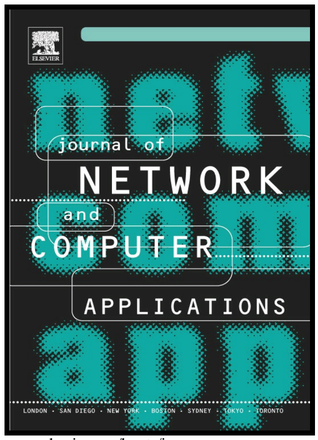
Author's Accepted Manuscript



www.elsevier.com/locate/jnca

Energy efficient clustering protocol based on improved metaheuristic in wireless sensor networks

Palvinder Singh Mann, Satvir Singh

PII: S1084-8045(17)30051-6
DOI: http://dx.doi.org/10.1016/j.jnca.2017.01.031
Reference: YJNCA1847

To appear in: *Journal of Network and Computer Applications*

Received date: 22 February 2016
Revised date: 16 August 2016
Accepted date: 24 January 2017

Cite this article as: Palvinder Singh Mann and Satvir Singh, Energy efficien clustering protocol based on improved metaheuristic in wireless sensor networks *Journal of Network and Computer Applications* http://dx.doi.org/10.1016/j.jnca.2017.01.031

Energy efficient clustering protocol based on improved metaheuristic in wireless sensor networks

Palvinder Singh Mann[a,*], Satvir Singh[b]

[a] *Research Scholar, I K Gujral Punjab Technical University, Kapurthala, Punjab, India.144008.*
[b] *S B S State Technical Campus, Ferozepur, Punjab, India.152004.*

Abstract

Energy efficient clustering is a well accepted NP-hard optimization problem in Wireless sensor networks (WSNs). Diverse paradigm of Computational intelligence (CI) including Evolutionary algorithms (EAs), Reinforcement learning (RL), Artificial immune systems (AIS), and more recently, Artificial bee colony (ABC) metaheuristic have been used for energy 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, which is comparably poor at exploitation and require storage of certain control parameters, contributes to its insufficiency. Thus, we present an improved Artificial bee colony (iABC) metaheuristic with an improved solution search equation to improve its exploitation capabilities. Additionally, in order to increase 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, hence 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, make it suitable for limited hardware requirements of WSNs. Further, an energy efficient clustering protocol based on iABC metaheuristic is introduced, which inherit the capabilities of the proposed metaheuristic to obtain optimal cluster heads (CHs) and improve energy efficiency in WSNs. Simulation results shows 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: Energy efficient clustering, Wireless sensor networks, improved Artificial bee colony (iABC) metaheuristic.

*Corresponding author.Tel.:+91-9888395367 fax:+91-0181-2205852
E-mail addresses: psmaan@hotmail.com (Palvinder Singh Mann), drsatvir.in@gmail.com (Satvir Singh)

1. Introduction

Advances in wireless communication and electronics has led to the development of 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 [1] [2]. SNs can have more than one sensor to capture data from the physical environment, wherever deployed. A sensor with limited storage and computation capabilities receive the sensed data through analogue to digital Converter (ADC) and process it further for transmission to a main location, known as *Base Station* (BS), where the data can be analysed for decision making in variety of applications [3]. 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 network due to excessive usage[4]. 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 [5]. 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 [6, 7], which increase network life by effectively utilize 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 send it to the BS for further processing. Clustering has various significant advantages over classical schemes [6]. 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 [8]. 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[9]. Thus, Computational intelligence (CI) [10] 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 [11]. 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 [12]. 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, specially while handling multi-modal optimization problems [13]. 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, *Beecluster* 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) [14], 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, 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) [15], is another classsical clustering protocol that selects CHs based on hybridisation of node residual energy and node proximity to its neighbours 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) [16], selects CHs without additional overhead of wireless communication and supports adaptive multi-level clustering for both location-unaware and location-aware WSNs but with high latency and low scalability thus make it suitable only for small networks. T-ANT [17], a swarm-inspired clustering protocol which exploit two swarm principles, namely separation and alignment, through pheromone control to obtain a stable and near uniform distribution for selection of CHs. Energy-Efficient Multi-level Clustering (EEMC) [18], achieve less energy consumption and minimum latency in WSNs by forming multi-level 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) [19], 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. Multi-path Routing Protocol (MRP) [20], 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) [8], presents a distributed clustering algorithm where CHs are selected based on a three-way message exchange between each sensor and its neighbours while possessing maximum residual energy and degree. However the protocol donot support multi-level clustering and consider small transmission ranges. Mobility-based clustering (MBC) protocol [21] support 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 set-up phase. UCFIA [22], 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 balances the energy consumption of CHs. Distributed Energy-Efficient Clustering with Improved Coverage (DEECIC) [23], 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 prolong network lifetime and improve network coverage. Energy-Aware Evolutionary Routing Protocol (ERP) [24], 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) [25], 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 optimize the energy distribution for WSNs. BeeSensor [26], 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. [27] 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 Particle swarm optimization (PSO). Their proposed algorithms demonstrate their proficiency in terms of network life, energy consumption, and delivery of data packets to the BS. Besides this some authors have presented extensive study on various aspects of WSNs as: [28] discussed the design of various data collection approaches in WSNs and presented a thematic taxonomy of energy saving techniques used in various hybrid WSN data gathering approaches. An extensive survey is presented by [29] to evaluate the diverse hierarchical routing protocols which are derived from LEACH, directly or indirectly. [30] provide a realistic insights on the practical advantages and limitations of using established routing techniques for tactical WSNs particularly used for military establishments. [31] presents a survey on opportunistic routing, as a new paradigm in routing for WSNs which selects the node closest to the target node for forwarding the data and increases the efficiency, throughput and reliability of sensor networks. Further, [32] presents a rule-driven multi-path 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 presents a relative comparison of these protocols, highlighting their features and limitations for a better insight:

Table 1: Relative comparison of protocols in WSNs

Protocol	Classification	Energy-efficiency	Features	Limitations
LEACH [14]	Classical	Average	Self-organization	High communication cost
HEED [15]	Classical	Average	Low communication cost	High latency
PEACH [16]	Classical	Average	Load balancing	High latency, low scalability
T-ANT [17]	Computational Intelligence	Good	Fast convergence, low overhead	Low coverage
EEMC [18]	Classical	Average	Minimum overhead, low latency	Only uniform node distribution
EEHC [19]	Classical	Good	Support node heterogeneity	Low scalability
MRP [20]	Computational Intelligence	Good	Prolong network lifetime	Need parameters adjustment
EECF [8]	Classical	Average	Prolong network lifetime	Low transmission range
MBC [21]	Classical	Average	High node mobility, low packet loss	High communication cost
UCFIA [22]	Computational Intelligence	Good	Prolong network lifetime	Need parameters adjustment
DEECIC [23]	Classical	Average	Better network coverage	Low scalability
ERP [24]	Computational Intelligence	Average	Better network lifetime	Non-cluster based approach
HSACP [25]	Computational Intelligence	Good	Fast convergence	No load balancing
BeeSensor [26]	Computational Intelligence	Good	Low processing cost	Non-cluster based approach
PSO [27]	Computational Intelligence	Good	Better packet delivery	Network overhead

It is very much clear from the comparison that classical as well as CI based approaches has 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 shown to be good in energy-efficiency with prolong 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 D. Karaboga [13] for optimizing multi-variable and multi-modal 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 meta-heuristics [33]. In ABC, there are three type of bees: employed bees, onlookers and scout bees [34]. 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, there after onlooker bee will select a food source based on a probability function related to the richness of that food source, whereas scout bee *explore* new food sources randomly around the hive. When a scout or onlooker bee finds a new food source, they become employed again, on the other 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 $x_{ij} = \{\ x_{i1}, x_{i2}, ...x_{iD}\}$ is the i-th food source, where $j = 1, 2......D$, is obtained by:

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

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 obtain a new solution v_{ij} from x_{ij} using expression:

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

where k is randomly obtained from { 1,2..... 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:

$$p_i = \frac{f_i}{\sum_{n=1}^{SN} f_i} \tag{3}$$

where f_i is the fitness of x_{ij}. Onlooker bee choose a food sources 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. ABC variants

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 [35] introduced search equations which are based on Differential evolution (DE) [36] algorithms, to solve numerous real-world optimization problems.

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

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

where $\phi_{i,j}$ is a uniform random number and $x_{r,j}$ is the jth component of a random solution. In addition, Gao [37] introduced another variant later as

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

Other search equations are proposed by [38] [39]

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

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

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

where x_{sbj} is the j coefficient of the second best solution. However, the effectiveness of these DE based equations will critically depends 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 [40] into Eq.(10), where w_{ij} is the relative weight and θ_1, θ_2 are parameters to control step size.

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

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

Scout bee phase also witnessed some improvements [41] in form of new solution search equations (11)(12) as:

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

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

In addition, some improved versions of the ABC metaheuristic [42] include some parameters, like Modification Rate (MR) and Scale Factor (SF), MR controls the neighbourhood 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 proposes an improved artificial bee colony (*iABC*) metaheuristic with better sampling technique using **Students-t** distribution; a compact probability density function (cPDF), which requires only one control parameter to be stored on memory. *Students-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 improve 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 extend 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 multi-modal optimization problems. To overcame these limitations we proposes an improved artificial bee colony ($iABC$) metaheuristic with an improved initialization phase for better sampling and improved solution search equation, named 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 [43] through family of Estimation of Distribution Algorithms (EDA) [44] framework by considering compact probability density functions (cPDFs). Therefore, we propose $Student's-t$ distribution [45] ; 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 needed to be stored on memory. This cPDF is being introduced first time in population-based metaheuristic due to its compact nature.

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

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:

$$
\begin{aligned}
\int_{-x}^{x} f(x)dx &= \frac{1}{\sqrt{\kappa}B\left(\frac{1}{2},\frac{\kappa}{2}\right)} \int_{-x}^{x} \left(1+\frac{x^2}{\kappa}\right)^{-(\kappa+1)/2} dx \\
&= \frac{2}{\sqrt{\kappa}B\left(\frac{1}{2},\frac{\kappa}{2}\right)} \int_{0}^{x} \left(1+\frac{x^2}{\kappa}\right)^{-(\kappa+1)/2} dx \\
&= \frac{-2}{\sqrt{\kappa}B\left(\frac{1}{2},\frac{\kappa}{2}\right)} \int_{1}^{\kappa/\kappa+x^2} \frac{x^{\kappa+1/2}\kappa\sqrt{x}}{2x^2\sqrt{\kappa}\sqrt(1-x)} dx \\
&= \frac{1}{B\left(\frac{1}{2},\frac{\kappa}{2}\right)} \int_{\kappa/\kappa+x^2}^{1} (1-x)^{-1/2} x^{\kappa/2-1} dx \\
&= \frac{1}{B\left(\frac{1}{2},\frac{\kappa}{2}\right)} \left(B\left(\frac{\kappa}{2},\frac{1}{2}\right) - B_{\frac{\kappa}{\kappa+x^2}}\left(\frac{\kappa}{2},\frac{1}{2}\right)\right) \\
&= 1 - I_{\frac{\kappa}{\kappa+x^2}}\left(\frac{\kappa}{2},\frac{1}{2}\right) = I_{\frac{x^2}{\kappa+x^2}}\left(\frac{1}{2},\frac{\kappa}{2}\right)
\end{aligned} \tag{14}
$$

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.

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

$$
Q_j(x) = \frac{1}{2} + \frac{1}{2} I_{\frac{x_{ij}^2}{\kappa+x_{ij}^2}}\left(\frac{1}{2},\frac{\kappa}{2}\right) \qquad for \quad 0 \le x < 1 \tag{16}
$$

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) [46] employs most powerful stochastic real-parameter algorithms to solve multi-modal 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** [47] [48]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.

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

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:

$$ABC/rand-to-opt/1: \quad v_{ij} = x_{ij} + \phi_{ij}(x_{opt,j} - x_{ij}) + \psi_{ij}(x_{r_1j} - x_{r_2j}) \tag{18}$$

where r_1, r_2 are random variables from $1, 2, ..., SN$, 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 :

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

Then a selection operation will be performed as :

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

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 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 uni-modal and functions f_5 to f_8 are multi-modal functions as shown in table [2] :

Table 2: Benchmark functions used in experiment

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

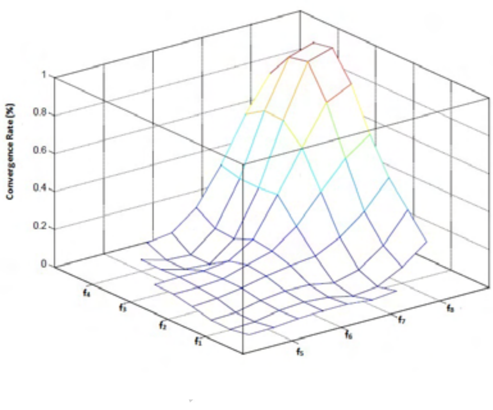


Fig. 1: Convergence Rate of iABC

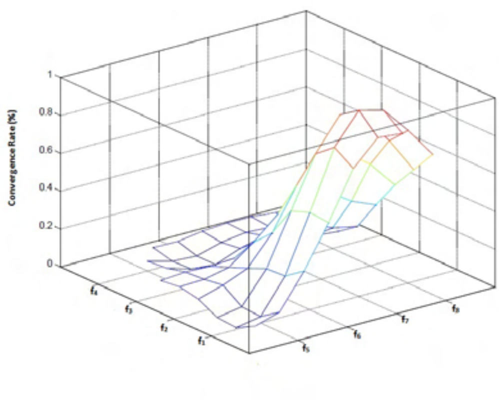


Fig. 2: Convergence Rate of ABC

Graphs [1] [2] show that the proposed *iABC* metaheuristic convergence fast with optimal or closer-to-optimal solutions on the uni-modal as well as complex multi-modal 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 $S = \{ s_1, s_2, ...s_n\}$, 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 $l_{i,j} \epsilon$ L represents wireless link between node s_i and s_j.
3. Set of Cluster Heads (CH's) are denoted by $S_{ch} = \{ ch_1, ch_2, ...ch_k\}$ where $S_{ch} \epsilon$ S.
4. $D_{s_i}^{s_j}$(max) denotes the maximum distance between a senor node s_i and s_j , is calculated by squared Euclidean distance between them as

$$\begin{aligned} D_{s_i}^{s_j}(max) &= Max\{\ dis(s_i, s_j)\} \quad &| \ \forall\ s_i, s_j\ \epsilon\ S \\ &= \|s_i - s_j\|^2 = \sum (s_i - s_j)^2 \quad &| \ \forall\ s_i, s_j\ \epsilon\ S \end{aligned} \tag{21}$$

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

$$\begin{aligned} D_{s_i}^{s_{n+1}}(max) &= Max\{\ dis(s_i, s_{n+1})\} \quad &| \ \forall\ s_i\ \epsilon\ S \\ &= \|s_i - s_{n+1}\|^2 = \sum (s_i - s_{n+1})^2 \quad &| \ \forall\ s_i\ \epsilon\ S \end{aligned} \tag{22}$$

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

$$\begin{aligned} D_{s_i}^{ch_j}(max) &= Max\{\ dis(s_i, ch_j)\} \quad &| \ \forall\ s_i,\ ch_j\ \epsilon\ S \\ &= \|s_i - ch_j\|^2 = \sum (s_i - ch_j)^2 \quad &| \ \forall\ s_i,\ ch_j\ \epsilon\ S \end{aligned} \tag{23}$$

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

$$\begin{aligned} D_{ch_j}^{s_{n+1}}(max) &= Max\{\ dis(ch_j, s_{n+1})\} \quad &| \ \forall\ j\ \epsilon\ S_{ch} \\ &= \|ch_j - s_{n+1}\|^2 = \sum (ch_j - s_{n+1})^2 \quad &| \ \forall\ j\ \epsilon\ S_{ch} \end{aligned} \tag{24}$$

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

$$P_{tran_i} = \frac{1}{k}\left(\frac{\gamma.T_{delay}}{dc}\right)^{\alpha} \tag{25}$$

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

$$T_{delay} = T_{que} + T_{tran} + T_{ack} \tag{26}$$

The first component; T_{que} is the queuing delay, second component; T_{tran}, is the transmission delay and third one 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 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 [14], 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}.

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

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

$$E_{trans}(l, d) = l.E_{elec} + E_{amp}(l, d) \tag{28}$$

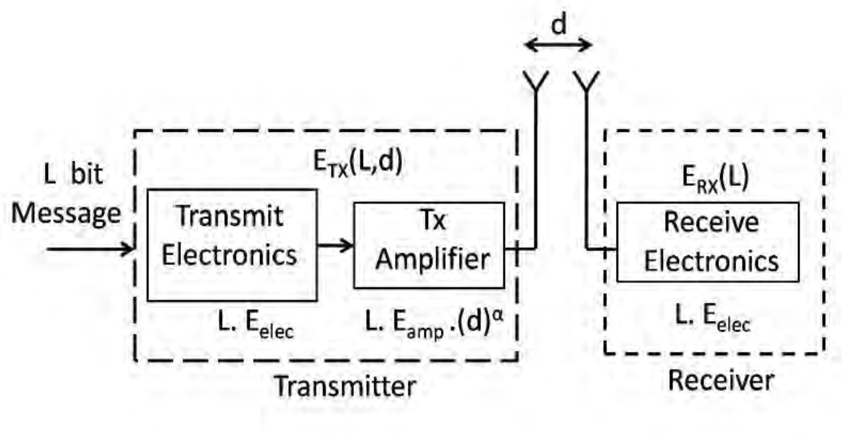


Fig. 3: Radio Model for energy analysis

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 $\tau_{ack} = \dfrac{l_{ack}}{l}$, is the ratio between length of acknowledgement packet to data packet.

therefore, residual energy of each senor 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 $\dfrac{n}{k}$ nodes per cluster. Out of these, there will be one CH node and remaining $\dfrac{n}{k} - 1$ 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 :

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

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 :

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

therefore, energy consumed in whole network per round

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

Lemma 1: $\forall$ { $s_1, s_2, ...s_n$} ϵ S near to BS, the total energy consumption for transmission of l 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 { $s_1, s_2, ...s_n$} ϵ S.

Proof: Let { $s_1, s_2, ...s_n$} ϵ 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 :

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

as signals from all the nodes suffer free space loss while transmitting data, therefore:

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

so, Eq.(38) becomes:

$$E_{round} = l\left((2n-k).E_{elec} + n.E_{da} + (n-k)\epsilon_{fs}(\frac{M^2}{2\pi k}) + (\frac{n}{k}-1)E_{ack} + \epsilon_{fs}\sum_{j=1}^{k}(D_{ch_j}^{s_{n+1}})^2\right) \tag{40}$$

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

$$E_{total}(l, D_{s_i}^{s_{n+1}}) = l.\frac{n}{k}\left((2n-k).E_{elec} + n.E_{da} + (n-k)\epsilon_{fs}(\frac{M^2}{2\pi k}) + E_{ack}\right) + l.\epsilon_{fs}\sum_{j=1}^{n}(D_{ch_j}^{s_{n+1}})^2 \tag{41}$$

therefore,

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto \quad l.\epsilon_{fs}\sum_{j=1}^{n}(D_{ch_j}^{s_{n+1}})^2 \tag{42}$$

thus, from Eq.(24):

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto \quad Min\left(l.\epsilon_{fs}\sum_{j=1}^{n}(ch_j - s_{n+1})^2\right) \quad | \quad \forall\, j \,\epsilon\, S_{ch} \tag{43}$$

further, to obtain optimal location of BS, we need to minimize $E_{total}(l, D_{s_i}^{s_{n+1}})$ and sum of square of Euclidean distances will be minimum at the centroid [49]. Thus, $E_{total}(l, D_{s_i}^{s_{n+1}})$ will be minimum if BS will be placed at centroid of $\{\ s_1, s_2, ...s_n\} \,\epsilon$ S.

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

Proof: As $\{\ s_1, s_2, ...s_n\} \,\epsilon$ 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, $E_{total}(l, D_{s_i}^{s_{n+1}})$ will be given by:

$$E_{total}(l, D_{s_i}^{s_{n+1}}) = l.\frac{n}{k}\left((2n-k).E_{elec} + n.E_{da} + (n-k)\epsilon_{fs}(\frac{M^2}{2\pi k}) + E_{ack}\right) + l.\epsilon_{fs}\sum_{j=1}^{n}(D_{ch_j}^{s_{n+1}})^4 \tag{44}$$

therefore,

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto \quad l.\epsilon_{fs}\sum_{j=1}^{n}(D_{ch_j}^{s_{n+1}})^4 \tag{45}$$

thus, from Eq. (24):

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto \quad Min\Bigg(l.\epsilon_{fs}\sum_{j=1}^{n}(ch_j - s_{n+1})^4\Bigg) \quad | \quad \forall\ j\ \epsilon\ S_{ch} \tag{46}$$

further, to obtain optimal location of BS, we need to minimize $E_{total}(l, D_{s_i}^{s_{n+1}})$ and polynomial of Euclidean distances will be minimum at the centroid only [50]. Thus, $E_{total}(l, D_{s_i}^{s_{n+1}})$ will be minimum if BS will be placed at centroid of $\{\ s_1, s_2, ...s_n\}\ \epsilon$ S.

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

Proof: As $\{\ s_1, s_2, ...s_a\}\ \epsilon$ S are deployed near to BS and $\{\ s_1, s_2, ...s_b\}\ \epsilon$ 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, $E_{total}(l, D_{s_i}^{s_{n+1}})$ will be given by:

$$E_{total}(l, D_{s_i}^{s_{n+1}}) = l.\frac{n}{k}\Bigg((2n-k).E_{elec} + n.E_{da} + (n-k)\epsilon_{fs}(\frac{M^2}{2\pi k}) + E_{ack}\Bigg) + l.\epsilon_{fs}\sum_{i=1}^{x}(D_{ch_i}^{s_{n+1}})^2 + l.\epsilon_{fs}\sum_{j=1}^{y}(D_{ch_i}^{s_{n+1}})^4 \tag{47}$$

$$E_{total}(l, D_{s_i}^{s_{n+1}}) = l.\frac{n}{k}\Bigg((2n-k).E_{elec} + n.E_{da} + (n-k)\epsilon_{fs}(\frac{M^2}{2\pi k}) + E_{ack}\Bigg) + l.\epsilon_{fs}\Bigg(\sum_{i=1}^{x}(D_{ch_i}^{s_{n+1}})^2 + \sum_{j=1}^{y}(D_{ch_i}^{s_{n+1}})^4\Bigg) \tag{48}$$

therefore,

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto l.\epsilon_{fs}\Bigg(\sum_{i=1}^{x}(D_{ch_i}^{s_{n+1}})^2 + \sum_{j=1}^{y}(D_{ch_i}^{s_{n+1}})^4\Bigg) \tag{49}$$

thus, from Eq. (24):

$$E_{total}(l, D_{s_i}^{s_{n+1}}) \quad \propto \quad Min\left(l.\epsilon_{fs} \left(\sum_{j=1}^{n} (ch_j - s_{n+1})^2 + \sum_{j=1}^{n} (ch_j - s_{n+1})^4 \right) \right) \qquad | \quad \forall\, j \,\epsilon\, S_{ch} \tag{50}$$

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

7.3. Optimal CH selection phase

CH selection is one of the crucial task for cluster formation in WSNs as it effect 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 multi-modal 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), corresponding food sources (SN) are initialised 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 $r\epsilon$ [0, 1] according to uniform distribution and obtain x_{ij} as :

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

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

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

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

$$f_i \propto \frac{1}{Min(D_{ch_j}^{s_{n+1}}(max) + P_{tran_i})} \tag{53}$$

aggregating Eqs.(52)(53) as

$$f_i \propto \frac{Max(E_{res})}{Min(D_{ch_j}^{s_{n+1}}(max) + P_{tran_i})} \tag{54}$$

$$f_i = K \frac{Max(E_{res})}{Min(D_{ch_j}^{s_{n+1}}(max) + P_{tran_i})} \tag{55}$$

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

$$f_i = \frac{Max(E_{res})}{Min(D_{ch_j}^{s_{n+1}}(max) + P_{tran_i})} \tag{56}$$

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 equation (19) of proposed $iABC$ metaheuristic as:

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

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 $x_{opt,j}$ 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:

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

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

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

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.

The detail 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**

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

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

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

  else

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

    $Evaluate\ fitness\ f_i(x_{ij})$

$trial(s) \leftarrow 0$

$round++$

end if

end for

repeat

until

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

Generate v_{ij} *according to Eq.*

$Evaluate\ fitness\ f_i(v_{ij})$

$round++$

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**

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

$Ch_j \leftarrow x_{opt,j}$

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**

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

$Ch_j \leftarrow x_{opt,j}$

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 10000 rounds and obtained number of CHs per round.

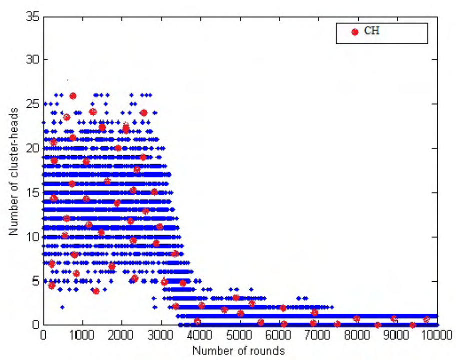


Fig. 4: CHs selected per round in $iABC^2$.

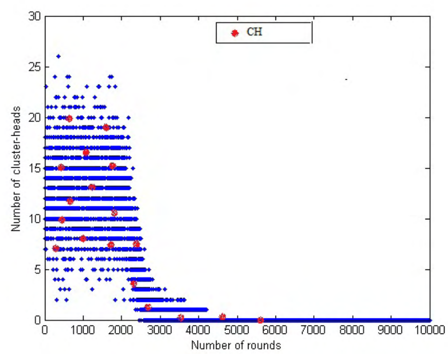


Fig. 5: CHs selected per round in ABC.

graphs [4] [5] shows that the proposed metaheuristic clearly supports more number of CHs per round compared to its peer, which in terns 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 .

7.4. Cluster Formation phase

After selection of CHs, each CH will advertise a Join-Request (J-REQ) message to all its neighbour 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 became member of the cluster. During this communication, all CH nodes must keep their receivers on and listen to the channel. If a particular node receive multiple J-REQ message from same CH then it discard 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.

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 BeeSensor [26] , MRP [20] and ERP [24] protocols using ns-2 simulator. The protocols are simulated over two different BS position scenarios to assess their behaviour towards 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:

Table 3: Simulation Parameters

Parameter	**Value**
Terrain size	150 * 150 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/m4
Propagation limit	-111 dBm
Receiver sensitivity	-89
Data rate	2 Mbps
Packet Size	3000 bits
Message Size	400 bits

In the first scenario WSN # 1, a network of sensor nodes ranging from 100 to 700 are deployed randomly over an area of size 150 $*$ 150 m^2 with a BS, located at optimal position calculated through proposed evaluation of energy equations [7.2] within the network field, whereas in second scenario WSN # 2, a BS will be placed at position (100 m, 200 m) outside the network field. First, we execute the protocols to compare Packet delivery ratio (PDR) in the network for both the scenarios.

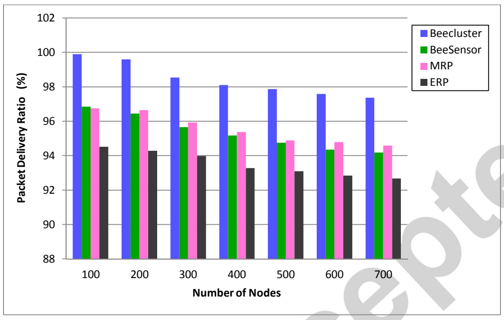


Fig. 6: Packet delivery ratio in WSN # 1

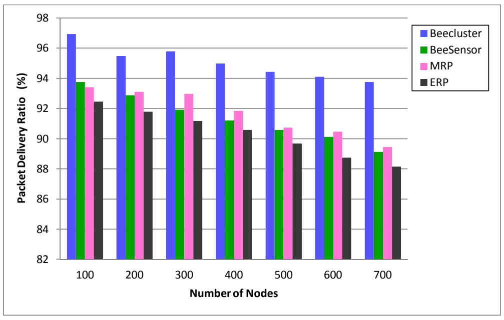


Fig. 7: Packet delivery ratio in WSN # 2

In scenario WSN # 1, Fig.[6] shows that the proposed protocol deliver 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 position in WSN # 1 scenario. Even in WSN # 2 scenario, Fig.[7] shows that *Beecluster* has highest PDR as compared to BeeSensor, MRP and ERP at nodes ranging from 100 to 700. It is further analysed from above Figs.[6] [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.

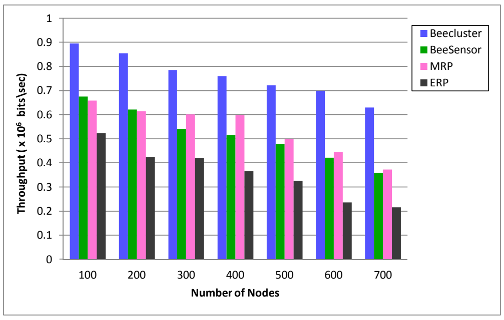


Fig. 8: Throughput in WSN # 1

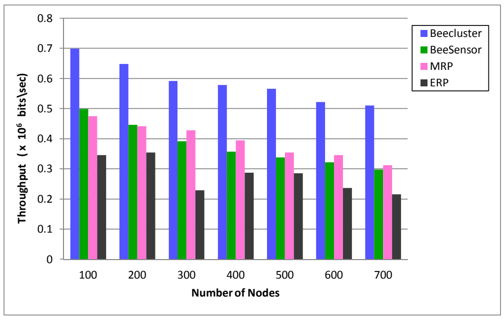


Fig. 9: Throughput in WSN # 2

Throughput is a measure of robustness for any protocol and Figs.[8] [9] shows that *Beecluster* is delivering highest number of packets per second even at 700 nodes deployed in the both scenarios. The low performance of BeeSensor and ERP is due to the fact that they are non-cluster based protocols, hence lack in performance while delivering data from nodes to BS. Additionally, it is observed that optimal BS position scenario WSN # 1 Figs.[8], witness highest throughput compared to WSN # 2 Figs.[9], thus placing BS at optimal position will increase data delivery per second in WSNs.

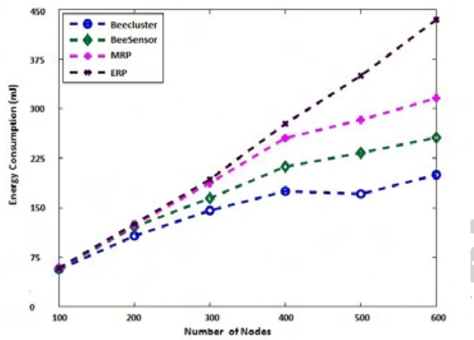


Fig. 10: Energy consumption in WSN # 1

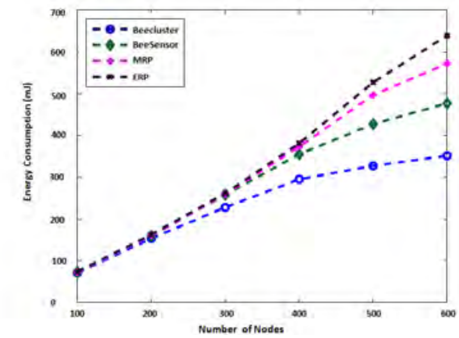


Fig. 11: Energy consumption in WSN # 2

Fig.[10] shows that in scenario WSN # 1, energy consumption of the proposed protocol is approximately 29 % , 67 % , 127% less then BeeSensor, 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 minimise energy consumption in the network. Even in scenario WSN #2 Fig. [11], *Beecluster* consumes 36 % less energy as compared to its contender BeeSensor, which clearly shows the effectiveness of the proposed metaheuristic *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 consume less energy and as a result the overall energy consumption of the network becomes lesser than other protocols. In MRP, all CHs are inevitably used as a relay node to forward the data packets to the BS, therefore consume more energy whereas BeeSensor and ERP are non-cluster based protocols, and are not able to optimise the energy usage in complex WSN scenario. Moreover, placement of BS at optimal location will decrease the energy consumption as evident form Figs. [10] and [11] respectively. Further Fig.[12] shows the average percentage reduction in energy consumption, when BS is placed at optimal position in a square field with nodes ranging form 0 to 1000. The results shows that average reduction in energy become highest when number of nodes varies from 10 to 70 with length of square field between 210 to 480 meters. Nodes which are placed below 200 meters will suffer free space loss whereas nodes placed above 500 meters may suffer multi-path loss as well.

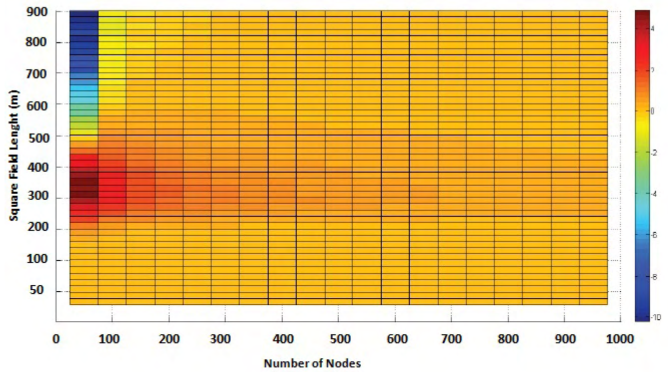


Fig. 12: Average % reduction in energy consumption at optimal BS position

Fig.[13] and Fig.[14] shows that *Beecluster* extend the average network lifetime by approximately 19 % and 14 % compared to BeeSensor 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.

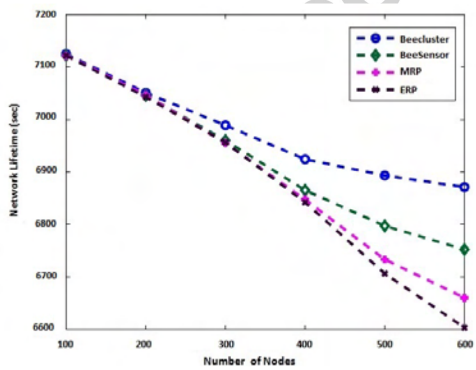


Fig. 13: Network lifetime in WSN # 1

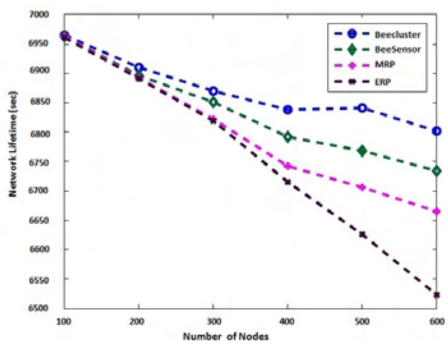


Fig. 14: Network lifetime in WSN # 2

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 reduces 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 analysed that every 1 % increase in network lifetime of the proposed protocol, will increase the data delivery by 2.2 % , thus increase networks robustness.

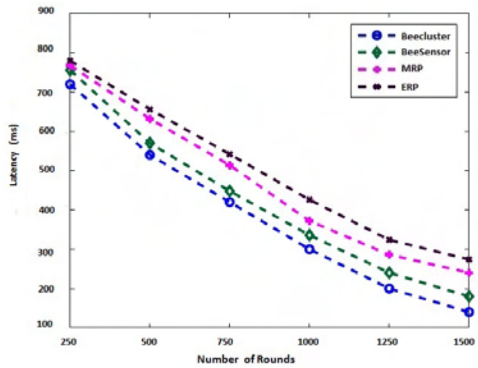


Fig. 15: Average latency in WSN # 1

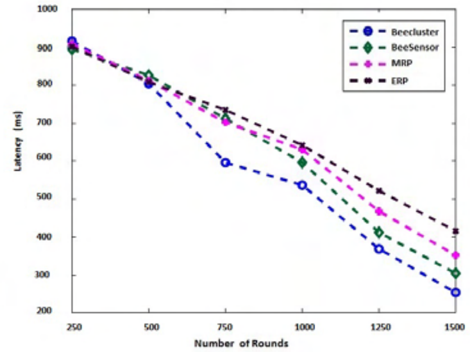


Fig. 16: Average latency in WSN # 2

Fig.[15] and Fig.[16] compares 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 increase 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 maintains 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 delivery the data packets with minimum delay successfully. In BeeSensor and MRP, data will be transmitted to BS using maximum number of hop-count ultimately exhaust the network with unnecessary end-to-end delay.

9. Conclusions

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 a well as convergence rate of existing ABC metaheuristic. To demonstrate its improvement, we evaluated the proposed metaheuristic on scalable uni-modal and multi-modal 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 29 % to 127% less energy as compared to other protocols and prolong network life approximately by 14 % to 19 % while delivering highest number of packets with minimum end-to-end delay in diverse WSNs scenarios. The significance of optimal BS position is also analysed with percentage reduction in energy consumption over varying distance with number of sensor nodes. Further, the proposed protocol needs to be implemented on real test bed scenario of sensor nodes, which are deployed to work on a real world application framework to judge its performance.

10. Acknowledgement

The authors acknowledge I K Gujral Punjab Technical University, Kapurthala, Punjab, India.

11. References

[1] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Computer networks 52 (2008) 2292–2330.

[2] Z. C. S. L. H. G. Zhu, C, A survey on coverage and connectivity issues in wireless sensor networks, Journal of Network and Computer Applications 35 (2012) 619–632.

[3] J. N. Al-Karaki, A. E. Kamal, Routing techniques in wireless sensor networks: a survey, Wireless Communications, IEEE 11 (2004) 6–28.

[4] K. Akkaya, M. Younis, A survey on routing protocols for wireless sensor networks, Ad hoc networks 3 (2005) 325–349.

[5] E. Gaura, Wireless Sensor Networks: Deployments and Design Frameworks, Springer, 2010.

[6] A. A. Abbasi, M. Younis, A survey on clustering algorithms for wireless sensor networks, Computer communications 30 (2007) 2826–2841.

[7] S. Tyagi, N. Kumar, A systematic review on clustering and routing techniques based upon leach protocol for wireless sensor networks, Journal of Network and Computer Applications (2012).

[8] A. Chamam, S. Pierre, A distributed energy-efficient clustering protocol for wireless sensor networks, Computers & electrical engineering 36 (2010) 303–312.

[9] E. A. Khalil, B. A. Attea, Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks, Swarm and Evolutionary Computation 1 (2011) 195–203.

[10] R. V. Kulkarni, A. Forster, G. K. Venayagamoorthy, Computational intelligence in wireless sensor networks: A survey, Communications Surveys & Tutorials, IEEE 13 (2011) 68–96.

[11] A. K. S. Das, A. Abraham, Metaheuristic clustering, Studies in Computational Intelligence 178 (2009).

[12] A. A. L. Samrat, S. Udgata, Artificial bee colony algorithm for small signal model parameter extraction of mesfet, Eng. Appl. Artif. Intelligence 11 (2010) 1573–1592.

[13] D. Karaboga, B. Akay, A comparative study of artificial bee colony algorithm, Applied Mathematics and Computation 214 (2009) 108–132.

[14] W. B. Heinzelman, A. P. Chandrakasan, H. Balakrishnan, et al., An application-specific protocol architecture for wireless microsensor networks, IEEE Transactions on wireless communications 1 (2002) 660–670.

[15] O. Younis, S. Fahmy, Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks, IEEE Transactions on Mobile Computing 3 (2004) 366–379.

[16] S. Yi, J. Heo, Y. Cho, J. Hong, Peach: Power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks, Computer communications 30 (2007) 2842–2852.

[17] S. Selvakennedy, S. Sinnappan, Y. Shang, A biologically-inspired clustering protocol for wireless sensor networks, Computer Communications 30 (2007) 2786–2801.

[18] Y. Jin, L. Wang, Y. Kim, X. Yang, Eemc: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks, Computer Networks 52 (2008) 542–562.

[19] D. Kumar, T. C. Aseri, R. Patel, Eehc: Energy efficient heterogeneous clustered scheme for wireless sensor networks, Computer Communications 32 (2009) 662–667.

[20] J. Yang, M. Xu, W. Zhao, B. Xu, A multipath routing protocol based on clustering and ant colony optimization for wireless sensor networks, Sensors 10 (2009) 4521–4540.

[21] S. Deng, J. Li, L. Shen, Mobility-based clustering protocol for wireless sensor networks with mobile nodes, Wireless Sensor Systems, IET 1 (2011) 39–47.

[22] S. Mao, S., C.-l. Zhao, W., Unequal clustering algorithm for wsn based on fuzzy logic and improved aco, The Journal of China Universities of Posts and Telecommunications 18 (2011) 89–97.

[23] Z. Liu, Q. Zheng, L. Xue, X. Guan, A distributed energy-efficient clustering algorithm with improved coverage in wireless sensor networks, Future Generation Computer Systems 28 (2012) 780–790.

[24] B. A. Attea, E. A. Khalil, A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks, Applied Soft Computing (2012).

[25] D. Hoang, P. Yadav, R. Kumar, S. Panda, Real-time implementation of a harmony search algorithm-based clustering protocol for energy efficient wireless sensor networks, IEEE Transactions On Industrial Informatics (2014).

[26] M. Saleem, M. Farooq, Beesensor: A bee-inspired power aware routing protocol for wireless sensor networks, in: Applications of Evolutionary Computing, Springer, 2012, pp. 81–90.

[27] P. Kuila, P. K. Jana, Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach, Engineering Applications of Artificial Intelligence 33 (2014) 127–140.

[28] A. H. A. Gaddafi Abdul-Salaam, A comparative analysis of energy conservation approaches in hybrid wireless sensor networks data collection protocols., Telecommunication Systems 61 (2016) 159–179.

[29] M. S. Vishal Kumar Arora, Vishal Sharma, A survey on leach and otherŠs routing protocols in wireless sensor network, Optik - International Journal for Light and Electron Optics (2016).

[30] K. A. W. Preetha Thulasiraman, Topology control of tactical wireless sensor networks using energy efficient zone routing, Digital Communications and Networks 2 (2016) 1Ű14.

[31] R. S. Payal Jadhav, A survey on opportunistic routing protocols for wireless sensor networks, Procedia Computer Science 79 (2016) 603Ű609.

[32] R. C. K. H. Yongsheng Ding, A multi-path routing algorithm with dynamic immune clustering for event-driven wireless sensor networks, Neurocomputing (2016).

[33] D. Karaboga, B. Basturk, On the performance of artificial bee colony (abc) algorithm, Applied soft computing 8 (2008) 687–697.

[34] R. Zhang, C. Wu, An artificial bee colony algorithm for the job shop scheduling problem with random processing times, Entropy 13 (2011) 1708–1729.

[35] S. L. W. Gao, Improved artificial bee colony algorithm for global optimization, Inform. Process. Lett 111 (17) (2011) 871–882.

[36] G. I. Ferranate Neri, Compact optmization, Optmization Algorithm (2001).

[37] L. H. W. Gao, S. Liu, A global best artificial bee colony algorithm for global optimization, J. Comput. Appl. Math 236 (11) (2012) 2741–2753.

[38] J. M. S. A. Abro, Enhanced global-best artificial bee colony optimization algorithm, Sixth UKSim-AMSS European Symposium on Computer Modeling and Simulation (2012) 95–100.

[39] L. H. W. Gao, S. Liu, A novel artificial bee colony algorithm based on modified search equation and orthogonal learning, IEEE Trans. Cybernet 43 (3) (2013) 1011–1024.

[40] X. X. G. Li, P. Niu, Development and investigation of efficient artificial bee colony algorithm for numerical function optimization, Appl. Soft Computing 12 (1) (2013) 320–332.

[41] J. L. P. Guo, W. Cheng, Global artificial bee colony search algorithm for numerical function optimization, Seventh International Conference on Natural Computation 3 (2011) 1280–1283.

[42] D. K. B. Akay, A modified artificial bee colony algorithm for real-parameter optimization, Inform. Science 192 (2012) 120Ű142.

[43] D. N. E. Mininno, F. Cupertino, Real-valued compact genetic algorithms for embedded microcontroller optimization, IEEE Trans. Evol. Computer 12(2) (2008) 203–219.

[44] J. L. P. Larranaga, Estimation of distribution algorithms: a new tool for evolutionary computation, Kluwer (2001).

[45] C. Walck, Statistical Distributions for experimentalists, Particle Physics Group, 2007.

[46] P. K. Storn R, Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces, Journal of Globa lOptimization 23 (2010) 689–694.

[47] P. N. S. Swagatam Das, Differential evolution: A survey of the state-of-the-art, IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION 15 (2011).

[48] K. C. V. V. Gonuguntla, R.Mallipeddi, Differential evolution with population and strategy parameter adaptation, Mathematical Problems in Engineering (2015).

[49] M. A. M. R. Apostol, Sum of square of distance in m-space, The Mathematics Asso. of America (2003) 516–526.

[50] R. Chen, Location problem with cost being sum of power of euclidean distances., The Journal of computers and operational research 11 (3) (1984) 285–294.