CHAPTER 6

CUCKOO SEARCH FOR DATA GATHERING IN SENSOR NETWORKS

6.1 INTRODUCTION

Clustering and classification techniques, a method of hierarchical routing afford a new dimension to aggregation of data by considering with factors like routing update frequency, routing algorithms and network configuration. Hence efficient data clustering techniques must be used to reduce the data redundancy and in turn reduce overhead on communication (Akojwar and Patrikar 2008). Inspired by the collective behavior of biological species, self-organizing clustering algorithms for large-scale networks are developed. By using simple rules most of the nodes in a network can determine their role as either cluster head or cluster member based on purely local decisions and by using only a limited amount of broadcast messages in the network (Jacobsen et al 2011). In this research, new paradigm of computational intelligence, the Cuckoo search is employed for data aggregation in Wireless Sensor Networks. Cuckoo Search is applied for both cluster formation as well as routing of gathered information to the base station. The traditional Cuckoo Search is modified as per the requirements of the proposed problem.
6.2  CUCKOO SEARCH

Cuckoo search (CS) is an optimization algorithm developed by Xin-She Yang and Suash Deb (2009). It is a novel algorithm which is inspired by the obligate brood parasitism of some cuckoo species by laying their eggs in the nests of other species. In the multi dimensional space where the optimal solution is sought, the CS is carried out for a maximization problem, where the quality or fitness of a solution can simply be proportional to the value of the objective function. Cuckoo Search has similarity to the hill climbing algorithm. It is more efficient than Genetic Algorithm and Particle Swarm Optimization to adapt to wider class of optimization problems (Yang and Deb 2010). The CS is based on three idealized rules:

1. Each cuckoo lays one egg at a time, and dumps its egg in a randomly chosen nest.

2. The best nests with high quality of eggs will carry over to the next generation.

3. The number of available host’s nests is fixed, and the egg laid by a cuckoo is discovered by the host bird with a probability $P_a$. The worst nests are discovered and dumped from further calculations.

The fitness function is related to the difference in the solutions. The new solutions are updated towards the optimal point by adding the product of levy angle and step size to the existing solution. $X$ is related to the new solution, for an $i^{th}$ cuckoo and then levy is performed as given in Equation (6.1) and (6.2).

\[
X_i^{(t+1)} = X_i^{(t)} + \alpha \cdot \text{Levy}(\tilde{n}) \quad \text{or} \quad X_i^{(t+1)} = X_i^{(t)} + \alpha \cdot E_t
\]  

(6.1)

(6.2)
\( E_t \) is drawn from a standard normal distribution with zero mean and unity standard deviation for random walks, or drawn from Levy distribution for Levy flights. \( \alpha \) is the step size and is greater than zero. Step size is usually taken as 1 in most of the cases. Equations (6.1) and (6.2) are more sequential in regard to the pseudo code. It can also be written in biased way with random step sizes (Yang 2010). Equations (6.3), (6.4) and (6.5) gives cuckoo search’s biased random walk. Step size is determined as given below.

\[
\text{Step size} = \text{rand}\ast(\text{nest} (\text{randperm}(n,:),\text{-nest(randperm(n,:))}); \quad (6.3)
\]

\[
\text{new_nest=nest+stepsize.*K} \quad (6.4)
\]

where; \( K = \text{rand (size (nest)) > pa} \) \quad (6.5)

The pseudo code for cuckoo Search is given below.

Objective function: \( f(x) = (x_1, x_2, \ldots, x_d); \)

Generate an initial population of \( n \) host nests;

While \( (t<\text{Max Generation}) \) or (stop criterion)

Get a cuckoo randomly (say, \( i \)) and replace its solution by performing Lévy flights;

Evaluate its quality/fitness \( F_i \)

[For maximization, \( F_i \propto f(X_i) \)];

Choose a nest among \( n \) (say, \( j \)) randomly;

if \( (F_i > F_j) \),

Replace \( j \) by the new solution;

end if

A fraction (pa) of the worse nests is abandoned and new ones are built;

Keep the best solutions/nests;
Rank the solutions/nests and find the current best;

Pass the current best to the next generation;

end while

Post-processing the results and visualization;

6.3 PROBLEM FORMULATION AND NETWORK MODEL

As stated in section 2 of chapter 4, the cluster based routing protocols have in common cluster formation phase and routing phase. The methodology adopted in cluster formation and technique employed for routing varies as per applications. The following are the most widely used assumptions and model in sensor network simulation and analysis. Hence the cluster formation, assumptions and radio model considered in this research analysis are generalized as follows.

1. Nodes are dispersed randomly following a uniform distribution in a 2-dimensional space and the location of the Base Station (BS) is known to all sensors. The BS is considered as a powerful node having enhanced communication and computation capabilities without energy constraints. It is located outside the deployment area.

2. The nodes are capable of transmitting at variable power levels depending on the distance to the receiver. The nodes are unaware of their location, i.e.: they are not equipped with any Global Positioning System device.

3. The nodes can estimate the approximate distance by the received signal strength if the transmit power level is known, and the communication between nodes is not subject to multi-path fading.
4. A network operation model similar to that of LEACH and HEED is adopted here, which consists of rounds. Each round consists of a clustering phase followed by a data collection phase.

5. The sensor nodes are assumed to be capable of switching between sleep and active modes. The sensor nodes are immobile or stationary after deployment.

6. The energy model presented by Heinzelman et al. (2002) is adopted here.

6.4 CLUSTER FORMATION

The cuckoo-based data gathering consists of three main steps namely, initialization of parameters, formation of static clusters and communication to the base station.

Step 1: Initialization

Select the number of sensor nodes, cuckoo nests, eggs in nests to start the search. Each nest has multiple eggs representing a set of solutions. Initialize the location and energy of nodes and the location of base station.

Step 2: Formation of Static Clusters

The clusters are formed by Cuckoo Search technique. Each egg in a nest corresponds to a sensor node. A group of M nests are chosen with N eggs in it. The probability of choosing the best egg or quality egg is done by random walk. Step size and Levy angle is updated. In turn the nests are updated. The optimal solution i.e.; best egg – high energy node is taken as Cluster Head in context to energy, distance between the nodes and distance to the base station.
The worse nets are abandoned in normal Cuckoo Search. In order to compensate the unequal energy dissipation, the worse nets (or) least energy nodes are allowed to join the cluster as non cluster head nodes in the proposed approach. The least energy nodes join the proximity cluster heads to form cluster. The cluster formation is done by appropriate advertisement of Cluster-Head to all other nodes to join a particular cluster. The cluster head is not permanent. In each run, according to the residual energy of the nodes, the cluster head is periodically changed. This helps to eradicate the communication overhead and redundancy. The search tries to manage the uneven energy dissipation considering the least and maximum energies. This will result in lesser percentage of dead nodes in higher rounds. The cuckoo search doesn’t deteriorate with scalability and time. The fitness function \( f \) is considered in this problem for minimization of energy and maximization of lifetime of the nodes.

\[
f(dfi) = \sum_{i=1}^{n-1}(100 \times di) \quad (6.6)
\]

‘dfi’ is the energy function. The above equation is derived from Equation (3.1), by considering the value of \( \epsilon_{\text{mp}} \) as 100 pJ/bit/m\(^2\) for \( n=2 \), i.e.; communication range between the sensors.

After the clusters are formed, the Cluster Heads (CHs) fuse or aggregate the information before forwarding it to the base station. The energy model incorporates free space radio model followed by all nodes. The inter cluster and intra cluster routing via shortest path is to be performed based on the application. Intra cluster refers to communication between cluster head and non cluster head nodes within the cluster. Inter cluster communication refers to communication between the clusters. For intra cluster communication, the most widely used methodology as followed by the basic LEACH algorithm concept is TDMA Scheduling- Time Division Multiple Access Scheduling is adopted.
6.5 MODIFIED CUCKOO SEARCH

In the proposed Cuckoo Search the basic assumptions are made as follows. For inter cluster communication, Cuckoo Search routing scheme is explained as follows. The objective of using CS is to optimize route and extension of the network lifetime. The cuckoo search is modified by considering cognitive ability of the bird in routing.

Pigeon Race:

The Pigeons are set for race to reach a particular city (target) and return back to the source point. Each year the race is conducted and the pigeons are set to fly. The pigeons remember the shortest path by having the experiences of worst and best routes. Similar to Pigeons, the cuckoo bird is assumed to have memory of its previously visited best and worst routes. When the cuckoo is aware of its path, along with the worse nests, the levy flights and step size is varied. This gives additional capability to explore the search space and results in more optimal solution. The levy angle K is calculated using Equation (6.7).

\[
k = \text{rand}*[(\text{best route (randperm (n,:),)} - \text{rand (size (nest))})
+ (\text{rand (size (nest)) -bad route (randperm (n,:))})]
\]

(6.7)

Step 3: Communication to the Base Station

Normally, communication between two Wireless Sensor Nodes happens, when there is no other interfering node between two nodes. It is assumed that there exists wireless path and link between two nodes during communication. The node, (specifically the Cluster-Head) with the information to the base station initiates the routing task by transferring the data packets to the nearby Cluster-Head. Similarly the transactions take place from
X Cluster to Y cluster to base station. In Cuckoo Search the cuckoo tries to find the optimal path by minimum levy angle and random walk. Cuckoo traverses from a source node (Cluster) to the base station by travelling through neighbour clusters. The advantage of CS is the redundancy to check the ‘n’ search space will be eliminated and the optimal solution is obtained. After several numbers of iterations, the optimal path to transmit the data to the base station is being established.

### 6.6 SIMULATION RESULTS AND ANALYSIS

The modified Cuckoo Search algorithm is developed and deployed for energy efficient data collection in Sensor Networks. The performance of the Network is analyzed using MATLAB 7.0.0.19920(R14). Current direction, state probability, foraging behavior of the cuckoo birds are taken into consideration for cluster formation and routing of information to the base station. The number of nodes is related with the number of eggs in a nest. The nests are related to the number of clusters. The probability of finding the host egg ($P_a$) is set as 0.2 for 100 numbers of nests. The number of eggs in nests is set as 1-3. 100 simulation runs are performed and the best run is given as results. The simulation Parameters are listed in Table 6.1.

**Table 6.1 Simulation Parameters for Cuckoo Search**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor deployment area</td>
<td>100 m *100m</td>
</tr>
<tr>
<td>Base station location</td>
<td>(50m, 150m)</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>100-200</td>
</tr>
<tr>
<td>Data Packet size</td>
<td>100 bytes</td>
</tr>
<tr>
<td>Control Packet size</td>
<td>25 bytes</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>0.5J- 2J</td>
</tr>
<tr>
<td>$E_{\text{electrical}}$</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>$\varepsilon_s$</td>
<td>10pJ/bit/m²</td>
</tr>
<tr>
<td>$\varepsilon_g$</td>
<td>0.0013pJ/bit/m⁴</td>
</tr>
</tbody>
</table>
The traditional methods HEED and LEACH are compared with the proposed scheme, with regard to the parameters relevant to network lifetime and energy consumption. Since the comparison models are LEACH and HEED, the network operation model similar to them is taken for analysis. In the analysis each round consists of a clustering phase followed by a data collection phase. Figure 6.1 shows the percentage of active nodes versus the number of iterations. The initial energy of the nodes is considered as 0.5 joules. The simulation is executed for 100 rounds. The percentage of nodes alive at the end of 100 iterations is 32% for cuckoo Search. This is due to the fact that new solutions are updated around the best solution, to accelerate the local search.

![Network lifetime](image)

**Figure 6.1** Network lifetime
In Figure 6.2 the network lifetime is compared with regard to node density. The first node death in a cluster of the entire network is evaluated for 200 nodes. At round number 58 the first node is exhausted of energy by cuckoo Search. The life time of a network is usually defined by the number of the nodes alive or percentage of nodes die. Figure 6.3 shows the number of rounds until the last node death. At round number 220, all nodes are exhausted of energy by LEACH protocol. In Cuckoo Search, at round number 255, 100 nodes are exhausted of energy. The time the first node and last node dies are significant in determining the lifetime of the network. Cuckoo Search produces comparable results mainly due to search process by far field randomization, without getting trapped in local optimum.

![Figure 6.2 Effect of Initial node density on Network lifetime](image-url)
Figure 6.3 Effect of node density on Network lifetime

Figure 6.4 Energy consumption per round
In Figure 6.4 the average energy consumed per round/iteration is shown. The average energy consumption for Cuckoo search is 1.6 times less compared to the LEACH protocol.

6.7 CONCLUSION

Cuckoo Search is applied for balancing the minimal energy dissipation among the Sensor nodes. All the nodes are utilized with equal importance to balance the energy dissipation. This approach incorporates two significant metrics that makes it favorable with respect to energy efficiency when compared to the original Cuckoo search. The proposed Cuckoo Search is compared with the standard LEACH protocol and HEED protocol. The simulation results exhibits that, Cuckoo Search enhances the proportion of active nodes by minimum of fifteen percent. The developed optimal algorithm reduces complexity in cluster formation and prolongs sensor network lifetime. The average energy consumption of nodes per round is within one to ten percent. The results comply with the literature after several iterations. The random biased Search offers consistency in the cluster formation, minimal number of clusters, average energy consumption and energy consumption per rounds. The proposed approach will be even suitable for unequal initial energy distribution to the nodes.