NSGA-II with ENLU inspired clustering for wireless sensor networks

Gunjan, Ajay K Sharma & Karan Verma

Wireless Networks
The Journal of Mobile Communication, Computation and Information

ISSN 1022-0038

Wireless Netw
DOI 10.1007/s11276-020-02281-8

ONLINE FIRST

Springer

Wireless Networks
https://doi.org/10.1007/s11276-020-02281-8

NSGA-II with ENLU inspired clustering for wireless sensor networks

Gunjan[1] · Ajay K Sharma[2] · Karan Verma[1]

Abstract

Wireless sensor networks (WSNs) have a large number of existing applications and is continuously increasing. Thus it is envisioned that WSN will become an integral part of our life in the near future. Direct propagation, chain formation, cluster creation are various techniques by which data is communicated by sensor nodes to the sink. It has been proved that Clustering is an efficient and scalable method to utilize the energy of sensor nodes efficiently. Optimal election of cluster heads is an NP (non deterministic polynomial time)-Hard problem. In our proposed work, a multi-objective optimization algorithm, non dominated sorting genetic algorithm-II based clustering in wireless sensor networks has been proposed. Energy conservation, network lifetime, coverage and load balancing are the four conflicting objective functions used. Our proposed algorithm handles all of these multiple objectives simultaneously. To reduce the computational complexity of the algorithm, efficient non-dominated level update mechanism for sorting has been used, which eliminates the need of applying non dominated sorting from scratch every time. The algorithm returns a solution set consisting of multiple non dominated solutions, wherein every solution is a best solution according to some objective function, in a single run, from which any solution can be chosen based on user preferences. According to our simulation carried on MATLAB, the proposed approach outperforms the established clustering algorithms in terms of network characteristics such as network lifetime, energy consumption and number of packets received.

Keywords Wireless sensor network · Clustering · Evolutionary algorithm · Energy efficiency · Network lifetime

1 Introduction

Advances in Micro-Electro-Mechanical-Systems (MEMS) in the recent years have led to the development of sensors with smaller size and fewer cost [1]. These sensors measure environmental factors such as temperature, pressure, air contaminants, humidity and then the gathered measurements are converted into signals which describe the characteristics of the phenomenon of interest present in the area [2]. A sensor network comprises of a sensing unit, processing unit, computing unit and a power unit (mobility unit and position finding system may also be present). A Wireless Sensor Network (WSN) consists of a large number of these devices which cooperate with each other to communicate the gathered information to the sink/base station. There exists a plethora of WSN applications including environmental monitoring, agriculture, health care, military, home automation and many more [3].

These sensors may be deployed randomly in the field by throwing them from the sky or it may be uniform by placing them consciously in a certain area depending upon the reach-ability [4]. WSNs provide us the advantage that they can operate in conditions in which the typical computing devices cannot function e.g. rain, sleet, high temperature, high humidity etc [5]. Due to the operating conditions of sensors, the batteries of sensors are often not replaceable. Thus energy conservation is a major concern in wireless sensor networks.

Equation 1 describes the transmission energy of a sensor in transmitting a m bit message at a distance of l units:

$$E_{Tx}(m,l) = \begin{cases} m * E_{el} + m * E_{fs} * l^2 & if\ l < do \\ m * E_{el} + m * E_{mp} * l^4 & if\ l > \ = do \end{cases} \quad (1)$$

where d_0 is the threshold distance, given by Eq. 2 as:

✉ Gunjan
gunjan@nitdelhi.ac.in

1 Computer Science and Engineering, National Institute of Technology, Delhi, India

2 I.K.Gujral Punjab Technical University, Kapurthala, India

Published online: 06 March 2020

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (2)$$

E_{el} is defined as the electronics energy, used by the hardware circuitry. $E_{fs} * l^2$ or $E_{mp} * l^4$ is the energy which is expended in amplification. It is proportional to the distance between sender and receiver.

If this distance (l) is less than the threshold distance d_0, then the free space model is used, otherwise the multi-path model is used [6–8].

Equation 3 gives the energy expended to receive a m bit message :

$$E_{Rx}(m) = m * E_{e1} \quad (3)$$

The reception energy is independent of distance between sender and receiver.

Due to energy constrained nature of sensor nodes, it is required that the WSN protocols send the sensed data to the sink wisely.

Protocol stack assists the collaboration of sensor nodes to collect and forward the data. Due to the diverse applications of WSN, there exists a large number of protocol standards each with their own protocol stack.

The most popular and earlier solutions include ZigBee, Z-Wave, INSTEON, Wavenis. ZigBee protocol stack comprises of following four layers: physical (PHY)layer, medium access control (MAC) layer, network (NWK) layer, and application (APL) layer. PHY and MAC are defined by the IEEE 802.15.4 standard, while NWK and APL are established by the ZigBee specification. Three types of devices are defined: coordinator, router and end device [9]. A number of algorithms in WSN based on the ZigBee protocol stack have been developed [10–13]. The Z-Wave stack consists of PHY/MAC layer, transport layer, network routing layer, and the application layer. Two types of devices are defined : controllers and slaves. Controllers send commands to slaves, while slaves execute the commands and reply back to the controller. Routing at network layer is performed on the basis of source routing approach. Wavenis is another protocol stack, which was developed by Coronis Systems. Wavenis stack defines physical, data link and network layers [14]. The recent trend is to move towards IP based solutions, wherein internet connectivity is added to smart objects [15]. Due to the huge number of connected devices, a good choice is to use IPv6 based solutions over IPv4. An adaptation layer is needed to use IPv6 over the protocol stack, to perform tasks such as header compression, address auto-configuration etc. 6LoWPAN IETF Working Group defines the format for adaptation between IPv6 and IEEE 802.15.4. 6LoWPAN devices can inter-operate with other IP based devices while ZigBee based devices need a 802.15.4 IP gateway for interoperability.

Our work focuses on the network layer of the protocol stack. The main functions of the network layer are : network establishment, routing, neighbour discovery, adding and removing devices. Our main focus is to reduce the data communication cost as much as possible in order to improve the network longevity, without hampering the network performance, determined by coverage, number of packets sent etc.

To assist data propagation, various network topologies are supported by the network layer in each protocol stack. For example, IEEE 802.15.4 supports four different network topologies i.e. star, tree, cluster-tree and mesh topology whereas ZigBee use star, tree and mesh topology [16]. The star topology is the easiest one to achieve. The coordinator is the central node of the network, end devices communicate directly with the coordinator. There is no router in this topology. The limitation of star topology is that end devices must be in communication range of the coordinator. Also failure of central node (coordinator) will affect the whole network. The tree topology is more flexible as compared to star topology. In this topology, the network scale is extended by using routers as sub-devices. An end device joins the tree via a router, while a router joins the tree via another router(for scalability), or via the coordinator directly. The disadvantage of this topology is that no alternative route is available if any one of the links fails. Also even if two nodes are geographically close to each other, their direct communication is not permitted. In a cluster tree topology, Router along with end devices, is called a cluster. Each cluster is assosiated with a cluster id. End devices communicate with the router, router communicates with the coordinator directly or via another router/ cluster head. Parent node in each cluster acts as an aggregator as well as a router, and sends aggregated data to the coordinator. Levels are defined as distance, in terms of number of hops from the nodes to the sink (coordinator). The coordinator/sink is at level 0, children of coordinator are at level 1 and so on [17]. Mesh topology is similar to tree topology, except the network communications in mesh topology are more flexible. All routers are allowed to communicate with each other directly. The network routing algorithm picks up an alternative path from the available paths when any of them fails. Based on the application requirements, any of the above discussed topologies may be chosen, or a hybrid of them is used.

Based on the above discussed topologies, we can conclude that Direct propagation, chain formation, cluster creation are the various available techniques by which data is communicated by sensor nodes to the sink.

In direct propagation, each node sends its data to the base station (BS) directly. Direct communication is preferred when the location of sink is near the sensing field and the sensing region is small. When the sensing region is large or the sink/BS is located far away from the sensing field (greater than threshold distance d_0 apart), large amount of transmit power will be used by farther nodes to transmit its data, since d is large (from Eq. 1), thus leading to uneven energy consumption (far away nodes die faster than nearby nodes) and hence shorter network lifetime.

The second approach is to use a minimum energy routing protocol where data is transferred to base station through intermediate nodes which act as routers along with sensing to minimize long distance transmissions. However in this Minimum Transmission Energy (MTE) routing protocol it has been found that in transmitting the data through n hops rather than just one high energy transmission the total energy expended might actually be greater than Direct transmission. Also the nodes near to the base station act as routers along with sensing, thus they deplete their energy much faster than the nodes far away, which is usually termed as the hot-spot problem [18].

The solution protocol for load balancing and energy efficient communication is Clustering, in which the nodes organize themselves into clusters which communicate their data to local base station, which further transmits the data to the global base station. Thus number of long distance transmissions are minimized and these local base stations are closer to the nodes. It has been proved that clustering performs well than Direct communication when the base station is far away from the sensing nodes.

The most common clustering protocol is LEACH (Low energy adaptive clustering hierarchy), which uses stochastic election of cluster heads, such that the cluster heads are rotated periodically to balance the load equally among all sensor nodes [6]. There exists a plethora of clustering algorithms in the literature inspired by LEACH. TL-Leach [19], Leach-M [20], Leach-C [8], Leach-DT [21], N-Leach [7], TEEN [22], AP-TEEN [23], M-TEEN [24], HEED [25] are a few examples to cite.

All hierarchical clustering protocols suffer from the problem of early death of cluster heads (due to high load of cluster members). Thus to overcome this, Lindsey et al. [26] proposed Power Efficient Gathering in Sensor Information Systems (PEGASIS). PEGASIS also operates in rounds. At the beginning of each round, all the nodes form one single chain, and any one node among this chain is elected as a leader. Each node communicates only with a close neighbor and takes turns transmitting to the base station (BS), thus reducing the amount of energy spent per round. This algorithm assumes that nodes have global knowledge of the network. That is, every node knows about the location of every other node. Thus PEGASIS is not much scalable. Whenever a node dies in PEGASIS, the entire chain has to be reconstructed. This reconstruction incurs large delays. Thus, it cannot be used for delay sensitive applications. Thus, we can conclude that chaining is not suitable for large scale networks. Hence clustering is an efficient and scalable solution for large scale WSNs [27]. Optimal election of cluster heads is a NP-Hard problem. If the total number of sensors in the area of interest is n, the number of clusters possible is $2^n - 1$. This indicates that the computational complexity to find optimal cluster heads by a brute force technique is tremendous for large scale WSNs. Thus heuristic approach is needed to solve the wsn clustering problem. Also, one of the major concerns in any clustering algorithm is the optimal election of cluster heads that satisfies conflicting clustering goals [28, 29]. For e.g., if we attempt to improve network lifetime, there is a danger that it might increase delay also (not all good things go together). Thus, we need to find a trade-off solution among these conflicting objectives.

Multiobjective optimization algorithms have been used widely in the area of Wireless Sensor Networks to find a trade-off solution. Traditional multi-objective algorithms use weighting methods [30–34] to obtain one single optimum answer by assigning different weights to different objectives and then perform single objective optimization. Weights can be assigned to multiple conflicting objectives through direct assignment, eigen vector method, entropy method and minimal information method, etc. The major drawback of using these methods is that the algorithm needs to be run again every time user preferences (weights) change.

The solution is to use population based methods. The advantage of using population-based multi-objective optimization algorithms is that a set of optimal answers are returned in a single run of the algorithm, instead of a single solution, from which any one can be chosen. Hence we don't need to assign any weights (or priorities) in the beginning. Due to the presence of multiple objectives, we obtain a set of pareto optimal solutions instead of a single solution. The general aim of this research work is to develop an efficient and scale-able clustering strategy for wireless sensor networks, taking into account the various issues which occur in a clustered architecture. Based on the above mentioned improvements, this paper has the following objectives:

- To develop an evolutionary clustering algorithm which obtains a trade off solution among following conflicting objectives i.e energy consumption, network lifetime, coverage and load balancing.
- The algorithm is able to return all the possible solutions in a single run of the algorithm, from which the user can chose any solution from the solution set returned, based on user preferences.

There exists a large number of multi-objective optimization algorithms to solve multiple conflicting objectives. We have used non-dominated sorting genetic algorithm NSGA-II, to solve the WSN clustering problem, because of its fast non-dominated sorting procedure, an elitist-preserving approach, and a parameter-less niching operator [35]. The computational complexity of NSGA-II is $O(m * N^2)$, where m is the number of objective functions and N is the size of population. When the number of objectives and the size of population is large, this can be extremely time consuming. To overcome this limitation, we have used the efficient non-dominated level update mechanism [36]. This mechanism does not apply the non-dominated sorting from scratch each time, rather it updates the non-domination level structure of only a limited number of solutions by knowledge extraction from the current structure. The computational complexity of ENLU mechanism is O(m) in the best case and $O(m * N^2)$ in the worst case. Therefore, in our proposed algorithm, NSGA-II with ENLU mechanism has been considered to obtain pareto optimal solutions with less computational complexity and good convergence and uniform diversity.

The rest of the paper is organised as follows. Section 2 presents a state of the art of cluster based protocols in Wireless Sensor Networks. Preliminaries are given in Sect. 3. Our Proposed algorithm is presented in Sect. 4. Section 5 presents the simulation model. Results are presented in Sect. 6. Finally, the concluding remarks and future scope is presented in Sect. 7.

2 Related work

A number of clustering protocols have been proposed in the last two decades. Heinzelman et al. proposed LEACH [6], which is one of the earliest and renowned clustering protocol in wireless sensor networks. The sensing region is divided into various clusters. The protocol operates in rounds. Each round is composed of two phases: setup phase and steady state phase. Clusters are formed in the setup phase, while data transmission is done in the steady state phase.

In the set up phase, each node first computes a threshold value, and then computes a random number between 0 and 1. The threshold is specified in Eq. 4 as:

$$T(n) = \begin{cases} \dfrac{p}{1 - p * r(mod\dfrac{1}{p})} & if\ n \in G \\ 0 \quad otherwise \end{cases} \tag{4}$$

where n represents the sensor node number, p represents predefined percentage of cluster heads (e.g. $p = 0.05$), r represents the current round number and G represents the set of nodes which have not been cluster heads in the last $\frac{1}{p}$ rounds. If the generated random number has a lesser value than the threshold value, then the node is elected as a cluster member. This is done to ensure that each sensor node becomes a cluster head in some round within $1/p$ rounds. Once a node becomes a cluster head, it is again eligible to become a cluster head after every $1/p$ rounds. Thus, rotation of cluster heads is done to achieve load balancing. After a node is elected as cluster head, it broadcasts an advertisement message to all its neighbours. Non cluster head nodes chose the nearby cluster head based on RSSI (Received Signal Strength Indicator) value. Each cluster head then computes a TDMA (Time Division Multiple Access) schedule for its cluster members, assigning a time slot to each member. In the steady state phase, each non-cluster head node senses the phenomenon of interest and then sends this data to their respective cluster head. Each cluster head then aggregates the data received along with its own sensed data and sends it to the base station in a single hop.

However, LEACH has a few drawbacks. Due to random election of Cluster heads, low energy nodes may be elected as cluster heads. Due to heavy toll on cluster heads, these low energy nodes die out soon, leading to short network lifetime. Also, since the percentage of cluster heads remain fixed during the entire network operation, it may happen that no cluster head is elected in the later rounds of operation. This leads to no data being sent to the base stations during these rounds.

A number of refinements of LEACH have been proposed. LEACH-C (centralized) [8] is refinement of LEACH, based on centralized election of cluster head nodes. Higher energy nodes are elected as cluster heads by the base station to improve the network lifetime. Hybrid Energy Efficient Distributed clustering (HEED) protocol attempts to further improve network lifetime by incorporating residual energy as primary parameter and communication cost as secondary parameter for cluster head election. HEED ensures high energy cluster heads as well as non overlapping clusters. Fault tolerance is another important factor to be considered while designing protocols for wireless sensor networks, since WSNs are often subject to failures caused by energy depletion of nodes, software or hardware failure of nodes, environmental events etc [37]. LEACH-VH (LEACH with Vice Cluster head) protocol [38] is another refinement of LEACH to improve the network lifetime and ensure fault tolerance, wherein a vice cluster head is elected along with cluster head. Node with highest and second highest residual energy are elected as cluster head (CH) and vice cluster head (VH) respectively. The VH remains in sleep mode initially and when the energy of cluster head falls below a specific threshold, the VH wakes up to act as the CH. Number of clusters created is another important factor which affects the network lifetime. Centralised energy efficient distance based routing protocol for wireless sensor networks (CEED) [39] elects the optimum number of cluster heads mathematically by considering the energy dissipated in setup and steady state phases. Cluster head election is then done by including residual energy and distance of node to base station as cluster head election parameters. Multi-hop routing is performed between cluster heads in CEED. Recently, Balanced Energy and Adaptive Cluster Head Selection Algorithm for Wireless Sensor Networks (BEACH) [40] was proposed. Equal size clusters are created in this algorithm, to balance the load equally among all the clusters. Also multi-hop communication based on river formation dynamics is introduced between the cluster heads and the base station to reduce data communication cost. Mittal et al. proposed a stable energy efficient clustering protocol for wireless sensor networks [41], which aims to increase the stability period of the network. The protocol balances the load among the nodes and hence ensures higher stability. Another factor to be considered while setting up data communication path in clustering protocols is security. The authors in [42] propose an Instantaneous and Secure Clustering protocol (ISCP) for WSNs wherein, solution to the followings attacks are proposed for a cluster based protocol: Sinkhole attacks, Selective forwarding, Sybil attacks and Hello flood attack. Recently, a large number of algorithms have been developed by various researchers to ensure security in cluster based WSN [43–46].

The WSN algorithms often have conflicting goals. To find a trade-off solution among conflicting goals, Optimization algorithms have been used in Wireless sensor networks. Peiravi et al. developed M2NGA, a general Genetic Algorithm-based clustering algorithm, which simultaneously optimizes network lifetime and delay [47]. Cheng et al. proposed a framework for multi-objective optimization for clustered wireless sensor networks.The objective functions used are the energy consumption and duration of a data collection process (DCP). The proposed algorithm provides reasonable trade-off between these two conflicting objectives. Ozdemir et al. [48] developed multi-objective optimization algorithm to simultaneously optimize coverage and energy in sensor networks based on clusters using multi-objective evolutionary algorithm based on decomposition (MOEA/D). Ozdemir et al. [49] The authors in [50] have provided a review of multi-objective optimization techniques to solve various conflicting issues of wireless sensor networks related to the design, deployment, operation, placement, planning and management. Seri et al. proposed MOFCA (Multi-objective fuzzy clustering algorithm) for calculating the cluster head competition radius.The parameters considered were residual energy, distance and node density [51]. Authors in [52] proposed multi-objective clustering in wireless sensor networks using NSGA-II(Non dominated sorting genetic algorithm). Various energy parameters are taken as cost functions. The authors have shown that the proposed algorithm has five times longer lifetime and can transmit two times more packets compared to LEACH. Jameii et al. [53] proposed AMOF(Adaptive multi-objective optimization framework) for coverage and topology control in wireless sensor networks. The algorithm consists of two parallel modules-LA(learning automata) and NSGA module to handle several conflicting objectives in heterogeneous networks. Mazumdar et al. [54] proposed DFMCA (distributed fault tolerant multi-objective clustering algorithm). In this algorithm, each sensor takes its decision using local information only.

Unlike the above mentioned algorithms, this paper proposes the utilization of Non dominated sort genetic algorithm (NSGA-II) accompanied by Efficient non dominated level update (ENLU) mechanism to solve the wsn clustering problem considering four conflicting objective functions of energy conservation, network lifetime enhancement, coverage enhancement and load balancing simultaneously.

Table 1 summarises the clustering algorithms mentioned above on the basis of clustering objectives, strategies and parameters for cluster head election.

3 Preliminaries

In this section, we present the mathematical models used in representation of Wireless sensor networks, assumptions for simulation and problem definition.

3.1 Network model and assumptions

A wireless sensor network can be illustrated as a directed graph $G = (V, E)$, where V represents the set of sensor nodes, and E represents the links between the nodes, which depends upon the transmission range of sensor node. Each node has a unique id i and a node is represented as $node(i)$.

For simulation of a wireless sensor network, the following assumptions have been taken.

1. All the nodes as well as the base station are stationary.
2. Sensor nodes have a uniform random deployment in the network.
3. Sensor nodes batteries are irreplaceable while the base station has an infinite reserve of energy.

Table 1 Classification of clustering protocols

Cluster based protocol	Clustering strategy	Clustering objectives	Parameters for cluster head election
LEACH [6]	Probabilistic	None	None (random election)
HEED [25]	Probabilistic	Lifetime, coverage	Residual energy (primary), separation between nodes (secondary)
CEED [39]	Probabilistic	Energy, lifetime	Residual energy, distance to base station
Genetic algorithm based routing (GAR) [33]	Genetic algorithm (GA) based (single objective optimization)	Lifetime	Distance to base station
Kuila et al. [29]	GA based (single objective optimization)	Load balancing	Standard deviation of cluster head load
LEACH-DT [20]	Probabilistic	Energy	Distance to base station
LEACH-M [21]	Probabilistic	Energy	Residual energy, network address
M2NGA [47]	GA based (multi-objective optimization)	Delay, network lifetime	Energy, delay
Hacioglu et al. [52]	GA based (multi-objective optimization)	Energy, lifetime	Distance to base station, residual energy

4. All the nodes are homogeneous (i.e. all have the same initial energy)
5. Each node has the capability to aggregate data, received from its neighbors.
6. Nodes are not equipped with any GPS module.
7. The nodes can judge the distance from the signal source through the strength of the signal, and adjust the transmission power.
8. All error handling is done at the MAC layer.

3.2 Energy consumption model

Energy model used is first order radio energy model. The transmission and reception energies are given in Eqs. 1 and 2. In this paper, $Eelec = 50nJ/bit$, $Efs = 10nJ/bit/m2$, $Emp = 10nJ/bit/m4$. Each node is capable of doing data aggregation. Nodes send their respective data to the cluster heads. Cluster heads aggregate the data received and forward a single aggregated packet. Thus fewer packets need to be forwarded, leading to significant decrease in energy consumption. The data aggregation energy, $E_{da} = 5nJ/bit/signal$

The first order radio model and the parameters are chosen so as to make the comparisons simpler with existing algorithms.

3.3 Data aggregation model

For data aggregation, the infinite compressibility model is used [55]. Cluster heads receives data from its members and then aggregates the data received into a single packet of fixed length irrespective of the cluster size.

3.4 Challenges

Appropriate election of cluster heads with high residual energy, election of cluster heads such that no two cluster heads are in the communication range of each other, minimum energy clustering and load balancing among the cluster heads so as to increase network lifetime and to avoid creation of hot-spots are the various objective functions used. Our proposed algorithm handles all of these multiple objectives simultaneously. We considered the problem of selecting appropriate cluster heads in a wireless sensor networks which involves optimizing four conflicting criterion simultaneously as mentioned below :

- Maximize residual energy of cluster heads
- Minimize sum of distances between cluster heads and sink.
- Maximize separation between adjacent cluster heads.
- Minimize deviation of cluster load.

3.5 Concept of pareto dominance

For a problem with multiple objectives that might conflict with each other, finding a single best solution is not a feasible approach. It is imperative to use the concept of non-dominance to generate a set of optimal solutions, which can then be used to find a solution as per the practical significance.

A given solution x dominates, another given solution y, if all the values of x are better($> =$ or $< =$ depending on whether the problem is of maximization or minimization) than those of solution y. If x does not dominate y and y does not dominate x, then the solutions are said to be non-dominating. The pareto front is defined as the set of all the non-dominated solutions which are found from the solution space. The Pareto dominance relation is defined as:

Let us assume that $f(x) = [f_1(x), f_2(x), \ldots, f_m(x)]$ be the given m multiple objectives which are to be optimized. A solution x_1 dominates solution x_2 iff Eq. 5 holds:

$$\begin{aligned} &\forall i \in [1, 2, \ldots, m] : f_i(x_1) < = f_i(x_2) \\ &\in j \in [1, 2, \ldots, m] : f_j(x_1) < f_j(x_2) \end{aligned} \tag{5}$$

From the multiple solutions returned, the user can choose any of the solution from the set based on its preferences instead of assigning predefined weights as per the preferences beforehand.

4 Proposed evolutionary multi-objective optimization using NSGA-II-ENLU

There exists a multitude of evolutionary algorithms in the literature to solve the multi-objective optimization problem. In our proposed work, a multi-objective optimization algorithm, Non Dominated Sorting Genetic Algorithm-II [35] is used for electing cluster head and non cluster head nodes based on the objective functions mentioned in section 3 above. The procedure of NSGA-II is explained with the help of Fig. 1. The algorithm is chosen because of its fast non-dominated sorting technique, an elitist preserving technique, and niching operator which do not require any algorithm specific parameters.

The NSGA-II with ENLU [36] mechanism based cluster head election algorithm is described in algorithm 1.

Algorithm 1 : NSGA II with ENLU based multi-objective optimization algorithm for clustering in wireless sensor networks.

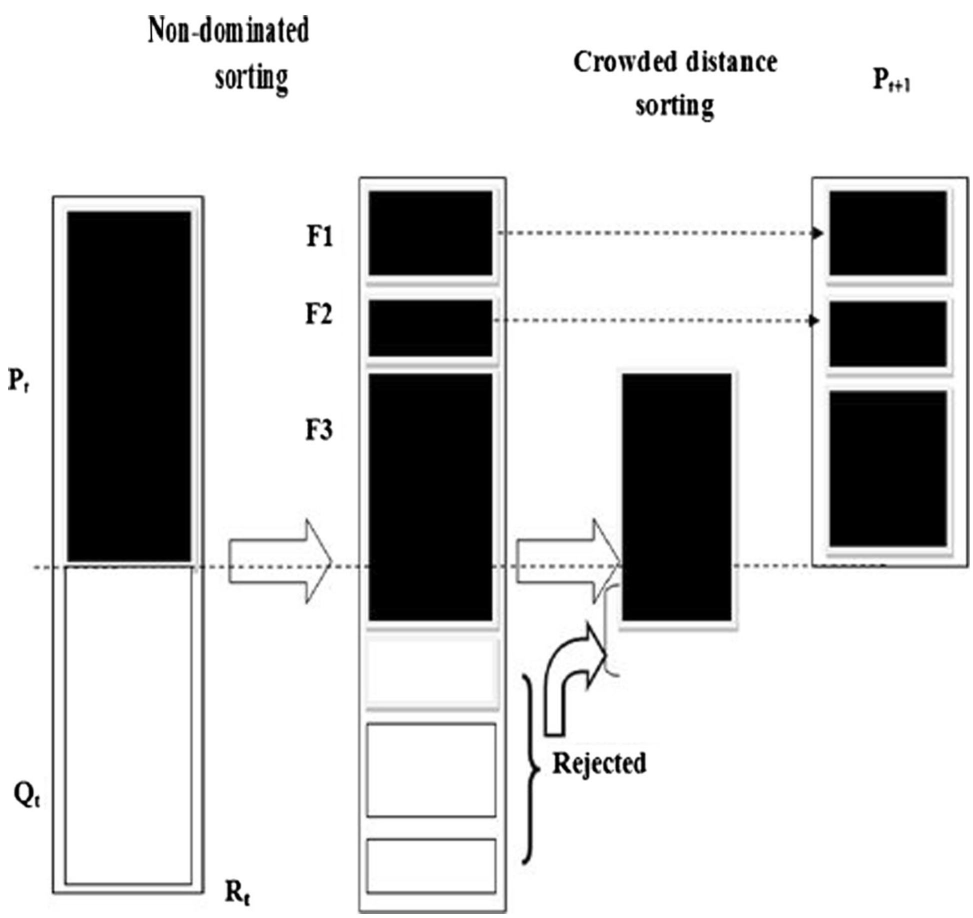
Input : Size of population (N), m objective functions $[f_1(x), f_2(x), \ldots, f_m(x)]$, maximum number of generations(t)

Output : Solution set S

Steps :

1. Randomly create initial population $P(0)$ of size N.
2. Allocate fitness values to each individual of Population $P(0)$ as per the four conflicting objective functions : residual energy of cluster heads, distance between cluster heads and sink, separation between cluster heads, deviation of cluster load, given by Eqs. 5–8.
3. Apply the fast-non-dominated sorting procedure on the population $P(0)$ to generate Pareto Fronts $(F_1, F_2, \ldots, F_l)$ for the individuals.
4. Initialize $t = 0$;
5. repeat

Fig. 1 NSGA-II procedure



Algorithm 1 : NSGA II with ENLU based multi-objective optimization algorithm for clustering in wireless sensor networks.

6. Use defined genetic operators for selection, crossover and mutation to create an offspring population $Q(t)$ from the population $P(t)$.
7. Update the non domination level structure by adding $Q(t)$ to the appropriate pareto front using Efficient non domination level update method given by algorithm 2 .
8. Remove worst solutions after applying the update method.
9. Until ($t <$ maximum number of generations).

The Efficient non dominated level update method mentioned in step 7 is described in algorithm 2 below.

Algorithm 2 : Efficient Non-domination Level Update (ENLU) method

Input : pareto fronts of current population, offspring population

Output : pareto fronts of combined population S

Steps :

1. For $i = 1$ to l
2. Analyze the dominance relation of Q_t with solutions of F_i
3. If Q_t is non dominated with all the solutions of F_i or is dominating some solutions in F_i then
4. Add Q_t to F_i and move to another non domination level F_j such that $j > i$.
5. Break
6. Else if Q_t dominates all solutions of F_i then increase the non domination level of each front by 1.
7. End for

The methodologies for population representation, fitness evaluation, genetic operators-selection, crossover and mutation, fast-non dominated sorting, update method are described in the subsections.

4.1 Chromosome representation

To resolve the cluster head election problem in a wireless sensor network, an individual in the population is represented as a string. We use the binary method. A chromosome is represented as a string of 0 and 1s where a 0 indicates that the node is a non-cluster head/member node and a 1 indicates that the node is a cluster head node. The length of each chromosome is equal to the number of sensor nodes. The following example illustrates this.

Example 1 Consider a wireless sensor network with 10 sensor nodes. $S = s1, s2, \ldots, s10$ represents the set of sensor nodes. Thus the length of each chromosome is 10. Table 2 shows the corresponding representation of the chromosome in which the gene value 1 at position 4 and at position 7 indicates that the sensor nodes $s4$, $s7$ are cluster head node, while all the remaining nodes with gene value 0 are non cluster head/member nodes.

This string is generated randomly while taking into consideration the application constraints.

4.2 Initial population generation

We create the initial population by generating N number of chromosomes, with each chromosome being a string of 0s and 1s generated randomly of length equal to the number of sensor nodes (Table 3).

Example 2 Consider a Wireless sensor network with 10 sensor nodes. After generation of initial population, fitness functions are used to evaluate the fitness of these individuals. Four fitness functions are used, residual energy of cluster heads, cluster head to sink distance, separation between cluster heads, deviation of cluster load, as given by Eqs. 5–8 in the next section.

4.3 Objective functions

- Residual energy of cluster head nodes

  The energy expended by a cluster head includes the energy consumed in receiving data from all the cluster members, aggregating the data received and sending the aggregated data to the sink. Higher energy nodes should be chosen as cluster heads. For a set of solutions, a chromosome having higher residual energy for the cluster heads is considered a better solution. In any chromosome, the sum of residual energy of cluster heads is denoted as:

$$res_{ch} = \sum_{i=1}^{n} node(i).e \tag{6}$$

  if $node(i)$ is a cluster head and $node(i).e$ represents the residual energy of $node(i)$
- Distance of cluster heads to sink

  Cluster head nodes send their aggregated data to the base station. Since the energy consumed in transmission

Table 2 Chromosome representation

Sensor node number	1	2	3	4	5	6	7	8	9	10
Sensor node type	0	0	0	1	0	0	1	0	0	0

Table 3 Initial population generation

Sensor node number	1	2	3	4	5	6	7	8	9	10
P_1	0	0	0	1	0	0	1	0	0	0
P_2	0	0	1	0	0	0	1	1	0	0
...	...	...	...	...	...	...	...	...	...	...
P_N	0	1	0	0	0	1	1	0	0	0

of this data is proportional to the distance, this distance must be minimized. In any chromosome, the sum of the distance of cluster heads to sink is denoted as:

$$dss_{ch} = \sum_{i=1}^{n} node(i).d \tag{7}$$

if $node(i)$ is a cluster head and $node(i).d$ represents the distance of $node(i)$ from the base station.

- Separation between cluster heads

  In order to ensure maximum coverage, the separation between any two cluster heads should be as large as possible. The cluster head separation can be calculated as:

$$sep = max(d(node(i), node(j))) \tag{8}$$

  where $node(i)$ and $node(j)$ are any two cluster heads.

- Deviation of cluster load Each node calculates its deviation from the ideal cluster head load. This deviation is calculated as:

$$dev = \sum_{i=1}^{n} (|l_i - \mu|)/\mu \tag{9}$$

  if $node(i)$ is a cluster head Here l_i is the load of cluster head i and μ is the ideal load for cluster head i and is defined in Eq. 9 as:

$$\mu = n/m \tag{10}$$

  where n is the total number of sensor nodes in the region of interest and m is the total number of cluster head nodes expected.

After evaluation of the fitness values, genetic operators are applied on the initial population (P_0)), to create the offspring population (Q_0).

4.4 Genetic operators

4.4.1 Selection

Selection operator is used to select those chromosomes which will be included in the next generation, by selecting better chromosomes (with higher fitness). We use Binary tournament selection [56] in our work because of its wider use and acceptability to efficiently select better solutions.

After selection process is completed, crossover operator is used as explained below:

4.4.2 Crossover

The crossover operator creates new solutions from the existing solutions by exchange of information between the solutions. It selects any two chromosomes randomly from the population and some portion of the strings is exchanged between the strings. The crossover point is also selected randomly. A probability of crossover is also introduced in order to give freedom to an individual solution string to determine whether the solution would go for crossover or not.

We have used single point crossover in our paper with crossover rate 0.5 as it is expected to give a faster convergence to the solution as per the problem space described. Crossover operator is illustrated with Example 3 below.

Example 3 Let us assume a wireless sensor network having 10 sensor nodes. The crossover point in this example is taken to be 4. Rest of the string after this point is exchanged between the two chromosomes. This is illustrated in Fig. 2.

The next step is mutation.

4.4.3 Mutation

Mutation is the introduction of new features into the population to maintain diversity in the population. The Mutation operator randomly selects a node in the string and

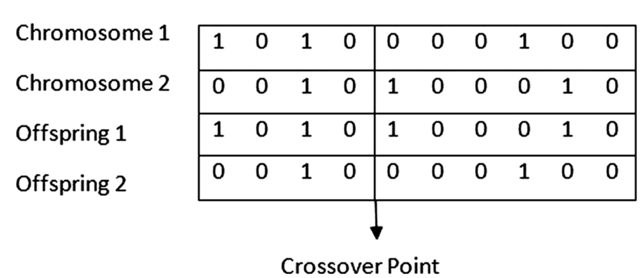


Fig. 2 Single point crossover

changes its type i.e it changes a cluster head to non-cluster head and vice verse. The mutation probability is generally kept low for steady convergence. A high value of mutation probability would search here and there like a random search technique. In this paper, we have used flip bit operator with mutation rate 0.05 [56]. Mutation operator is illustrated in the below example in Table 4:

Example 4 Assume a Wireless sensor network having 10 sensor nodes.

The initial population P(0) is sorted according to non-dominated sorting procedure given by deb et al(2002) in NSGA-II.

4.5 Fast non dominated sorting of initial population P(0)

P(0) is sorted according to non-domination into various fronts $F_1, F_2, \ldots$

Example 5 The example illustrates the concept of non-domination for our algorithm. We have four objective functions $obj_1 \ldots obj_4$

As seen in Table 5, the solution i dominates solution j, if the following relation is satisfied:

$obj_{ni} <= obj_{nj}$, where $n = 1$ to 4

The solutions which are not dominated by any other solutions are included in the first front F_1. The second front, F_2 consists of solutions which are only dominated by the solutions of the first front. The other fronts are also formed similarly until all the members of initial population are included. We have used fast non dominated sorting procedure for sorting of initial population only.

After the off-springs are generated, we use NSGA-II ENLU to add it to the appropriate pareto front because of its ability to cut off a significant amount of unnecessary comparisons which occur if we perform fast non dominated sorting each time a solution is generated. After the pareto fronts are updated due to the addition of current off-springs, the size of new population, (P_{t+1}) becomes $2N$. Since we can include only N members, the worst solutions are removed from the last front using crowded distance approach as demonstrated in Fig. 1.

Table 4 Flip bit mutation

Offspring1	1	0	1	0	0	0	0	1	0	0
Offspring2	0	0	1	0	1	0	0	0	1	0
Mutated 1	1	0	0	0	0	0	0	1	0	0
Mutated 2	0	0	1	0	0	0	0	1	1	0

Table 5 Non-dominated solution

Chromosome	obj_1	obj_2	obj_3	obj_4
i	obj_{1i}	obj_{2i}	obj_{3i}	obj_{4i}
j	obj_{1j}	obj_{2j}	obj_{3j}	obj_{4j}

4.6 Crowded distance sorting

We have to select exactly N members for the next generationP_{t+1}. If all the members of the last front F_l to be included cannot be added completely to P_{t+1}, these solution are sorted according to crowding comparison operator in descending order and then the remaining slots are filled. Each solution in this front f_l, is sorted according to each of the four objective functions. The solutions which lie at the boundaries are assigned a crowding distance of ∞.

Example 6 Let individual j and k get the highest and lowest value according to obj_1. The distances of these two individuals are taken as ∞ Let $i-1, i, i+1$ be other three solutions belonging to the same front. The distance for solution i is calculated as:

$$d_{i,1} = \frac{|obj_{1,i-1-obj1,i+1}|}{obj_1^{max} - obj_1^{min}} \tag{11}$$

Similarly, the values of $d_{i,2}, d_{i,3}, d_{i,4}$ are calculated. The crowding distance for this individual is then taken as:

$$cd_i = \sum_{z=1}^{4} d_{i,z} \tag{12}$$

The crowding distance for all the other solutions are calculated in the similar manner. The solutions are then sorted based on crowding distance and the solutions with larger crowding distance are chosen.

Thus, we can conclude that between any two solutions, the one having the lower rank is preferred, and between two solutions with the same rank, the solution belonging to less crowded region is preferred. The population returned after the crowded distance sorting is considered as the population corresponding to next iteration P_{t+1} From P_{t+1}, Q_{t+1} is generated in the similar manner using selection , crossover and mutation, and the process continues until the required number of generations.

The flowchart of the proposed algorithm is given in Fig. 3.

The flowchart is explained as follows : Initially, the proposed NSGA-II with ENLU based algorithm is run at

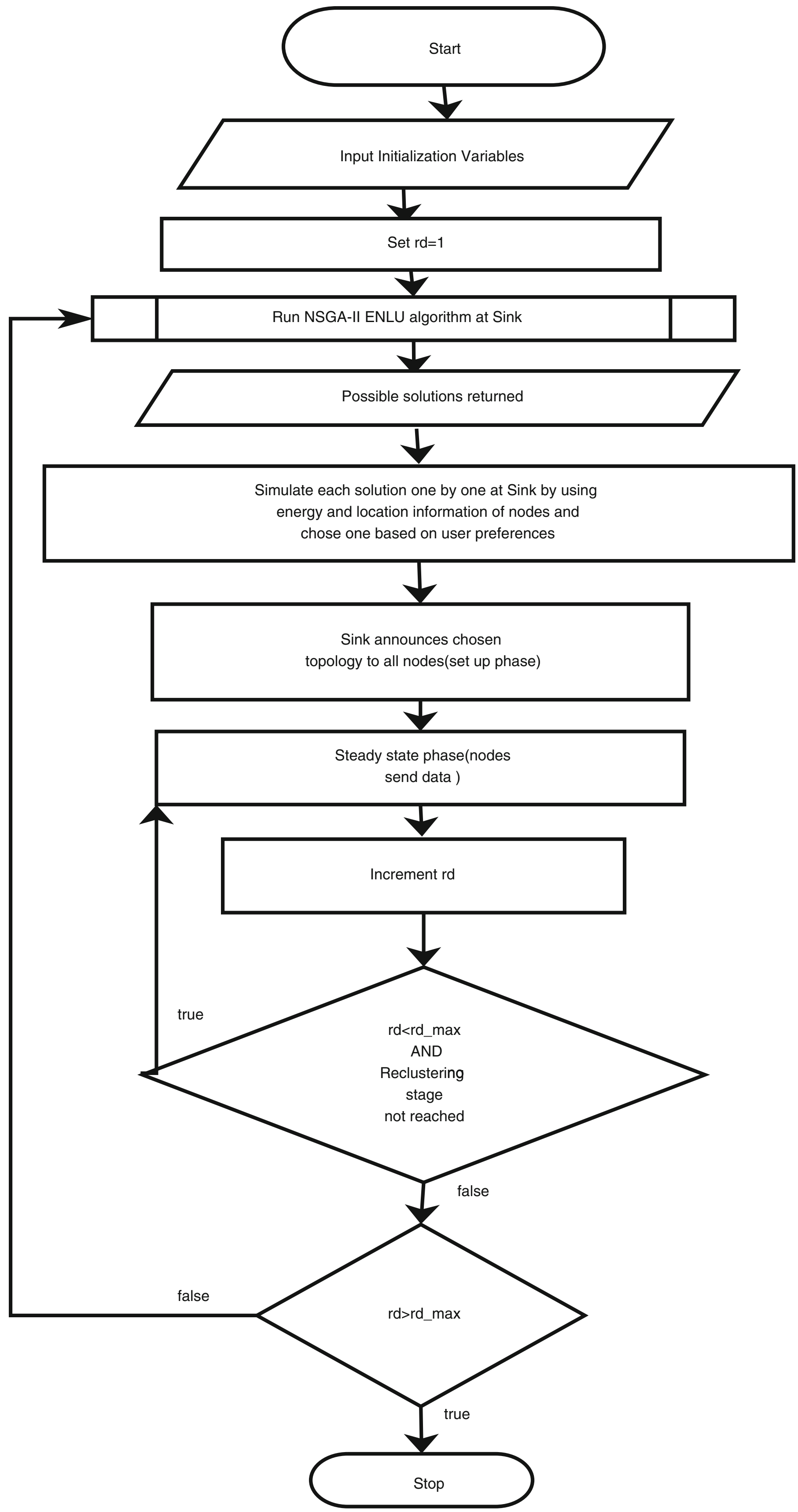


Fig. 3 Flowchart of proposed algorithm

the Sink node. Multiple non-dominating solutions are returned by the algorithm. Each solution is run at the Sink node and various network performance parameters are obtained. A single solution is chosen among them based on user preferences of network parameters, and the corresponding network topology is broadcast by the Sink to the nodes. This is done only for the first round of operation. For the next corresponding rounds, the topology remains fixed as announced in the first round. Nodes send their data in each round to their cluster heads based on TDMA schedule. Cluster heads then forward their data to the Sink node. This process continues until the re-clustering stage (once a cluster head dies or after every *m* number of rounds) is not reached. If the re-clustering stage is reached, nodes send their topology to the Sink node and the proposed algorithm is run again. This process continues until the maximum rounds (as defined initially) have been completed.

5 Simulation model

5.1 Simulation parameters

For our experiments, the sensors are deployed uniformly in a two dimensional area $200 \times 200m^2$. The sink node is placed at a location far away from the sensing field (100, 300) (since clustering is preferred when the base station is far away).

We use MATLAB for simulation. The radio parameters are described in Table 6:

The parameters used for simulation of proposed evolutionary multi-objective nsga-ii based approach are given in Table 7

Table 6 Radio model parameters

Parameter	Value
Initial energy of a sensor(E_{init})	$0.02J$
Free space energy dissipation(E_{fs})	$10pJ/bit/m^2$
Multi path energy dissipation(E_{mp})	$0.0013pJ/bit/m^4$
Threshold distance(d_0)	$87.7m$
Data packet length	2000*bits*
Data aggregation energy dissipation	5*nJ*/*bit*/*signal*
Transmitter electronics energy dissipation	50*nJ*/*bit*
Receiver electronics energy dissipation	50*nJ*/*bit*

Table 7 NSGA-II parameters

Parameter	Value
Size of initial population	50
Maximum number of generations	100
Initial population	Random generation
Selection operator	Binary tournament selection
Crossover operator	Single point crossover
Crossover rate	0.6
Mutation operator	Flip Bit
Mutation rate	0.05

5.2 Single objective versus multi-objective optimization

For choosing an optimal clustering topology, we often have multiple conflicting objectives. Optimizing one of the solution leads to degradation in the other objectives and vice-verse. In our clustering algorithm, we have taken four conflicting objectives viz. high residual energy of the cluster head nodes, minimum communication energy, load balancing of the network and coverage. Optimizing one of them leads to poor performance in the other objectives, which may affect the overall network performance. We explain this through the following examples.

Example 1 If we consider minimum energy consumption as our only objective function.

Figure 4 shows the topology obtained, if we consider minimum energy consumption as our only objective function.

Through the resulting topology obtained in Fig. 4, we can conclude that

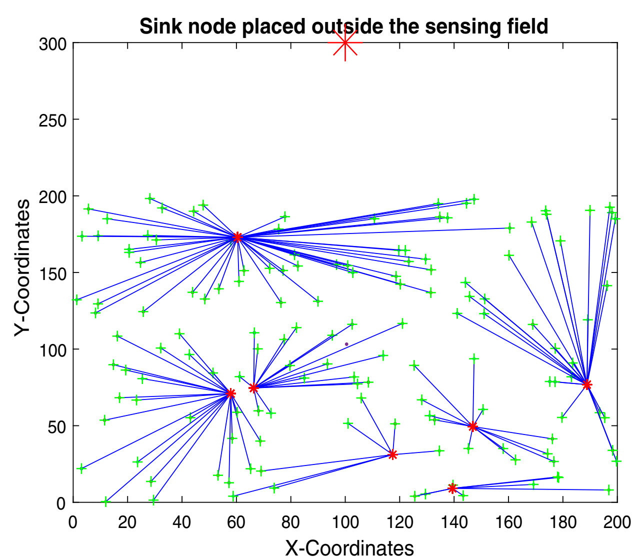


Fig. 4 Objective : minimum energy consumption

1. the nodes may not be load balanced in the obtained topology.
2. Cluster heads may be in each others communication range.

Example 2 If we consider load balancing of the cluster heads as our only objective function.

Figure 5 shows the topology obtained using the above mentioned objective function.

Through the topology obtained in Fig. 5, we can conclude that

1. the resulting topology may not lead to minimum energy consumption.
2. Cluster heads may be in each others communication range.

Example 3 If we consider coverage of the network as our only objective function.

Figure 6 shows the topology obtained using the above mentioned objective function.

Through the topology obtained in Fig. 6, we can conclude that

1. the resulting topology may not lead to minimum energy consumption.
2. The nodes may not be load balanced in the obtained topology.

Thus, we can conclude that single objective optimization cannot give us desirable performance. We need to consider all the multiple objectives in order to obtain the desired results. The problem with multiple objectives is that these objectives are often conflicting. Thus there may exist multiple optimal solutions to a problem rather than a single solution. Each of these optimal solution may give different preference value to each of the objective functions. Using NSGA-II, we can obtain all these optimal solutions in a single run. We can chose any one of these possible solutions based on our network requirements which is not possible by weighted approach.

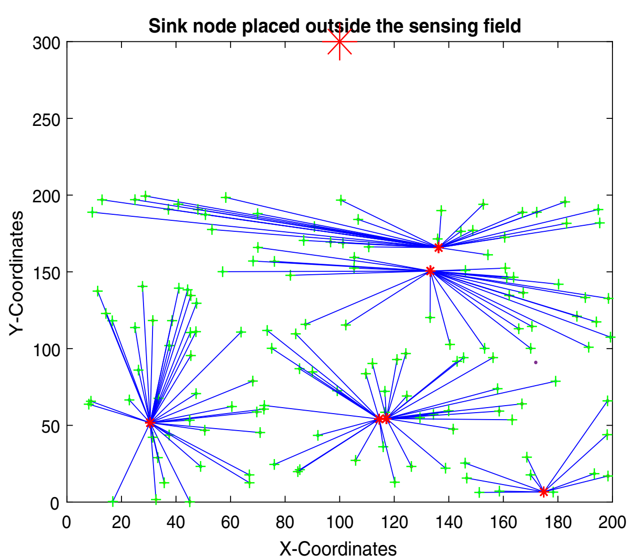


Fig. 5 Objective: standard deviation

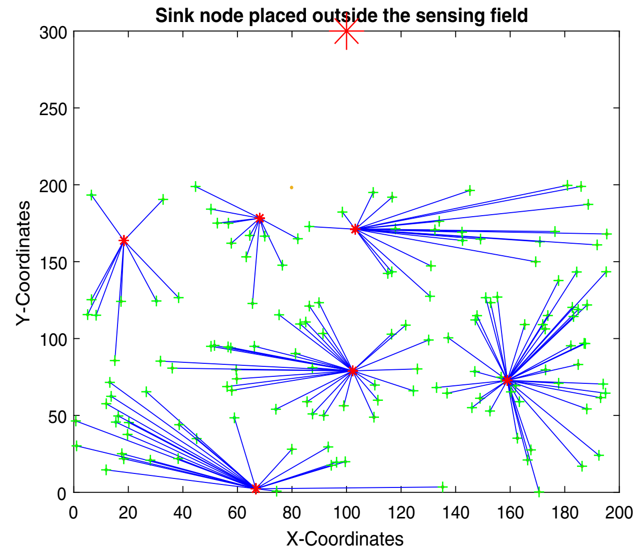


Fig. 6 Objective: maximum coverage

6 Results and discussion

We simulate our proposed method for single hop communication networks. For single hop networks, the results are compared with LEACH and direct communication. In our experiments, two scenarios are considered

- Re clustering is done every time a cluster head dies (until k percent nodes have not died).
- Re clustering is performed after every m rounds.

Lifetime of the network and the total number of packets received, are calculated for both the above scenarios. The clustering topology used is Single hop clustering. Flowchart in Fig. 3 represents the steps involved in the clustering process. Nodes are grouped into cluster head and non cluster head nodes.

We assume that the sink node has the information about the energy of all the nodes and their location information. NSGA-II with ENLU algorithm is implemented at the sink/base station node. From the solution set, one of the solutions is chosen by the sink. The selected topology is broadcast to all the nodes. Cluster heads chosen are broadcast by the sink.Non cluster heads join the nearest cluster head. This is termed as the set up phase. In the steady state phase, data collection takes place. Cluster members send their respective data to the corresponding cluster heads following a Time division multiple access

(TDMA) schedule and further these cluster heads aggregate the data received and send it to the base station following Code division multiple access (CDMA) schedule. In order to reduce clustering overhead, re-clustering is not performed in every round. After re-clustering, the new topology is again broadcast by the sink node and this process continues until the network is in operation or the maximum number of rounds have not been completed.

For our simulations, number of nodes are assumed to be 150 and the location of sink node is taken as (100, 300) (Fig 7).

6.1 Scenario 1

For the first scenario, we perform re-clustering every time a cluster head dies until 10 percent of the nodes are dead.

6.1.1 Network lifetime comparison

The analysis of network lifetime is done on the basis of following definition: Network Lifetime is defined as the time elapsed between the first transmission in the wireless sensor network and when the percentage of sensor nodes which have not terminated their residual energy falls below a specific predefined threshold, set according to the type of application (100 percent or less).

In Fig. 8, we have plotted the number of alive nodes in each round against the number of rounds. During simulation, the values of the parameter p considered for LEACH is $p = 0.05$. It is observed that using NSGA-II with ENLU increases the lifetime of the network when compared with Hacioglu et al., traditional LEACH and DIRECT communication protocols respectively. Extending lifetime is

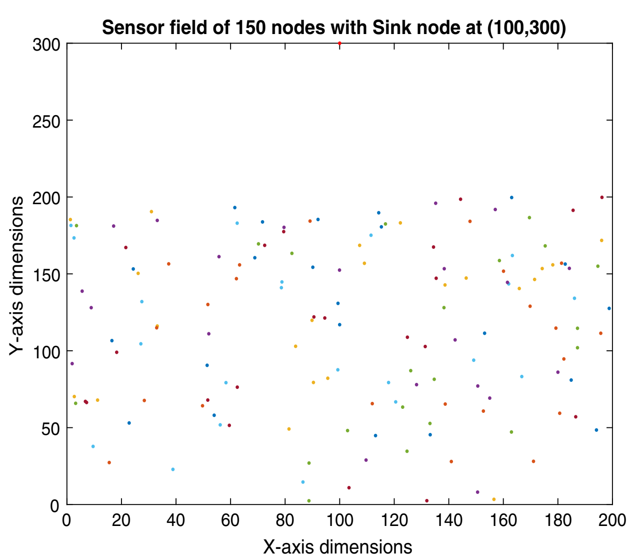


Fig. 7 Network model

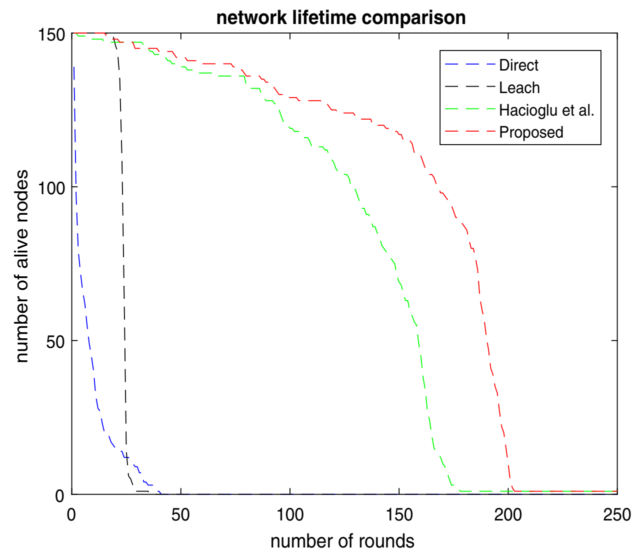


Fig. 8 Network lifetime comparison

achieved at the cost of properly balanced clustering in the proposed algorithm.

6.1.2 Comparison of percentage gain in network lifetime

We consider the data obtained from Fig. 8 in order to analyze the percentage gain of proposed algorithm against LEACH. The simulation is performed with homogeneous sensor nodes having probability of Cluster head p as 0.05 for LEACH.

The percentage gain compared to Hacioglu et al. is calculated as

$$gain\% = \frac{202 - 174}{174} \times 100 = 16$$

The percentage gain compared to LEACH is calculated as

$$gain\% = \frac{202 - 35}{35} \times 100 = 477$$

Similarly, the percentage gain compared to Direct propagation is calculated as

$$gain\% = \frac{202 - 28}{28} \times 100 = 621$$

6.1.3 packets received comparison

The total number of packets received by the base station in DIRECT communication, LEACH, Hacioglu et al. and our Proposed algorithm as depicted in Fig. 9 are 1074, 191, 1005 and 1773 respectively. It can be seen that the number of packets received by DIRECT communication is larger than those obtained in LEACH. This is due to the fact that every node transmits to the base station in DIRECT

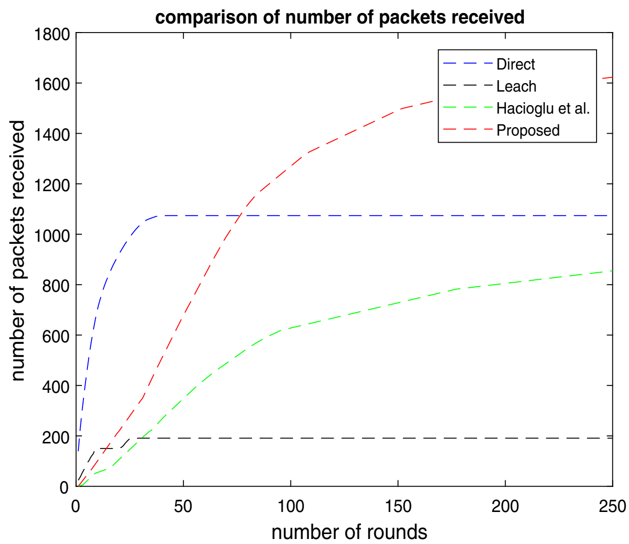


Fig. 9 Number of packets received at base station

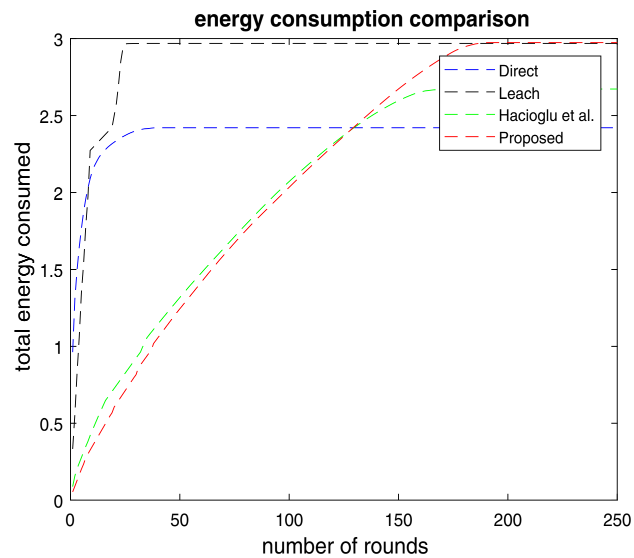


Fig. 10 Energy consumption comparison

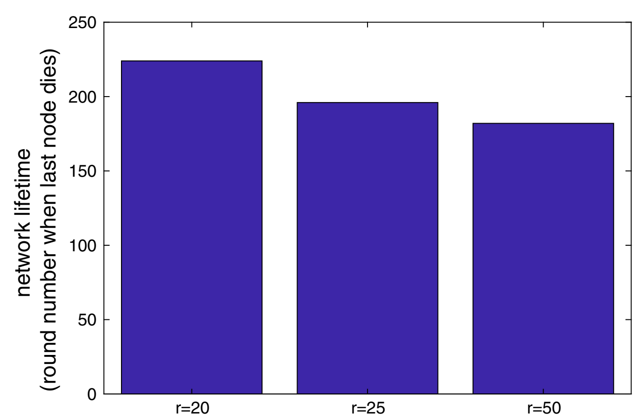


Fig. 11 Network lifetime comparison

communication. Thus there is no data aggregation in DIRECT communication and hence large amount of data is sent in DIRECT communication. While in LEACH, only the cluster head nodes transmit the data. Thus comparatively less data is communicated to the sink. This is clearly depicted in the Fig. 9. However due to stochastic clustering in LEACH, it does not take into account any factors like residual energy, number of nodes etc in cluster creation. This accounts for poor clustering in LEACH. Hacioglu et al. does not consider the issue of overlapping clusters and load balancing among the nodes which leads to early dying of nodes. Our proposed algorithm takes care of these issues.

6.1.4 Comparison of energy consumed

The total energy consumption by DIRECT communication, LEACH, Hacioglu et al. and our proposed algorithm is depicted in Fig. 10. The initial energy in all the four cases for each node is assumed to be 0.02 J. The number of nodes in the sensing area are assumed to be 150. Thus the total amount of energy available is

$$0.02 \times 150 = 3\,J$$

From the figure the energy consumed by the four algorithms are 2.4197, 2.9685 and 2.6715 and 2.9837 respectively. Thus although $3 - 2.4197 = 0.5803$ J energy still remains in DIRECT communication, there is no data received by the base station. This is due to the fact that the distance of nodes from the base station is large and as such,they are not able to forward any data directly. Thus although there is energy available at nodes, but it is not enough to make large distance communication. This problem is eliminated in clustering algorithms like LEACH, wherein data is communicated through intermediate nodes called as cluster heads to reduce the energy consumed in transmission, as transmission energy is proportional to the distance between transmitter and receiver (Eq. 2). In Our proposed algorithm, the remaining energy is lesser than that in LEACH and Hacioglu et al.. This is because of the fact that cluster heads are well separated in our proposed algorithm which reduces the distance between nodes and their corresponding cluster heads. Thus we can conclude that our proposed algorithm handles energy more efficiently.

6.2 Scenario 2

In the second scenario, we perform re-clustering after every k rounds. Three different values of k are considered. Re-clustering is done after every 20, 25 and 50 rounds.

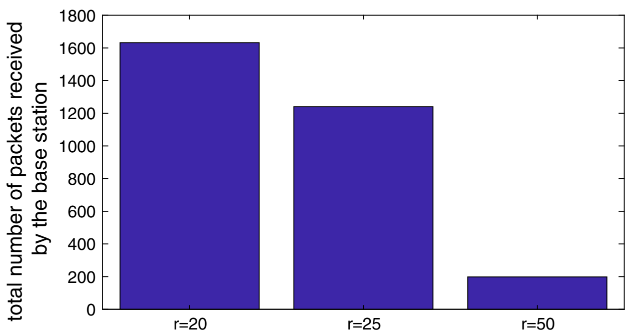


Fig. 12 Number of packets received

6.2.1 Network lifetime

The following Fig. 11 shows the network lifetime comparison for these three different scenarios.

It can be seen that as the value of k increases, there is a decrease in Network Lifetime. This is because some of the cluster heads might have utilized their energy completely and need to be rotated.

6.2.2 Number of packets

The following Fig. 12 shows the number of packets received by the base station for three different scenarios.

Thus on the basis of above simulation results, we can conclude that if we perform re-clustering very frequently, the network lifetime will be more but at the cost of large computational time (in performing re-clustering). Thus we should try to chose a value of k so as to achieve a balance between computational time and lifetime.

6.3 Performance comparison with related clustering algorithms

We provide an extensive analysis of related cluster creation algorithms proposed till date, in terms of network lifetime metrics : FND (first node dead), HND (half node dead) and CND (complete node dead) in Table 8. The results were extracted from the related protocols for the scenario in which the position of base station is outside the area of interest, that is far away from the sensor nodes, with random node distribution and single hop communication from cluster head to base station. We then ran our proposed algorithm using the same network metrics (network area, number of nodes, radio model, data packet length etc.) as specified in the comparative algorithm and then determined the network lifetime metrics for our algorithm. It is found that the proposed algorithm outperforms all the related protocols in HND and CND metrics.

Table 8 Network lifetime comparison with related protocols

Article	Description	Network lifetime metrics			Proposed		
		FND	HND	CND	FND	HND	CND
Hussain et al. [57]	proposed a weighted genetic algorithm based approach for cluster head election using five fitness parameters : direct distance to base station,cluster distance, cluster distance standard deviation, transfer energy, number of transmissions	–	4200	8000	2043	16545	19228
Nguyen et al. [58]	proposed DB-LEACH (distance based), which considers geometric distance between cluster head and base station in cluster head election and DBEA-LEACH (distance based energy aware) which also considers residual energy of cluster head into consideration	–	–	831	457	2987	3655
Sert et al. [51]	proposed MOFCA (multiobjective fuzzy clustering algorithm) with residual energy, distance to base station and density for calculating cluster head competition radius, making use of fuzzy logic for handling uncertainties in WSN	117	271	–	109	879	1084
Xie et al. [59]	proposed CRT2FLACO, based on Type-2 fuzzy logic with residual energy, number of neighbours and distance to base station for cluster head election and cluster radius calculation and ant colony optimization for cluster head chain creation	1082	1099	1117	615	4927	5026
Barati et al. [60]	proposed EACHP (Energy Aware Clustering Hierarchy Protocol) with residual energy, communication cost, density and distance as cluster head election parameters	105	170	831	231	1854	2077
Hacioglu et al. [52]	proposed NSGA-II based clustering protocol which considers distance and energy parameters for cluster head election	09	151	174	22	173	202
Zhao et al. [20]	proposed modified cluster head selection algorithm based on LEACH, LEACH-M, which considers residual energy and network address of a ZigBee nodes for cluster head creation	500	590	650	327	2632	2924

7 Conclusion

Efficient clustering is a major challenge in Wireless Sensor Networks in order to improve energy conservation. In this paper, we have proposed a multi-objective clustering algorithm based on NSGA-II with ENLU mechanism. We have considered high residual energy cluster heads, minimum total energy consumption, maximum cluster head separation and load balancing as simultaneous optimization goals. Although there are large number of multi-objective clustering algorithms available, but none of them have considered all of them as simultaneous optimization goals. The proposed algorithm shows approximately 5 times improvement in network lifetime as compared to LEACH and 16 percent improvement as compared to Hacioglu et al., which has not been achieved by any other existing clustering algorithm. In future, we will try to simulate our proposed algorithm and compare it to LEACH and other protocols on a discrete event simulator like NS2. Also, to capture a real life scenario, we will try to verify the performance of our proposed algorithm with a more realistic energy transmission model [61–63]. Mobility of nodes and the base station can also be taken into account.

References

1. Arampatzis, T., Lygeros, J., & Manesis, S. (2005). A survey of applications of wireless sensors and wireless sensor networks. In *Proceedings of the 2005 IEEE international symposium on intelligent control, Mediterrean conference on control and automation*. IEEE.
2. Al-Karaki, J. N., & Kamal, A. E. (2004). Routing techniques in wireless sensor networks: A survey. *IEEE Wireless Communications*, *11*(6), 6–28.
3. Puccinelli, D., & Haenggi, M. (2005). Wireless sensor networks: Applications and challenges of ubiquitous sensing. *IEEE Circuits and Systems Magazine*, *5*(3), 19–31.
4. Younis, M., & Akkaya, K. (2008). Strategies and techniques for node placement in wireless sensor networks: A survey. *Ad Hoc Networks*, *6*(4), 621–655.
5. Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks*, *52*(12), 2292–2330.
6. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd annual Hawaii international conference on system sciences*. IEEE.
7. Tripathi, R. K., Singh, Y. N., & Verma, N. K. (2012). N-LEACH, a balanced cost cluster-heads selection algorithm for Wireless Sensor Network, 2012. In *National conference on communications (NCC)* (pp. 1–5). Kharagpur.
8. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, *1*(4), 660–670.
9. Aju, O. G. (2015). A survey of zigbee wireless sensor network technology: Topology, applications and challenges. *International Journal of Computer Applications*, *130*(9), 47–55.
10. Gaddour, O., et al. (2010). Z-Cast: A multicast routing mechanism in ZigBee cluster-tree wireless sensor networks. In *2010 IEEE 30th international conference on distributed computing systems workshops*. IEEE
11. Koubâa, A., et al. (2008). TDBS: a time division beacon scheduling mechanism for ZigBee cluster-tree wireless sensor networks. *Real-Time Systems*, *40*(3), 321–354.
12. Tsai, C.-H., & Tseng, Y.-C. (2012). A path-connected-cluster wireless sensor network and its formation, addressing, and routing protocols. *IEEE Sensors Journal*, *12*(6), 2135–2144.
13. Menzel, K., et al. (2008). Towards a wireless sensor platform for energy efficient building operation. *Tsinghua Science and Technology*, *13*(S1), 381–386.
14. Pei, Z., et al. (2008). Application-oriented wireless sensor network communication protocols and hardware platforms: A survey. In *2008 IEEE international conference on industrial technology*. IEEE.
15. Mainetti, L., Patrono, L., & Vilei, A. (2011). Evolution of wireless sensor networks towards the internet of things: A survey. In *19th international conference on software, telecommunications and computer networks SoftCOM*. IEEE.
16. Kumar, T., Mane, P. B., & ZigBee Topology. (2016). A survey, 2016. In *International conference on control, instrumentation, communication and computational technologies (ICCICCT)* (pp. 164–166). Kumaracoil.
17. Cuomo, F., et al. (2008). Topology formation in IEEE 802.15. 4: Cluster-tree characterization. In *Sixth annual IEEE international conference on pervasive computing and communications (PerCom)*. IEEE
18. Jaichandran, R., & Irudhayaraj, A. A. (2010). Effective strategies and optimal solutions for hot spot problem in wireless sensor networks (WSN). In *10th International conference on information science, signal processing and their applications (ISSPA 2010)*. IEEE
19. Loscri, V., Morabito, G., & Marano, S. (2005). A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH). In *IEEE 62nd vehicular technology conference VTC-2005-Fall* (vol. 3). IEEE.
20. Zhao, L., Shaocheng, Q., & Yi, Y. (2018). A modified cluster-head selection algorithm in wireless sensor networks based on LEACH. *EURASIP Journal on Wireless Communications and Networking*, *2018*(1), 287.
21. Kang, S. H., & Nguyen, T. (2012). Distance based thresholds for cluster head selection in wireless sensor networks. *IEEE Communications Letters*, *16*(9), 1396–1399.
22. Manjeshwar, A., & Agrawal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In *Proceedings 15th international parallel and distributed processing symposium (IPDPS)* (pp. 2009–2015). San Francisco, CA.
23. Manjeshwar, A., & Agrawal, D. P., APTEEN. (2002). A hybrid protocol for efficient routing and comprehensive information retrieval in wireless. In *Proceedings 16th international parallel and distributed processing symposium* (p. 8). Ft. Lauderdale, FL.
24. Gunjan, M., P., & Sharma, A. K. (2018). Modified TEEN for handling inconsistent cluster size problem in WSN, In *2018 international conference on wireless communications, signal processing and networking (WiSPNET)*, Chennai (pp. 1–6).
25. Younis, O., & Fahmy, S. (2004). HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, *3*(4), 366–379.
26. Lindsey, S., & Raghavendra, C. S. (2002). PEGASIS: Power-efficient gathering in sensor information systems. In *IEEE Aerospace conference proceedings* (vol. 3). IEEE.
27. Li, C., et al. (2011). A survey on routing protocols for large-scale wireless sensor networks. *Sensors*, *11*(4), 3498–3526.

28. Geetha, V. A., Kallapur, P. V., & Tellajeera, S. (2012). Clustering in wireless sensor networks: Performance comparison of LEACH and LEACH-C protocols using NS2. *Procedia Technology*, *4*, 163–170.
29. Kuila, P., & Jana, P. K. (2012). Energy efficient load-balanced clustering algorithm for wireless sensor networks. *Procedia Technology*, *6*, 771–777.
30. Ferentinos, K. P., & Tsiligiridis, T. A. (2007). Adaptive design optimization of wireless sensor networks using genetic algorithms. *Computer Networks*, *51*(4), 1031–1051.
31. Hu, X.-M., et al. (2010). Hybrid genetic algorithm using a forward encoding scheme for lifetime maximization of wireless sensor networks. *IEEE Transactions on Evolutionary Computation*, *14*(5), 766–781.
32. Yoon, Y., & Kim, Y.-H. (2013). An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks. *IEEE Transactions on Cybernetics*, *43*(5), 1473–1483.
33. Gupta, S. K., Kuila, P., & Jana, P. K. (2013). GAR: an energy efficient GA-based routing for wireless sensor networks. In *International conference on distributed computing and internet technology*. Berlin: Springer.
34. Rathee, M., & Kumar, S. (2016). Quantum inspired genetic algorithm for multi-hop energy balanced unequal clustering in wireless sensor networks. In *2016 Ninth international conference on contemporary computing (IC3)*. IEEE.
35. Deb, K., et al. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197.
36. Li, K., et al. (2014). Efficient non-domination level update approach for steady-state evolutionary multiobjective optimization. Department of Electtrical and Computer Engineering, Michigan State University, East Lansing, USA, Technical Report COIN Report 2014014.
37. Hu, S., & Li, G. (2018). Fault-tolerant clustering topology evolution mechanism of wireless sensor networks. *IEEE Access*, *6*, 28085–28096.
38. Mehmood, A., et al. (2015). Improvement of the wireless sensor network lifetime using LEACH with vice-cluster head. *Ad Hoc & Sensor Wireless Networks*, *28*(1–2), 1–17.
39. Gawade, R. D., & Nalbalwar, S. L. (2016). A centralized energy efficient distance based routing protocol for wireless sensor networks. *Journal of Sensors*, *2016*, 8313986.
40. Guravaiah, K., & Velusamy, R. L. (2018). BEACH: Balanced energy and adaptive cluster head selection algorithm for wireless sensor networks. *Adhoc & Sensor Wireless Networks*, *42*(3–4), 199–225.
41. Mittal, N., Singh, U., & Sohi, B. S. (2017). A stable energy efficient clustering protocol for wireless sensor networks. *Wireless Networks*, *23*(6), 1809–1821.
42. Sindjoung, M. L. F., et al. (2018). ISCP: An instantaneous and secure clustering protocol for wireless sensor networks. *Network Protocols & Algorithms*, *10*(1), 65–82.
43. Mehmood, A., et al. (2017). Secure knowledge and cluster-based intrusion detection mechanism for smart wireless sensor networks. *IEEE Access*, *6*, 5688–5694.
44. Mezrag, F., Bitam, S., & Mellouk, A. (2017). Secure routing in cluster-based wireless sensor networks. In *IEEE global communications conference GLOBECOM 2017–2017*. IEEE.
45. Zhou, H., et al. (2016). A security mechanism for cluster-based WSN against selective forwarding. *Sensors*, *16*(9), 1537.
46. Fang, W., et al. (2019). CSDA: A novel cluster-based secure data aggregation scheme for WSNs. *Cluster Computing*, *22*(3), 5233–5244.
47. Peiravi, A., Mashhadi, H. R., & Javadi, S. H. (2013). An optimal energy-efficient clustering method in wireless sensor networks using multi-objective genetic algorithm. *International Journal of Communication Systems*, *26*(1), 114–126.
48. Cheng, C.-T., & Leung, H. (2012). A multi-objective optimization framework for cluster-based wireless sensor networks. In *2012 International conference on cyber-enabled distributed computing and knowledge discovery (CyberC)*. IEEE.
49. Ozdemir, S., Bara'a, A. A., & Khalil, Ö. A. (2013). Multi-objective evolutionary algorithm based on decomposition for energy efficient coverage in wireless sensor networks. *Wireless Personal Communications*, *71*(1), 195–215.
50. Iqbal, M., et al. (2015). Wireless sensor network optimization: multi-objective paradigm. *Sensors*, *15*(7), 17572–17620.
51. Sert, S. A., Bagci, H., & Yazici, A. (2015). MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks. *Applied Soft Computing*, *30*, 151–165.
52. Hacioglu, G., Kand, V. F. A., & Sesli, E. (2016). Multi objective clustering for wireless sensor networks. *Expert Systems with Applications*, *59*, 86–100.
53. Jameii, S. M., Faez, K., & Dehghan, M. (2016). AMOF: Adaptive multi-objective optimization framework for coverage and topology control in heterogeneous wireless sensor networks. *Telecommunication Systems*, *61*(3), 515–530.
54. Mazumdar, N., & Om, H. (2017). A distributed fault-tolerant multi-objective clustering algorithm for wireless sensor networks. In *Proceedings of the international conference on nano-electronics, circuits communication systems*. Singapore: Springer.
55. Gamwarige, S., & Kulasekere, C. (2006). Application of the EDCR algorithm in a cluster based multi-hop wireless sensor network. In *2006 International symposium on communications and information technologies*. IEEE.
56. Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*(2), 65–85.
57. Hussain, S., Matin, A. W., & Islam, O. (2007). Genetic algorithm for hierarchical wireless sensor networks. *JNW*, *2*(5), 87–97.
58. Nguyen, T. G., So-In, C., Nguyen, N. G. (2014). Two energy-efficient cluster head selection techniques based on distance for wireless sensor networks. In *2014 International computer science and engineering conference (ICSEC)*. IEEE.
59. Xie, W.-X., et al. (2015). A clustering routing protocol for WSN based on type-2 fuzzy logic and ant colony optimization. *Wireless Personal Communications*, *84*(2), 1165–1196.
60. Barati, H., Movaghar, A., & Rahmani, A. M. (2015). EACHP: Energy aware clustering hierarchy protocol for large scale wireless sensor networks. *Wireless Personal Communications*, *85*(3), 765–789.
61. Halgamuge, M. N., et al. (2009). An estimation of sensor energy consumption. *Progress in Electromagnetics Research*, *12*, 259–295.
62. Abo-Zahhad, M., Amin, O., Farrag, M., & Ali, A. (2014). Survey on energy consumption models in wireless sensor networks. *Open Transactions on Wireless Sensor Networks*, *1*(1).
63. Wang, Q., Hempstead, M., & Yang, W. (2006). A realistic power consumption model for wireless sensor network devices. In *2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks (SECON)* (Vol. 1, pp. 286–295). Reston, VA: IEEE.

Gunjan is currently pursuing her Ph.D. in Computer Science and Engineering Department from National Institute of Technology, Delhi, India. She received her Bachelor's in Technology in 2012 from Maharishi Dayanand University, Haryana, India; Masters in Technology from National Institute of Technology, Jalandhar, Punjab, India in 2014. Her research interest lies in the domain of energy efficient techniques in wireless sensor networks.

Ajay K. Sharma is working as Vice-Chancellor IKG Punjab Technical University, Jalandhar, Punjab, India. He received his B.E. in Electronics and Electrical Communication Engineering from Punjab University Chandigarh, India in 1986, M.S. in Electronics and Control from Birla Institute of Technology (BITS), Pilani in the year 1994 and Ph.D. in Electronics Communication and Computer Engineering in the year 1999. His Ph.D. thesis was on "Studies on Broadband Optical Communication Systems and Networks". After serving various organizations from 1986 to 1995, he has joined National Institute of Technology (Erstwhile Regional Engineering College) Jalandhar as Assistant Professor in the Department of Electronics and Communication Engineering in the year 1996. From November 2001, he has worked as Professor in the ECE Department and thereafter he has worked as Professor in Computer Science and Engineering from 2007 to 2013 in the same institute. His major areas of interest are broadband optical wireless communication systems and networks, dispersion compensation, fiber nonlinearities, optical soliton transmission, WDM systems and networks, radio-over-fiber (RoF) and wireless sensor networks and computer communication.

Karan Verma is working as Assistant Professor in Computer Science and Engineering Department, NIT Delhi. He received his B.E. and M. Tech degrees in Information Technology and Computer Science and Engineering, respectively, from the University of Rajasthan in 2008, and Indian Institute of Technology, Roorkee in 2010. He received his Ph.D. degree from Department of Computer and Information Science at the Universiti Teknologi PETRONAS, Malaysia, and Post DoC (PDF) in Computer Science and Engineering from University of Californian, Davis, USA. His research interests include information security, wireless sensor networksand VANET.