Abstract—Wireless Sensor Networks (WSNs) enable new applications and need non-conventional paradigms for the protocol because of energy and bandwidth constraints. In WSN, sensor node’s life is a critical parameter. Research on life extension is based on Low-Energy Adaptive Clustering Hierarchy (LEACH) scheme, which rotates Cluster Head (CH) among sensor nodes to distribute energy consumption over all network nodes. CH selection in WSN affects network energy efficiency greatly. This study proposes an improved CH selection for efficient data aggregation in sensor networks. This new algorithm is based on Bacterial Foraging Optimization (BFO) incorporated in LEACH.

Keywords—Bacterial Foraging Optimization (BFO), Cluster Head (CH), Data-aggregation protocols, Low-Energy Adaptive Clustering Hierarchy (LEACH).

I. INTRODUCTION

WIRELESS SENSOR NETWORK (WSN) [1] is a spatially distributed autonomous device using sensors to monitor physical/environmental conditions like pressure, temperature, vibration, sound and motion pollutants at locations. WSN has hundreds to thousands of low-power multi-functioning sensor nodes, in an unattended environment with restricted computational/sensing capabilities. These nodes collect, process, and pass collected information to a central location. WSNs unique characteristics are Low duty cycle, Redundant Data Acquisition, Power constraints, Limited battery life, Sensor Nodes Heterogeneity, Nodes Mobility, Dynamic Network Topology, etc. [2].

A data aggregation algorithm uses data from a sensor node and aggregates data using aggregation algorithms like Low Energy Adaptive Clustering Hierarchy (LEACH), centralized approach, and Tiny Aggregation (TAG). The aggregated data is sent to a sink node by selecting an efficient path [3].

The aim of Data Aggregation Algorithm is collecting and aggregating data in an energy efficient manner to enhance network life. WSN offers data gathering in Distributed System Architectures and Dynamic access through wireless connectivity [4].

The aim of Data Aggregation Protocols is to eliminate redundant data transmission and improves energy constrained in wireless sensor network’s life. In WSN data transmission happens in a multi-hop fashion where a node forwards data to a neighbour node near the sink. In this approach, as closely placed, nodes may sense the same data, which is not energy efficient. An improvement over it is clustering where a node sends data to Cluster-Head (CH) which performs aggregation on raw data and sends to a sink [5].

Routing determines a path between source node and sink (destination) node during data transmission. In WSNs, a network layer implements incoming data routing. Generally in multi-hop networks a source node cannot reach a sink directly. So, intermediate sensor nodes relay on packets. Routing table implementation is the solution. They have lists of node options for any packet destination. Routing table is a routing algorithm task aided by a routing protocol for construction/maintenance [6]. Routing paths are established through Proactive, Reactive or Hybrid ways. Proactive protocols compute routes before need and store them in a node’s Routing table. Reactive protocols compute routes only when it is needed. Hybrid protocols combine both the ideas [7].

Proactive routing protocols maintain network nodes consistent and accurate routing tables through periodic routing information dissemination. In this routing category, all routes are computed before being needed. Hierachical proactive routing is the solution to meet Routing demands for large ad hoc networks. Reactive routing strategies do not maintain Network Node’s global information, but route establishment between source and destination is based on a dynamic search according to demand. To discover a route from the source to destination, a route discovery query and reverse path are used for query and replies. The hybrid strategy is applicable to large networks. Hybrid routing strategies have proactive and reactive routing strategy. It uses clustering to make a network stable and scalable. A network cloud is divided into a cluster which is maintained dynamically when a node is added or leaves a cluster. Hybrid routing has Network overhead which is required to maintain clusters [8].

BFO technique optimizes the positions of multiple base stations randomly in a network to improve the likelihood of sensor node packets reaching at least one base station (BSs) due to the presence of large black hole regions, thereby
ensuring high-success [8]. BFO is a population-based numerical optimization algorithm. Recently, bacterial foraging behaviour was a rich source for solutions to many engineering applications and computational models. It was applied to solve practical engineering problems like optimal control and harmonic estimation channel equalization. [9].

BFOA is a part of nature which inspired optimization algorithms. Evolution and natural genetics inspired optimization algorithms like Genetic Algorithms (GA), Evolutionary Strategies (ES) dominated optimization algorithms for over several decades. Recently natural swarm inspired algorithms like Ant Colony Optimization (ACO), PSO, Artificial Bee Colony Optimization (ABC) entered the domain and proved their effectiveness. It was successfully applied to problems like optimization control, images detection and image quantization. The control system of bacteria dictates how foraging should proceed, is subdivided into four sections, i.e.; chemotaxis, swarming, reproduction, and elimination/dispersal [10].

Sensor nodes in clustering are partitioned into different clusters with each being managed by a node called CH. The other cluster nodes do not communicate to base station directly. They pass collected data to a CH which aggregates data from cluster nodes and transmit it to a base station by reducing energy consumption and communicate messages to the base station. The number of active nodes in communication is also lowered. The result of clustering sensor nodes is prolonged network life. The clustering architecture in WSN is shown in Fig. 1.

WSN issues like Localization, Node deployment, Energy-aware clustering and Data-aggregation are formulated as optimization problems. Traditional analytical optimization techniques need enormous computational effort, which increases exponentially as problem size increases. An optimization method needs moderate memory and computational resources which produces good results as desirable, especially for implementation on individual sensor nodes. Bio-inspired optimization methods are computationally efficient alternatives. This paper presents a Bacterial Foraging Optimization Algorithm (BFA). This BFA model adopts a combination of random movements and a straight line to reach nutrient rich locations [11].

This paper is proposed an improved CH selection for efficient sensor network data aggregation. This new algorithm is based on BFO incorporated in LEACH. Section II represents related works in WSNs. Section III gives the details of the Methodology and Section IV provides the Results and Discussion. Section V concludes the paper.

II. LITERATURE REVIEW

A WSN sensor communicates directly with other sensors within radio range in a cluster. Many clustering algorithms like LEACH, DEEC, and SEP were proposed with the aims of route-path selection, energy minimization, increased connectivity and network longevity. The problems of cluster formation and CH selection between different protocols for data aggregation and transmission were compared by [12]. The authors focused on the problem’s two aspects: (i) how to guess the clusters needed to consume available sources for a sensor network proficiently, and (ii) how to select CHs to cover up sensor networks efficiently. Simulation compared the performance of different protocols to find optimal solutions for the above problems. A CH selection algorithm for adapting clusters and rotating CH positions to distribute energy load
among nodes was proposed by [13]. The new model was extended to a LEACH's stochastic CH selection algorithm by modifying a node’s probability to become CH based on sensor nodes remaining energy level for transmission. A new scheme to cluster for data aggregation called Efficient CH Selection Scheme for Data Aggregation (ECHSSDA) in WSN presented by [14] was compared to LEACH clustering algorithm.

A new fuzzy multiple criteria decision-making approach based on trapezoidal fuzzy AHP and hierarchical fuzzy integral (FAHP) introduced by [15] optimized CHs selection to develop a distributed energy-efficient clustering algorithm. An efficient request-oriented coordinator method for hierarchical sensor networks presented by [16] considered the request type to ensure suitable coordinator selection. A request that needs live nodes, E selection is a good choice. The coordination selection is based on a node knowing its energy information. A new optimization algorithm; multi-colony bacterial foraging optimization (MC-BFO), to solve complex Radio Frequency Identification (RFID) network planning problems was proposed by [17], whose main idea was extending the single population bacterial foraging algorithm to interacting with a multi-colony model by relating a single bacterial cell’s chemotactic behaviour to cell-to-cell communication of a bacterial community. Its performance compared to GA and PSO on an RFID network planning problem, proved its superiority.

BFO algorithm mimics how bacteria forage over a nutrients landscape to perform parallel non-gradient optimization. A tutorial on BFO, including bacterial foraging and pseudo-code that models this process was provided by [18]. The algorithm’s features were compared to those in GA, other bio-inspired methods, and non-gradient optimization. BFO applications and future directions were presented.

A multidimensional scaling localisation algorithm based on BFO was proposed by [19]. The Multidimensional Scaling (MDS) algorithm got unknown nodes’ initial coordinates. BFO algorithm obtained unknown nodes final coordinates by optimising a local cost function. Sharma [20] used BFO algorithm. Other bio-inspired algorithms like ACO, Artificial Immune system and GA (significant time/power consuming) are also compared to BFO; but reducing computational complexity remains. A comparative study of different algorithms computational complexity was analysed. BFO was first for WSNs to enhance network life of sensor nodes. To validate the algorithm, simulations were carried out in MATLAB. Results showed that BFO had better performance compared to other clustering protocols.

Kavitha and Wahidabanu [21] proposed foraging optimization for CH Selection by which an improved CH selection for efficient data aggregation in sensor networks was realised. The algorithm was based on LEACH incorporated BFO with energy saving objective function. The new BFO had the high average throughput, low delay in seconds and low DATA dropping compared to LEACH.BFO algorithm was presented by [22]. It was widely accepted as an optimization algorithm for control and optimization. An analysis of BFOA with GA was presented.

Improving WSN’s LEACH Protocol using Fuzzy Logic was proposed by [23]. An improved LEACH (LEACH-C) algorithm called partition-based LEACH (pLEACH), which partitions a network into optimal sectors, and then selects a node with the highest energy as head for a sector, using centralized calculations was proposed by [24]. Simulation and analysis showed that pLEACH achieved much better WSN performance regarding energy dissipation, network life, and communication quality. LEACH-B (LEACH-Balanced) was presented by [25] where at every round, after first selection of CH according to LEACH protocol, a second selection modified cluster heads considering a node's residual energy. So the cluster heads are constant and near optimal per round.

III. METHODOLOGY

This study proposes an improved cluster head selection for efficient sensor networks data aggregation. BFO algorithm is incorporated in LEACH.

A. Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH by [26] is a famous clustering protocol that is the basis for many clustering protocols. The aim of LEACH’s is to have CHs reduce the energy cost of transmitting data from normal nodes to distant Base Stations [27]. The LEACH operation is divided into a Setup Phase and a Steady state phase. Each round begins with a set-up (clustering) phase where clusters are organized, followed by a steady - state (transmission) phase where data packets are transferred from nodes to CHs. After data aggregation, CHs transmit messages to a Base Station. In Setup Phase, a node decides whether to become a cluster head for the current round. The cluster head election is done with a probability function: a node selects a random number between 0 and 1 which if less than T(n), elects the node as a cluster head for current round:

\[ T(n) = \begin{cases} \frac{p}{1-p} \frac{(r-\text{mod}\ p)}{p} & \text{if } n \in G \\ 0 & \text{Otherwise} \end{cases} \]  

where, P is cluster head probability, r is a number of current round and G a set of nodes that have not been CHs in last \( \frac{r}{p} \) rounds. After the CH election, every CH prepares a Time Division Multiple Access (TDMA) schedule and transmits it to all nodes in their cluster. This completes LEACH’s set up phase.

Steady State Phase: Here nodes send collected data to CH at once per frame allotted to them assuming that a node always has data to transmit. The node goes to sleep mode after transmission till the next allotted transmission slot, to save energy. A CH must keep its receiver on always to receive cluster nodes data. After reception of data, CH aggregates it and transmits it to a base station.

LEACH’s strength is its CH rotation mechanism and data aggregation. However, a problem with LEACH is that it has no guarantee of placement and/or number of cluster head nodes in a round [27]. Thus, an energy level threshold is determined and then, it selects nodes (with higher energy than the threshold) as possible cluster heads.
Determining, an optimal number of cluster heads is an NP-Hard problem. LEACH-C uses Simulated Annealing [28] algorithm to offset this. After determining cluster heads of a current round, BS sends a message having cluster head ID for all nodes. If a node's CH ID matches its ID, the node is a CH; or else it is a normal node and goes to sleep till the data transmission phase. LEACH-C delivers 40% more data per unit energy than LEACH as BS has global knowledge of location and network nodes energy levels [26].

The CH node’s election in LEACH [26] has deficiencies like, both big and very small clusters that may exist in a network simultaneously, In CH selection the nodes have different energy ignores geographic location, residual energy, and further information, then it leads to CH node failing.

B. Bacterial Foraging Optimization (BFO)

BFO is a new class of biologically encouraged, stochastic, global search technique mimicking E. coli bacteria’s foraging behaviour. This method locates handles and ingests food. During foraging, a bacterium exhibits tumbling or swimming actions [29].

Chemotaxis movement continues till a bacterium reaches a positive-nutrient gradient. After specific swims, the population’s best half undergoes reproduction eliminating positive-nutrient gradient. After specific swims, the population’s best half undergoes reproduction eliminating positive-nutrient gradient. After specific swims, the population’s best half undergoes reproduction eliminating positive-nutrient gradient. After specific swims, the population’s best half undergoes reproduction eliminating positive-nutrient gradient.

Chemotaxis foraging behaviour has a common type of bacteria with a diameter of 1 µm and length of about 2 µm that under correct circumstances reproduces in 20 min. The ability to move is from a setup of to six rigid 100–200 rps spinning flagella, driven by a biological motor. When flagella rotate clockwise, they operate as propellers and so, an E. coli runs or tumbles.

Chemotaxis Actions are:

(A1) In neutral medium, alternate tumbles and runs ⇒ search.
(A2) If swimming (up a gradient nutrient or out of noxious substances), swim longer (Climb-up nutrient gradient or down noxious gradient) ⇒ seek increasingly favourable environments.
(A3) If swimming down gradient nutrient (or up noxious substance gradient), then search ⇒ to avoid unfavourable environments.

Bacteria swarm S behaves as follows [30]:

1. Bacteria are randomly distributed in nutrient’s map.
2. Bacteria move to high-nutrient regions on a map. Those in noxious substance regions or low-nutrient regions die/disperse. Bacteria in convenient regions reproduce (split).
3. Bacteria are located in promising regions of nutrients map as they try to attract other bacteria by generating chemical attractants.
4. Bacteria are now located in a highest-nutrient region.
5. Bacteria now disperse to look for new nutrient regions on a map.

The procedures implemented are:

\[
f = \beta f_1 + (1 - \beta) f_2 \quad (2)
\]

where, \(f_1\) is nodes maximum average Euclidean distance with associated CHs and \(f_2\) is the ratio of nodes total initial energy to CH candidate’s total energy expressed as:

\[
f_1 = \max_{k=1,2,3...K} \left( \frac{d(n_i,CH_{p,k})}{\sum_{i \in C_{p,k}} d(n_i,CH_{p,k})} \right) \quad (3)
\]

\[
f_2 = \frac{\sum_{p=1}^{K} E(n_i)}{\sum_{p=1}^{K} E(C_{p,k})} \quad (4)
\]

Here, \(N\) is a number of nodes of which \(K\) is elected as CHs.

\(\sum_{p=1}^{K} E(n_i)\) are nodes that belong to cluster \(C_k\) in particle \(p\), ensuring that only nodes with above average energy resources are elected as CHs, with the minimum average distance between nodes and CHs. LEACH uses energy as the objective function and is based on the energy threshold \(T\). In this work a novel objective function is proposed and given by:

\[
\min f_i (x) = a_1 (\min (PLR)) + a_2 \left( \frac{E_i^r}{E_{initial}} \right) \quad (5)
\]

where \(E_i^r\) is the remaining energy in node \(i\); \(E_{initial}\) is the initial energy in the node; PLR is the packet loss rate.

IV. RESULTS AND DISCUSSION

The experiments conducted for varying number of nodes with the single base station in a 2sq. km area. The number of nodes in a network ranges from 30 to 180. The simulations conducted evaluate the performance of the proposed BFO, LEACH and method proposed by [21], for clusters formed, average end to end delay, average packet drop ratio, and lifetime computation. The proposed method is compared with LEACH.

### TABLE I

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>LEACH</th>
<th>Cluster formation using Bacterial Foraging Algorithm</th>
<th>Cluster formation using BFO [21]</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>60</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>90</td>
<td>15</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>120</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>150</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>180</td>
<td>22</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>LEACH</th>
<th>Cluster formation using Bacterial Foraging Algorithm</th>
<th>Cluster formation using BFO [21]</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.001242</td>
<td>0.001198</td>
<td>0.001196</td>
</tr>
<tr>
<td>60</td>
<td>0.001138</td>
<td>0.00149</td>
<td>0.001505</td>
</tr>
<tr>
<td>90</td>
<td>0.011368</td>
<td>0.013081</td>
<td>0.013278</td>
</tr>
<tr>
<td>120</td>
<td>0.018795</td>
<td>0.016424</td>
<td>0.016731</td>
</tr>
<tr>
<td>150</td>
<td>0.043275</td>
<td>0.036754</td>
<td>0.037391</td>
</tr>
<tr>
<td>180</td>
<td>0.04613</td>
<td>0.037921</td>
<td>0.038646</td>
</tr>
</tbody>
</table>
When the number of nodes is 30, the percentage of the number of cluster formation of Bacterial Foraging Algorithm is increased by 13.33% than LEACH. When the number of nodes is 180, the percentage of the number of cluster formation of Bacterial Foraging Algorithm is increased by 4.44% than LEACH.

When the number of nodes is 30, the percentage of Average End to End Delay (sec) of Cluster formation using Bacterial Foraging Algorithm is decreased by 4.87% than LEACH.

When the number of nodes is 30, the percentage of Average packet drop ratio of Cluster formation using Bacterial Foraging Algorithm is decreased by 6.60% than LEACH. When the number of nodes is 180, the percentage of Average packet drop ratio of Cluster formation using Bacterial Foraging Algorithm is decreased by 6.42% than LEACH.

If the number of rounds is 200, then the percentage of lifetime computation of Cluster formation using Bacterial Foraging Algorithm is increased by 11.24% than LEACH. If the number of rounds 500 then the percentage of lifetime computation of Cluster formation using Bacterial Foraging Algorithm is increased by 82.92% than LEACH. It is also seen that our method objective improves over Kavitha and Wahidabanu who have used energy based objective function [21].

V. CONCLUSION

LEACH chooses too many CHs at a time or randomly selects CHs far from a base station without considering nodes' residual energy. Hence, some cluster heads drain energy early thereby reducing the WSN’s life. This paper proposed to improve cluster head selection for efficient sensor networks data aggregation. This new algorithm is based on BFO incorporated in LEACH scheme. Experiments were conducted

![Fig. 2 Number of Clusters formed](image)

![Fig. 3 Average End to End Delay (sec)](image)

![Fig. 4 Average packet drop ratio](image)

![Fig. 5 Lifetime computation](image)
on 40 nodes with on base station in a 2sq. km area. This proposed algorithm shows better performance in terms of selection of the higher number of Clusters, Average End to End Delay (sec), Average packet drop ratio, and Lifetime computation.

REFERENCES


