Data Aggregation on Single & Multiple data sinks in Heterogeneous Wireless Sensor Network using BATOC and MS- BATOC Protocol

Jitendra Bahadur Singh¹ and R.C. Tripathi²
¹Research Scholar, Department of Electronics Engineering, NGBU, Prayagraj (Uttar Pradesh), India.
²Dean Research (Rtd. Professor, IIT Allahabad), Department of Engineering, NGBU, Prayagraj (Uttar Pradesh), India.

(Received 25 May 2020, Revised 23 June 2020, Accepted 02 July 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The biggest issue of Wireless Sensor Networks (WSN) is energy consumption. Battery-powered sensor networks turn out to be dead during a given period. Therefore, it is more difficult to increase data dissipation in an energy-efficient way to maximize the duration of the sensor devices. The clustering approach has also been proved capable of maintaining or increasing WSN’s life cycle. The selection of CH (cluster head) in every cluster considers in cluster model is the effective routing process, which minimizes the delay in transmission of WSN. In the case of Heterogeneous WSN (HWSN), the selection of CH is already left with a broad area for further development of its optimization capability. Also, Hot-Spot problem & to limit communications from nodes to sinks are big issues. The key challenge was however to take the best CH that speeds up network operation. So far, further work has been undertaken to address this issue by taking into consideration different constraints. In this situation, this paper tries with several key parameters including energy, delay, distance, and security to create a new clustering model with optimal selection of the cluster heads. In this paper, we used bat algorithm based optimized clustering for single sink (BATOC) and multiple sink (MS-BATOC) for appropriate cluster head selection in the cluster by integrating the distance parameter by calculating fitness function toward the sink node. It distributes threshold method dependent data from cluster member nodes for transmit the data. Some will forward the data to the relevant Cluster Head, where the Check value is higher than the predefined soft threshold value before threshold conditions are met. From this method the result obtained is better in compare with existing GA based optimized clustering. Finally, in contrast to traditional models, schemes are carried out based on energy consumption, number of alive nodes, no. of dead nodes and efficiency.

Keywords: Bat optimization, CH selection, Clustering, Data sinks, Genetic Algorithm, Heterogeneous Wireless Sensor Network, Multiple sink.

I. INTRODUCTION

A significant number of sensor nodes (SN) and sinks are collected in a sensor network. As an entrance base Station usually serves to many other networks. It provides efficient data collection, backup centers and a point of entry to the network sensor nodes. Sensor nodes sense their surroundings, gather sensed information and send it to BS. The power, machine resources, and memory are therefore limited. Placing some of heterogeneous nodes into the WSN would further improve network life and reliability. WSN Heterogeneity can be used to improve network life and reliability. Many of the latest energy-efficient protocols for heterogeneous networks have been developed by the clustering method, which can be extended to scalability and energy saving in WSNs [1]. According to the differences of node computation, sensing, connectivity and energy factors WSN divided into four different heterogeneous types: link, node energy, computing energy and network protocol heterogeneous type [2]. The nodes will submit data through multiple hops in a single sink WSN. In a broad WSN, energy consumption is very inefficient and all information is stored into a solitary sink [3]. The highest energy usage occurs in data transferring from cluster members to sinks. During the transmission to the sink, several sinks are used for reducing energy consumption. If several sinks are put, the distance from cluster node to the sink decreases, meaning that the distance between the perceived data is not shortened by decreasing the energy demand significantly [4]. The sensed data will then travel dramatically. The energy gap between cluster nodes are close to the sink & the cluster nodes that are further off [5] are another drawback of a WSN single sink. Network is reorganized with increasing no. of linked nodes to sink. A new algorithm for network optimization is suggested by the latest research in a many sink WSN for decrease the utilization of energy & improve the operation life of the network. This network optimization energy management optimizes the life of the network. There are also very fewer numbers of non-connected nodes [6].

A WSN is typically based on a multifaceted communication model that involves multiple sensors that transfers data to one specific sink. The proposed research indicates that multiple sinks in the WSN are of significant. A routing path between a SN and a sink can be shortened to have more than one sink. If the sensing field is enormous or sensors are used arbitrarily, then very efficient routing protocol fails. Although clustering is utilized, nodes are inefficient of energy transmitting data to long distances [7]. The installation of several sinks enhances WSN to allow multiple sinks to take over and avoid WSN from crashing even though one of the sinks dies out. The best way to exchange data is to directly relay nodes to the sink node or base station (BS). But, a node is easily decayed by an unsustainable power consumption if distance among sink & network is high [8]. The
algorithms cluster reduces unexpected resource usage by sorting the network into clusters while transmitting the data toward the BS. A CH that transmits data to BS is allocated to each cluster. Mechanism for election by CH [9] should have even an allocation of energy among cluster member nodes is an essential part of clustering algo [10].

The cluster head energy consumption, data transmission and increasing network lifetime depend upon the cluster strategy i.e. uniform cluster (UC) and non-uniform cluster (NUC). The imbalanced energy utilization between nodes & node placement is the main factor affecting cluster lifetime. Energy utilization between nodes depends on cluster types. To balance energy in uniform clustering entire clusters had an approximate number of nodes and the same coverage area. So, that energy consumption between cluster head and members are to be less. Finally, prolong the network lifetime because of the uniform distribution and cluster size [11].

The optimized selection of CH for the acquisition of energy efficiency at WSN is of primary concern with the hotspot issue where the early death of overall network life is inevitable in wide area of the network. Because of these issues the efficiency and stability of the network are degraded. The proposed approach improves heterogeneous sensor networks and enhances the network life.

The article is organized in as following sections. Section II contains the requisite context and research to related topic. Section III addresses our proposed model & algorithm. Section IV analyzes the results of the simulation. Finally, in Section V the article is concluded.

II. LITERATURE SURVEY

Xiu-wu et al., (2020) for heterogeneous WSNs a clustering routing algorithm based on the WOLF pack algorithm (CLWPA) is suggested to improve HWNS & efficiently control the network costs. In the first instance, heterogeneous nodes have been optimally implemented to become a concern of mixed-integer programming. The method is achieved by using the WPA to boost logistics & levy flight based heterogeneous Routing (LWPA) algo is suggested. Second, the edge grade principle is implemented to strengthen the DEEC algorithm to address the issue of preset paths in the LWPA. Improved DEEC algo is utilized to cluster the heterogeneous network-specific nodes dynamically, & data transmission is done in clustering node collection. Finally, the CLWPA algorithm significantly increases the stable time and life of the network in comparison with three other heterogeneous network routing algorithms and improves energy efficiency [12].

Verma et al., (2019), this paper is intended in the Genetic Algorithm Optimized Clusters (GAOC) protocol to optimize the collection of CH with the incorporation of residual energy, sink distance & node Density in their formulated fitness function. Indeed, multiple data sinks depends on the GAOC (MS-GAOC), is proposed for an agreement on the Hot Spot problem and a reduction of the communication gap between nodes and sinks. The MS-GAOC methodological review takes place with protocols built for different data sinks so that fair comparative research can be carried out. The GAOC with Ms-GAOC surpass up to date standards in the area with various performance measures that were inferred from a simulation study which are Stability, network life, no. of dead node round, throughput & residual energy of the network [13].

Zhang et al., (2019) uses fuzzy reasoning to take into account several clustering considerations to prolong network life. Important considerations are considered i.e. the relative density of network nodes and relative length from nodes to the base station, along with initial energy for a complex range of the cluster center. Proposed clustering algo shows that it can regulate the energy consumption of network nodes but efficiently increase the network’s survival duration, thereby guaranteeing the accuracy of the data aggregation [14]. Baradaran and Navi (2019) provides a framework for the generation of high-quality clustering algorithms (HQCA). HQCA technique utilizes metric for cluster efficiency calculation that can increase intercluster and intracluster distances and reduce the error rate through clustering. Optimized CH is focused on fuzzy logic and is focused on specific criteria for residual energy, lowest and highest energy levels in each cluster, minimal & max distances amid cluster nodes & BS. The key benefits for large-scale networks are their highly efficient clustering, a low error rate, independence from key CHs improved scalability and efficiency. The significance of the consistency of clusters is also evaluated by external and internal parameters. Results of simulation revealed that the HQCA-WSN approach would increase energy efficiency and network life dramatically. In contrast to existing approaches, this approach considerably increases the FND & LND metrics [15].

Almogahed and Abdelrahman (2018) they also proposed an ODEC (Optimized Distributed Energy Intensive Clustering System) to expand the length and reliability of the network. This approach suggests an optimization of CH selection by increasing the heterogeneity probability function. The network area is categorized in two section i.e. inner & outer. The sensor nodes opt to be CH nodes or not depending on their Base Station (BS) locations and heterogeneity probabilities. Simulation tests demonstrate that ODEC achieves better efficiency in terms of quality as well as network life and reliability between related protocols [16].

Hu et al., (2018) supported an underwater heterogeneous sensor network (UHSN) mobile sink routing scheduling process. Firstly, clustering topology NW in that sensor nodes gather narrow level data, onward to CH for energy through a multihop acoustic connection in the cluster, CHs gather broad level data, upload optical connection to BS, that can save power on the network but extend the network coverage. Finally, the monitor center will transmit a remote path plan guide to the intercluster network route on mobile sink and change its direction for network maintenance, with a view to remote control of BS through monitoring center. This paper obtained proficient data transmission through dual-modal UHSN communications as well as a mobile sink path plan but solved an underwater speed but distance dispute bottleneck problem [17].

Xie and Wang (2017) suggest an enhanced energy-intensive distributed clustering algo for heterogeneous wireless sensor networks (IDEEC). A multi-level energy model is considered by IDEEC. The likelihood threshold is reduced, the likelihood of placement of cluster head is increased and an average energy calculation is optimized for network. Also, IDEEC takes the least running time which creates it easier for users to implement it in practice, in comparison with present clustering Protocols, in terms of reliability, amount of msgs, variance & mean of CHs. [18].
Clustering is an optimized strategy used to fulfill the exclusive performance specifications of broad level WSNs. The researcher performed a performance study of clusters-based WSN with various communicating patterns. In [19], a fuzzy clustering, k-mean and c-mean clustering technique. It is seen that overhead in cluster-based protocol isn’t a lot of ward upon up-to-date time. Simulation outcome demonstrates that a cluster-based protocol has low correspondence overheads contrasted and the speed based protocol.

In HWSN, CH selection is most demanding research area for selecting one node as a lead node from various cluster member node which collect the information from these cluster members and send to it BS. The sensed information sent in an energy efficient manner to the CH from cluster members and to BS. Since we know each sensor node have their limited battery lifetime so there is need of enhancing network lifetime as in term of residual power. In section II, we present various energy efficient CH election mechanisms and provide a systematic literature to identify the problems on that work is going on. At first, we start the study with some of the protocols that are Genetic Algorithm Optimized Clusters protocol, fuzzy based clustering, high-quality clustering algorithms, distributed clustering algo, Optimized distributed Energy Intensive Clustering System, multi-hop acoustic clustering and WOLF pack algorithm. From this literature study finding out that the energy utilization for CH election is utmost key issue in WSN. But the researchers and authors are not able to notice this issue wherever they can give their contribution to enhancing the energy efficiency. This study also defines the limitations of the existing algorithms that are barrier in the selection of energy efficient optimized CH in WSN.

III. PROPOSED METHODOLOGY

The optimized selection of CH for the acquisition of energy efficiency at WSN is of primary concern. There is also the hot spot where the premature death of network life is inevitable in a broad area of the network. The efficiency and stability of the network are also a greater issue that degrades the output of the network. The proposed work intends to reduce energy usage by implementing multiple sinks.

A. Methodology

We used the optimized clustering approach based on bat algorithms for single or multiple data sinks to solve these issues. CH selection consists of the metaheuristic bat algorithm to maximize the fitness functions globally by the integration of the distance parameter into the sink in BATOC. Inspired by microbats, echolocation behaviour, it has different emission and loudness pulse frequencies. BATOC protocols and MS-BATOC protocols are applied in two stages which are setup phase and steady state phase. During the clustering process of MS-BATOC, 4 data sinks are used outer to the network at the perimeter as opposed to a single sink.

(i) Setup phase: CH election is conducted in the network formation in this phase. The following steps are discussed.

Step 1. The entire cycle is initialized by the randomly distribution of 3 heterogeneous energy nodes that consist fundamentally of normal nodes, intermediate or advance nodes. Only one data sink is used for BATOC in wide network, while multiple of data sinks are mounted outer of network for MS-BATOC Protocol.

Step 2. After network creation, the selection of the CH using the Bat Algorithm is followed by the clustering. Distance variables are the criteria for the collection of CH used in the fitness functions.

(ii) Steady-state phase: Data transmission is started in steady-state mode following the selection of cluster head and steps below are described.

Step 1. Instant CH is selected; it distributes data from cluster member nodes on the basis of threshold method. The nodes broadcast information depend upon the constraints & are evaluated on every node of the cluster if present sense value (C(V)) is greater than the hard threshold (H(T)) specified above. That’s it, no communication is leded. But, in or else case, Further, the testing value is determined by deducting current sense value S(V) to formerly sense value (C(V)) to newly sense value, whether C(v) > H(t) for either node. Some will send data to the relevant CH, where the Check value is higher than the predefined soft threshold value (S(T)) till threshold conditions are met.

Step 2. If data are obtained by CH, they collected data forward it to the corresponding sink using BATOC protocol, while for MS-BATOC, by calculating least Euclidean distance among nodes & multiple data sinks then they decide sink is nearest to them on the basis of least Euclidean distance.

Step 3. The Energy of a node are tested to decide if they’re equal or less than zero, whether they are, nodes that are considered dead nodes and signified by D N(i) that range from i to N (total node number).

Step 4. Also, if some nodes are less than the total number, D N is raised by 1 and the next node will be preceded by the CH selection again. However, the network ceases running if all nodes are dead.

B. Bat algorithm

Bat algorithm was taken very attentive by scientists as a modern intelligent search algorithm, developed and applied further. By echolocation action, the bat algorithm optimizes the purpose of an optimal function. During flight update duration, loudness, location and pulse rate, the following bat populations are represented mathematically. Bat also has frequency modulation which automatically switches from global to local search when the condition is satisfied and can control dynamically the interplay between global and local searches. Algorithms have basic model characteristics, fast convergence, parallel processing, straightforward working, and good parallel power [20].

To formulation of new Bat algorithm is depends upon bats behavior [20]. When implement the bat algorithm then there are following rules which are used:

1. By using of echolocation all bats sense the distances and guess the variation between their prey & obstacles of surroundings magically.
2. When bat searching his prey then y bats use the v_i velocity at random at x_i location with predefined frequency f_{min}, as well as unstable wavelength alongside loudness of A_0. The wavelength and pulse emission rates can be automatically adjust by the bats from emitted pulses (i.e. pulse emission r > [0; 1]) and this pulse emission rate depends upon the target & its proximity.
3. While the loudness will vary, we will presume it ranges from broad (positive) A0 to minimum constant Amin value.

The fate of an individual in BA is determined by the importance of fitness which is further used for the design of the fitness function. When interacting with
other nodes or sinks located somewhere distant, nodes consume their energy. Increased energy consumption covers more distance to node travels to connect and vice versa. Fitness Parameter (FP) for FF of pacts with distance factor for selection of CH is therefore given by Eq. (1).

$$FP = \sum_{i=1}^{N} \left( \frac{D_{N-S(i)}}{D_{AVG(N-S)}} \right)$$  \hspace{1cm} (1)

$$D_{AVG(N-S)} = \frac{\sum_{i=1}^{N} D_{N-S(i)}}{N}$$  \hspace{1cm} (2)

FP computes the cost of distance acquired for each $i^{th}$ node wherever $i$ extend of 1-N (total no. of network nodes). $D_{N-S(i)}$ is Euclidean distance from the sink of an $i^{th}$ node in Eq. (1), while Euclidean is $D_{AVG(N-S)}$ of the farthest node and sink. $D_{AVG(N-S)}$ denotes average Euclidean distance of all nodes to sink in Eq. (2). $D_{AVG(N-S)}$ is implemented in favor of the collection by CH of nodes of less than or equal to $D_{AVG(N-S)}$ near the sink. More FP value is observed as more likely to be elected as CH by node as it certify the minimum probable distance between sinks [21].

Fig. 1. Flow chart of bat algorithm.

C. Proposed Algorithm

Step 1: Initiated the network

Step 2: Three heterogeneous energy levels (the normal nodes, intermediate nodes & advanced nodes) are randomly deployed.

Step 3: A single data sink is used in BATOC in centre of network and MS-BATOC comprises multiple data sinks outside the network.

Step 4: Apply bat algorithm based optimized clustering after formed the network

Step 5: Cluster head (CH) election also done by taking distance factor as a fitness function in Bat algorithm

Step 6: CH disseminates data on threshold response from cluster nodes

Step 7: If data are obtained by CH, they connect the data forward it to the corresponding sink using the BATOC protocol, while for MS-, by calculating a minimum Euclidean distance between nodes but several data sinks, they decide sink is nearest to them.

Step 8: Check energy of every node to calculate the number of dead nodes

Step 9: If selected CH nodes are dead then perform re-clustering & select the CH

Step 10: network stops running if all nodes are dead

Step 11: Stop

Fig. 2. Proposed Work Flowchart.

IV. RESULT AND DISCUSSION

The simulation experiment of this proposed work done by MATLAB 2018 for analyzing result. BATOC & MS-BATOC modeling research was carried out based on performance measurements. First is to analyze simulation parameters. The simulation results would then be evaluated against various data sinks. Table 1 lists simulation parameters for BATOC & MS-BATOC.
Table 1: Simulation Parameters.

<table>
<thead>
<tr>
<th>Network model</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of Network Area</td>
<td>100 *100 m²</td>
</tr>
<tr>
<td>No. of Nodes</td>
<td>100, 200</td>
</tr>
<tr>
<td>BATOC &amp; MS-BATOC number of data sinks</td>
<td>1 &amp; 4</td>
</tr>
<tr>
<td>Nodes initial energy</td>
<td>0.5 J</td>
</tr>
<tr>
<td>Node type energy heterogeneity</td>
<td>regular, mid-level or advanced</td>
</tr>
<tr>
<td>Power requires for transmitter or receiver service</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>Threshold distance</td>
<td>87 m</td>
</tr>
<tr>
<td>The energy essential for amplification is lesser</td>
<td>10 PJ/bit/m²</td>
</tr>
<tr>
<td>The energy essential for the amplification is higher</td>
<td>0.0013 PJ/bit/m²</td>
</tr>
<tr>
<td>No. of the intermediate (m) &amp; the mature (mo) nodes</td>
<td>m = 0.1, mo = 0.2</td>
</tr>
<tr>
<td>Intermediate (β) or intermediate (α) nodes energy fraction</td>
<td>β = 1, α = 2</td>
</tr>
<tr>
<td>Energy consumption after processing of data</td>
<td>5 nJ/bit/signal</td>
</tr>
<tr>
<td>size of Data packet</td>
<td>2000 bits</td>
</tr>
<tr>
<td>Population Size (P)</td>
<td>20</td>
</tr>
<tr>
<td>Population Generations</td>
<td>500</td>
</tr>
<tr>
<td>Loudness</td>
<td>0.5</td>
</tr>
<tr>
<td>Pulse rate</td>
<td>0.5</td>
</tr>
<tr>
<td>Frequency range</td>
<td>0-2</td>
</tr>
</tbody>
</table>

A. Simulation Analysis
First of all, we have made a GUI for input the network coordinate values shown in Fig. 3. In this enter length and widths of the yard also enter the number of nodes.

Fig. 3. GUI.

BATOC & MS-BATOC data transfer scene is shown in Fig. 5 (a) and (b), respectively. In the centre of BATOC sink while four sink are situated in MS-BATOC.

Fig. 4. Network initialization.
(a) BATOC
(b) MS-BATOC
Fig. 5. Data transmission scenario.

(a) BATOC
(b) MS-BATOC
Fig. 6. Comparison of Alive nodes vs rounds.

Number of dead nodes against rounds: The efficiency of the network is measured according to factor that provides the position of some dead nodes during rounds. Network output is to be improved if no. of alive nodes in excess of several rounds is exceeded. In Fig. 7 depicts a comparison of dead nodes over no. of rounds for a single sink and multiple sink based on bat algorithm.
Network’s Energy Variance: Total capacity, i.e. number of all node’s energy resources, is slowly decreased by the energy consumption of nodes when interacting with other nodes or with sink as data transmission progresses. In Fig. 8 displays variations in energy over the number of rounds.

Throughput: No. of data packets transferred efficiently to the sink is called throughput. In Fig. 9 represents the number of packets over no. of rounds.

V. CONCLUSION
Optimized selection of CH is of vital important topic for WSN. So that this can be learned, two new protocols for reactive routing are introduced, namely BATOC and MS- BATOC for Heterogeneous WSN. The fitness function in both protocols is used to shape parameter distance to sink. Thus simulation and comparative study of BATOC/MS-BATOC protocols with existing GAOC & MS-GAOC were carried out based on the different metrics of performance. Detailed tests found that BATOC increased performance or network life in compared to GAOC protocol. This change is accomplished by considering distance fitness parameters. The throughput of MS-BATOC compared to aforementioned protocol is also substantially improved. This increase is accomplished not only by raising efficient communication distance among nodes as well as the corresponding sink node except as well by distance-based selection of CH.

REFERENCES


