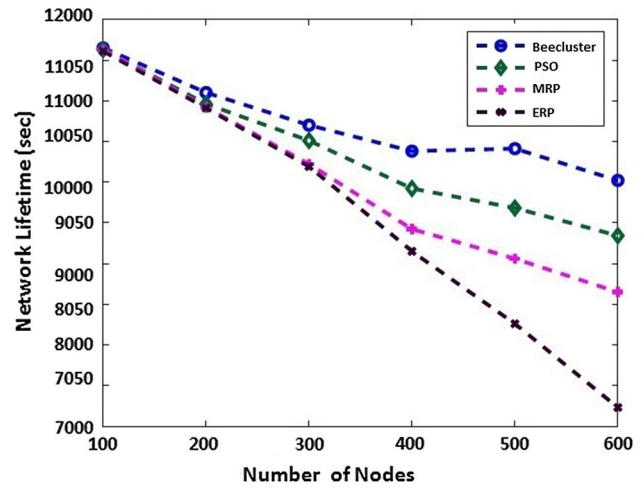


Fig. 13 Network lifetime in WSN #2

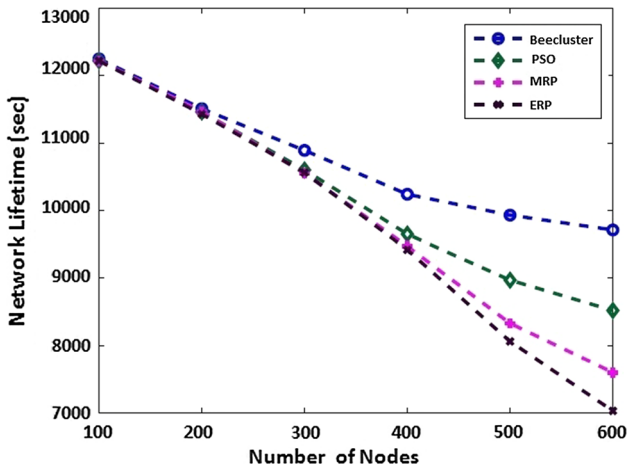


Fig. 12 Network lifetime in WSN #1

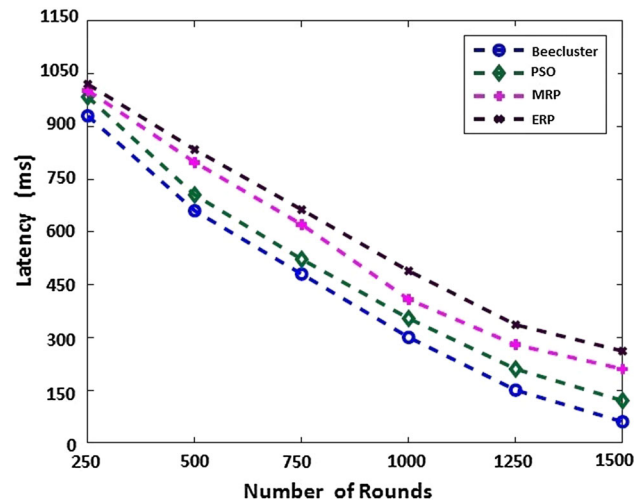


Fig. 14 Average latency in WSN #1

In MRP, all CHs are inevitably used as a relay node to forward the data packets to the BS; therefore, consume more energy whereas PSO and ERP are non-cluster-based protocols and are not able to optimize the energy usage in complex WSN scenario. Moreover, placement of BS at optimal location will decrease the energy consumption as evident from Figs. 10 and 11, respectively.

Figures 12 and 13 shows that Beecluster extend the average network lifetime by approximately 21 and 18 % compared to PSO and MRP in WSN #1 and WSN #2, respectively, which is the effect of nodes surplus energy availability due to less computation and an optimal selection of CHs with proposed metaheuristic.

The energy thus saved will prolong the network lifetime, and the nodes will be able to transmit data for a longer duration. In MRP, due to un-symmetric data forwarding effects on the CHs, those near to the BS will die quickly thus reducing network lifetime. ERP is having the least network lifetime among all its peers, due to absence of a clear data aggregation

and communication framework, specially for WSN #2 like scenarios. It is further analyzed that every 1 % increase in network lifetime of the proposed protocol will increase the data delivery by 3.8% , thus increasing networks robustness.

Figures 14 and 15 compare the average latency in both scenarios after number of rounds ranging from 250 to 1500. It is clearly visible that Beecluster deliver data packets with minimum latency in both the scenarios among other protocols which ultimately increases reliability of the network. In WSN #1, average latency decreases sharply with increase in number of rounds in Beecluster, which is due to the fact that the proposed protocol deliver data packets to the BS with minimum relay after calculating the optimal possible distance for the next hop; moreover, the CHs are placed at optimal distance to BS, thus maintaining a trade-off between transmission distance and hop count. Also in WSN #2, when the BS is located at a far distance from sensor nodes, the proposed protocol will be able to deliver the data packets with minimum delay successfully. In PSO and MRP, data will be

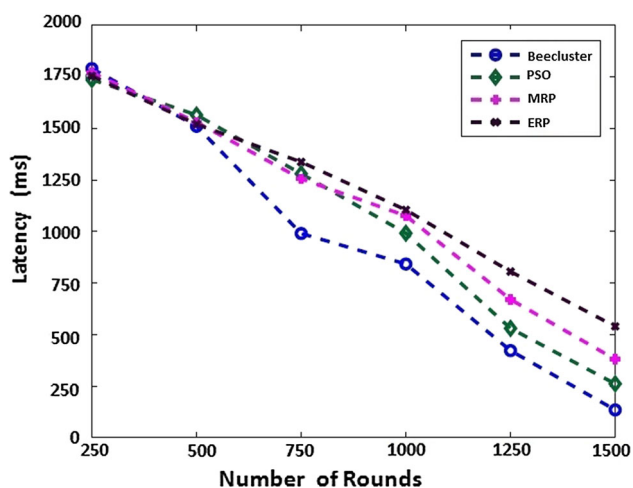


Fig. 15 Average latency in WSN # 2

transmitted to BS using maximum number of hop count ultimately exhausting the network with unnecessary end-to-end delay.

9 Conclusion

This paper presents an iABC metaheuristic, based on first-of-its-kind Student's t-cPDF and DE-inspired improved solution search equation ABC/rand-to-opt/1, to improve exploitation capabilities as well as convergence rate of existing ABC metaheuristic. To demonstrate its improvement, we evaluated the proposed metaheuristic on scalable unimodal and multimodal benchmark functions with standard metaheuristic. Further, we proposed Beecluster, a clustering protocol based on the proposed metaheuristic for WSNs, which selects optimal CHs based on an improved search equation and an efficient fitness function. In addition to this, we calculated optimal position of BS, through analytical evaluation of energy equations in WSNs. At last, we compare the performance of the proposed protocol with other protocols to prove its validness over various performance metrics. Simulation results show that Beecluster consumes approximately 32–131% less energy as compared to other protocols and prolongs network life approximately by 18–21% while delivering highest number of packets with minimum end-to-end delay in diverse WSNs scenarios.

Acknowledgements The authors acknowledge IKG Punjab Technical University, Kapurthala, Punjab, India.

Compliance with ethical standards

Conflict of interests The authors declare that there is no conflict of interests regarding the publication of this paper.

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

References

- Abro AG, Mohamad-Saleh J (2012) Enhanced global-best artificial bee colony optimization algorithm. In: Sixth UKSim-AMSS European symposium on computer modeling and simulation, pp 95–100
- Abbasi AA, Younis M (2007) A survey on clustering algorithms for wireless sensor networks. *Comput Commun* 30(14):2826–2841
- Ari AAA, Yenke BO (2016) A power efficient cluster-based routing algorithm for wireless sensor networks: honeybees swarm intelligence based approach. *J Netw Comput Appl*
- Akkaya K, Younis M (2005) A survey on routing protocols for wireless sensor networks. *Ad Hoc Netw* 3(3):325–349
- Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. *Wirel Commun IEEE* 11(6):6–28
- Attea BA, Khalil EA (2012) A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Appl Soft Comput* 12(7):1950–1957
- Akay B, Karaboga D (2012) A modified artificial bee colony algorithm for real-parameter optimization. *Inf Sci* 192(120):142
- Chamam A, Pierre S (2010) A distributed energy-efficient clustering protocol for wireless sensor networks. *Comput Electr Eng* 36(2):303–312
- Chen R (1984) Location problem with cost being sum of power of euclidean distances. *J Comput Oper Res* 11(3):285–294
- Das S, Sugantha PN (2011) Differential evolution: a survey of the state-of-the-art. *IEEE Trans Evolut Comput* 15(1):4–31
- Das S, Abraham A, Konar A (2009) Metaheuristic clustering. *Stud Comput Intell* 178:252
- Deng S, Li J, Shen L (2011) Mobility-based clustering protocol for wireless sensor networks with mobile nodes. *Wirel Sens Syst IET* 1(1):39–47
- Ding Y, Chen R, Hao K (2016) A multi-path routing algorithm with dynamic immune clustering for event-driven wireless sensor networks. *Neurocomputing*
- Ferrante Neri GI (2001) Compact optimization. In: *Handbook of Optimization*, ISRL 38, pp 337–364
- Gao W, Liu S (2011) Improved artificial bee colony algorithm for global optimization. *Inf Process Lett* 111(17):871–882
- Gao W, Liu LHS (2012) A global best artificial bee colony algorithm for global optimization. *J Comput Appl Math* 236(11):2741–2753
- Gao W, Liu LHS (2013) A novel artificial bee colony algorithm based on modified search equation and orthogonal learning. *IEEE Trans Cybernet* 43(3):1011–1024
- Gaura E (2010) *Wireless sensor networks: deployments and design frameworks*. Springer, New York
- Gonuguntla V, Mallipeddi R, Veluvolu KC (2015) Differential evolution with population and strategy parameter adaptation. *Math Probl Eng* 2015:287607. doi:10.1155/2015/287607
- Guo P, Cheng JLW (2011) Global artificial bee colony search algorithm for numerical function optimization. *Seventh Int Conf Nat Comput* 3:1280–1283
- Heinzlman WB, Chandrakasan AP, Balakrishnan H et al (2002) An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans Wirel Commun* 1(4):660–670
- Hoang D, Yadav P, Kumar R, Panda S (2014) Real-time implementation of a harmony search algorithm-based clustering protocol for energy efficient wireless sensor networks. *IEEE Trans Ind Inform* 10(1):774–783
- Jin Y, Wang L, Kim Y, Yang X (2008) Eemc: an energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks. *Comput Netw* 52(3):542–562
- Karaboga D, Akay B (2009) A comparative study of artificial bee colony algorithm. *Appl Math Comput* 214(1):108–132
- Karaboga D, Basturk B (2008) On the performance of artificial bee colony (abc) algorithm. *Appl Soft Comput* 8(1):687–697

- Khalil EA, Attea BA (2011) Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks. *Swarm Evolut Comput* 1(4):195–203
- Kuila P, Jana PK (2014) Energy efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach. *Eng Appl Artif Intell* 33:127–140
- Kulkarni RV, Forster A, Venayagamoorthy GK (2011) Computational intelligence in wireless sensor networks: a survey. *Commun Surv Tutor IEEE* 13(1):68–96
- Kumar D, Aseri TC, Patel R (2009) Eehc: energy efficient heterogeneous clustered scheme for wireless sensor networks. *Comput Commun* 32(4):662–667
- Larranaga P, Lozano JA (2001) Estimation of distribution algorithms: a new tool for evolutionary computation. Kluwer, Alphen aan den Rijn
- Li G, Niu XXP (2013) Development and investigation of efficient artificial bee colony algorithm for numerical function optimization. *Appl Soft Comput* 12(1):320–332
- Liu Z, Zheng Q, Xue L, Guan X (2012) A distributed energy-efficient clustering algorithm with improved coverage in wireless sensor networks. *Future Gen Comput Syst* 28(5):780–790
- Mao SS, Zhao CI W (2011) Unequal clustering algorithm for wsn based on fuzzy logic and improved aco. *J China Univ Posts Telecommun* 18(6):89–97
- Minno E, Cupertino DNF (2008) Real-valued compact genetic algorithms for embedded microcontroller optimization. *IEEE Trans Evol Comput* 12(2):203–219
- R Apostol MAM (2003) Sum of square of distance in m-space. *The Mathematics Asso of America*, pp 516–526
- Ozturk C, Hancer E (2015) Dynamic clustering with improved binary artificial bee colony algorithm. *Appl Soft Comput* 28(69):80
- Saleem M, Farooq M (2012) Beesensor: a bee-inspired power aware routing protocol for wireless sensor networks. In: Applications of evolutionary computing. *EvoWorkshops 2007. Lecture Notes in Computer Science*, vol 4448. Springer, New York, pp 81–90
- Samrat L, Udgata AAS (2010) Artificial bee colony algorithm for small signal model parameter extraction of mesfet. *Eng Appl Artif Intell* 11:1573–1592
- Selvakennedy S, Sinnappan S, Shang Y (2007) A biologically-inspired clustering protocol for wireless sensor networks. *Comput Commun* 30(14):2786–2801
- Storn RPK (2010) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J Glob Optim* 23:689–694
- Tyagi S, Kumar N (2012) A systematic review on clustering and routing techniques based upon leach protocol for wireless sensor networks. *J Netw Comput Appl* 36(1):623–645
- Walck C (1996) Handbook on statistical distributions for experimentalists. Internal report SUT-PFY/96–01. Stockholm
- Yang J, Xu M, Zhao W, Xu B (2009) A multipath routing protocol based on clustering and ant colony optimization for wireless sensor networks. *Sensors* 10(5):4521–4540
- Yi S, Heo J, Cho Y, Hong J (2007) Peach: power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks. *Comput Commun* 30(14):2842–2852
- Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. *Comput Netw* 52(12):2292–2330
- Younis O, Fahmy S (2004) Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans Mob Comput* 3(4):366–379
- Zhang R, Wu C (2011) An artificial bee colony algorithm for the job shop scheduling problem with random processing times. *Entropy* 13(9):1708–1729
- Zhu C, Zheng C, Shu L, Han G (2012) A survey on coverage and connectivity issues in wireless sensor networks. *J Netw Comput Appl* 35:619–632