Improved Cluster Head Selection for Efficient Data Aggregation in Sensor Networks

G. Kavitha and R.S.D. Wahidabanu

1Department of Computer Science and Engineering, Excel Engineering College, Kumarapalayam, India
2HOD/ECE and Government College of Engineering, Salem, India

Abstract: Large-scale Wireless Sensor Networks (WSN) is the focus of recent research and development efforts. Due to their benefits in monitoring physical environments, WSN find diverse applications from military usage to agriculture and scientific works. To maximize WSN’s network life, data transfer paths are selected so that total energy consumed on the path is minimal. To ensure high scalability and improved data aggregation, sensor nodes are grouped into disjoint, non-overlapping subsets known as clusters. This study proposes improved Cluster Head (CH) selection for efficient sensor networks’ data aggregation. The suggested hybrid algorithm is based on Bacterial Foraging Optimization (BFO) and Gravitational Search Algorithm (GSA). The proposed hybrid BFO is incorporated in Lower Energy Adaptive Clustering Hierarchy (LEACH).

Keywords: Bacterial Foraging Optimization (BFO), Cluster Head (CH) selection, Gravitational Search Algorithm (GSA), Lower Energy Adaptive Clustering Hierarchy (LEACH), Wireless Sensor Networks (WSNs)

INTRODUCTION

A sensor network comprises of sensing (measuring), computing and communication elements that ensures an administrator can instrument, observe and react to events in a specific environment (Sohraby et al., 2007). Sensor networking is multidisciplinary involving radio and signal processing, networking, database management, artificial intelligence, systems architectures for operator-friendly infrastructure administration, power management algorithms, resource optimization and platform technology. WSN is a network of distributed autonomous devices that sense/monitor physical/environmental conditions cooperatively (Kulkarni et al., 2011). WSNs have many small, inexpensive, disposable and autonomous sensor nodes deployed in an adhoc way for remote operation in geographical areas.

Sensor nodes operate with batteries and are usually deployed in hard-to-access or hostile environments, occasionally in large quantities. It is hard or nearly impossible to replace sensor nodes (Anju and Sourav, 2013) batteries. However, the sink has rich energy. As sensor energy is a precious WSN resource, efficient energy use is needed to prolong network life and this has been focused in most WSN research. With WSN, objects can be tracked by tagging them with small sensor nodes which are tracked as it moves through a sensor node field deployed at known locations.

WSNs are used in applications like environment and habitat monitoring, surveillance, structural health monitoring, home automation, healthcare and traffic control. It is common in surveillance and monitoring to have WSN nodes dropped from helicopters (Kulkarni and Venayagamoorthy, 2010). On the other hand, this is not the case when the monitored area is a hostile/dangerous territory. As WSN include hundreds/thousands of low-power multi functioning sensor nodes, using them in unattended environments, with limited computational and sensing capabilities is a challenge. Sensor nodes have small, irreplaceable batteries with limited power. WSNs have hundreds/thousands of small, cheap, battery-driven, nodes with a wireless modem to monitor/control tasks jointly (Zaman et al., 2009). A major concern is network life: when nodes lack power, connectivity decreases and network is partitioned, becoming dysfunctional.

Data aggregation aggregates sensor data using aggregation techniques. Data aggregation is an essential paradigm for sensor networks (Krishnamachari et al., 2002) wireless routing, the idea being to combine data from various sources enroute-eliminating redundancy, reducing transmissions thereby saving energy. Data aggregation protocols combine and summarize sensor nodes data packets to reduce data transmissions. In WSN, data aggregation benefits increase if intermediate sensor nodes perform data aggregation incrementally when data is forwarded to base station (Ozdemir and Xiao, 2009).

WSN routing is challenging due to characteristics which characterize such networks from other wireless networks like MANETs or cellular networks. To begin with, the relatively large number of sensor nodes makes
it impossible to build a global addressing scheme for deploying many sensor nodes due to high overhead of ID maintenance. Hence, traditional IP-based protocols are not applicable to WSNs (Al-Karaki and Kamal, 2004). Further, sensor nodes deployed in an adhoc manner have to be self-organizing as adhoc deployment requires forming connections to cope with resultant nodal distribution especially when sensor networks are un-attended.

Usually, WSN routing is split into flat-based routing, hierarchical-based routing and location-based routing based on network structure. Nodes are given equal roles or functionality in flat-based routing (Mohanty, 2010). In hierarchical-based routing, nodes play different roles in network whereas in location-based routing, sensor nodes’ positions are exploited to route data in networks. A routing protocol is adaptive if system parameters are controlled to adapt to current network conditions and available energy levels (Fig. 1). Routing paths are established in one of the 3 ways; proactive, reactive or hybrid. Proactive protocols compute routes before they are needed, storing them in each node’s routing table. When routes change, it is propagated throughout the network. As a WSN has thousands of nodes, routing tables to be maintained by nodes will be phenomenal and so proactive protocols do not suit WSNs (Kumar et al., 2012). Reactive protocols compute routes when needed. Hybrid protocols combine these ideas.

Clustering is important in large multi-hop WSN mechanism to obtain scalability, reduce energy use and achieve better performance. Clustering offers lower communication overheads and efficient resource allocations thereby reducing overall energy consumption and lowering interferences among sensor nodes (Kumar et al., 2011a, b). Many clusters congest an area with small size clusters and a small number of clusters exhaust Cluster Head (CH) with large amount of messages from cluster members.

Sensor nodes are partitioned into different clusters in clustering with each cluster being managed by a node called Cluster Head (CH) and other nodes being cluster nodes. Cluster nodes will not communicate with sink node directly. They pass collected data to CHs which aggregate data from cluster nodes and transmit it to base station (Patole, 2012) minimizing energy consumption and messages sent to base station. It also reduces active nodes in communication. The result of clustering sensor nodes is prolonged network life.

Clustering techniques introduce heterogeneity to the network’s service profile with a side effect of creating nodes considered ‘more critical’ than others (Blace et al., 2008). This is because CH nodes serve as central hubs/super-peers for node management functions like communications, organization and security. In clustered networks, every sensor node forwards communications to CHs assigned to it.

CH election selects a node in a cluster as leader node. WSN and MANET nodes have limited battery and CH maintains information related to it. This information includes a node list in the cluster and path to all nodes. The CH’s responsibility is to inform nodes in its cluster (Hussain et al., 2013). CH must communicate with nodes of other clusters. Communication with other cluster nodes can be direct through respective CH or via gateways.

Adaptive Cluster Head Selection (ACHS) divides unequal cluster size into equal cluster size to balance energy consumed in local clusters (Nam et al., 2010). When member nodes in a local cluster are more/less than average member nodes number, this cluster becomes an unequal cluster. In deterministic schemes, sensor node special attributes like their identification number (Node ID), number of neighbours (Node degree) and in adaptive schemes, resource information like remnant energy, energy dissipated in last round and initial nodes energy decide their role during data gathering rounds (Ramesh and Somasundaram, 2011).
Optimizing CH selection techniques increased network life. Lower Energy Adaptive Clustering Hierarchy (LEACH) is a popular distributed single-hop clustering protocol (Kumar et al., 2011a, b) where clusters are formed depending on received signal strength. The CH role is rotated periodically amongst sensor nodes in a cluster to ensure balanced energy consumption. This algorithm is inefficient when large area sensor networks indulge in single hop CHs communication to sink.

This study proposes an improved CH selection for efficient sensor networks data aggregation. The new algorithm is based on Bacterial Foraging Optimization (BFO) and Gravitational Search Algorithm (GSA). Hybrid BFO is incorporated in LEACH.

**LITERATURE REVIEW**

Clustering mechanism in LEACH protocol was analyzed, presenting improved approaches based on sensor node energy and distance between node and base station by Liu and Wu (2012). Performances were compared. Simulation results indicated that improved protocols balanced network load and prolonged network life.

An energy efficient MAC protocol suitable for a WSN having mobile nodes was proposed by Srikanth et al. (2011). This was based on TDMA and CSMA and adjusted frame size to adapt to mobility and traffic condition changes. The proposed protocol’s performance was compared to LEACH-C protocol and it outperformed LEACH-C regarding average energy consumption, packet delivery ratio and network life.

Xiangning and Yulin (2007) put forward energy-LEACH and multihop-LEACH protocols. Energy-LEACH protocol improved choice method of CH, made nodes with more residual energy CHs in next round. Multihop-LEACH protocol improved communication from single hop to multi-hop between CH and sink. Simulation results show that energy-LEACH and multihop-LEACH protocols perform better than LEACH protocols.

The problems in LEACH routing protocol were analysed by Long-long and Jian-Jun (2010). An improved LEACH CH multi-hops algorithm was proposed. The proposed algorithm considered the premise of node energy and optimum number of CHs, selecting cluster node, through limiting the number of nodes in every cluster to balance each node’s energy depletion. It balanced energy consumption and prolonged sensor network life by the use of this algorithm. Simulation results proved the effectiveness of the proposed algorithm.

The LEACH-SM protocol which modified prominent LEACH protocol to extend WSN life by providing optimal energy-saving spare management, including spare selection was proposed by Bakr and Lilien (2011). LEACH-SM added spare selection phase to LEACH with the authors presenting a quantitative comparison of energy consumed and WSN life for both protocols.

A new improved method called LEACH-N based on LEACH was proposed by Li et al. (2011) in which the problem of node selection as CH node depended on cluster nodes residual energy. This guarantees rationality when selecting head nodes. It also enhanced network robustness and network life cycle. Simulation results revealed that the new algorithm performed better than LEACH regarding numbers of life nodes, energy consumed and data transmission.

A new method to choose cluster-heads based on original LEACH protocol that reduced unnecessary energy consumption on computing each node during every round was proposed by Zhao et al. (2012). As traditional selection formula neglected the change of nodes’ energy which makes them act as cluster-heads. Many times they die early due to too much energy consumption. To make network energy distribution more even it considers dynamic change of sensor nodes' energy during CHs selection. Simulations showed the improved protocol performing better than original LEACH.

A revised cluster routing algorithm named E-LEACH to enhance LEACH’s hierarchical routing protocol was proposed by Xu et al. (2012). In E-LEACH algorithm, the original CHs selection was random and round time for selection was fixed. In E-LEACH algorithm, sensor nodes remnant power was considered to balance network loads and changing round time depending on optimal cluster size. Simulation showed that the new protocol increased network life by a minimum 40% compared to LEACH algorithm.

An Improved-LEACH where residual energy and distance of node from base station were parameters for CH selection was suggested by Gajjar et al. (2012). The authors suggested starting the steady state node operation only if value sensed by a node was greater than set threshold value to save energy. Threshold value is set by end user at application layer. Improved-LEACH was analyzed both qualitatively and quantitatively. Both metrics presented for comparison framework could analyze trade-offs produced by varying WSN protocols and design guidelines for new WSN protocols.

**MS-LEACH** to enhance security of S-LEACH by providing data confidentiality and node to CH authentication using pair wise keys shared by CHs and cluster members was proposed by El-Saadawy and Shaaban (2012). The proposed MS-LEACH’s security analysis showed it having efficient security properties and achieving all WSN security goals compared to existing LEACH protocol’s secured solutions. MS-LEACH’s simulation based performance evaluation demonstrated the effectiveness of MS-LEACH protocol.
revealing that it achieved desired security goals and outperformed other protocols regarding energy consumed, network life, network throughput and normalizing routing loads.

An improved LEACH algorithm titled partition-based LEACH (pLEACH), that partitioned a network into optimal number of sectors and selected a node with highest energy as head for each sector, using centralized calculations was proposed by Gou and Yoo (2010). Simulation results and analysis revealed that pLEACH achieved better WSN performance regarding energy dissipation, network life and communication quality.

M cluster-heads in each cluster to obtain an M diversity order in long distance communication instead of selecting a single cluster-head at network layer was proposed by Kong (2010). Due to wireless transmission’s broadcast nature cluster-heads received data from sensor nodes simultaneously. This ensured synchronization to implement a virtual MIMO based space time block code in cluster-head for sink node transmission. An analytical method to evaluate energy consumed based on BER curve was presented. Analysis/simulation results proved the new cooperative LEACH protocol saved greater energy over LEACH protocol with same bit error rate, data rate, delay and bandwidth requirements. The proposed protocol also achieved higher order diversity with improved spectral efficiency compared to other virtual MIMO based protocols.

A novel Fault-Tolerant Target Tracking (FTTT) protocol based on clustering was proposed by Bhatti et al. (2011). The results of an investigation regarding performance overheads and FTTT protocol’s scalable nature via comparative simulations with LEACH were discussed. Overall, in cases, FTTT consumed minimum 25% less energy than LEACH.

A Mobility-Based Clustering (MBC) protocol for WSN with mobile nodes was proposed by Deng et al. (2011) in which a sensor node elected itself as cluster-head based on residual energy and mobility. Simulation results revealed that MBC protocol reduced packet loss by 25% compared to cluster-based routing protocol and by 50% compared to Low-Energy Adaptive Clustering Hierarchy-mobile (LEACH-mobile) protocol. It also outperformed cluster-based routing protocol and LEACH-mobile protocol regarding average energy consumed and average control overhead and adapted better to highly mobile environments.

**METHODOLOGY**

**LEACH:** LEACH is a popular clustering algorithm with distributed cluster formation for WSNs (Shankar and Shanmugavel, 2013) which randomly selects CHs and rotates roles to distribute energy consumed. TDMA/CDMA MAC are used by LEACH to reduce inter-cluster and intra-cluster collisions and data collection is centralized with specific periods. Most nodes transmit data to CHs in LEACH. The CHs collects and compresses data forwarding it to base station. In this protocol, all nodes organize into local clusters according to procedure, with many nodes acting as CHs and other nodes acting as cluster members (Sribala and Virudhunagar, 2013). The CH node uses more energy than member nodes. It runs many rounds in a network’s life, with each round having cluster formation and cluster steady phases. In cluster formation phase, many nodes act as local cluster-heads with specific probability.

All sensor nodes have same probability in LEACH protocol to be CHs, which make network nodes, consume energy in a relatively balanced way to prolong network life (Xiangning and Yulin, 2007). In the set-up phase, CH nodes are randomly chosen from sensor nodes and many clusters are constructed dynamically. In steady data transmission phase, member nodes of clusters send data to their CH, the latter compresses data received from member nodes and forwards compressed data to sink node. LEACH protocol periodically elects CH nodes, re-establishing them according to a round time thereby ensuring each node’s energy dissipation are relatively even.

LEACH divides a network into many constructed clusters by using localized coordination and control to reduce data transmitted to the sink and also to make routing and data dissemination scalable and robust (Singh et al., 2010). LEACH uses randomize rotation of high-energy CH position instead of a static selection, to ensure chances for all sensors to act as CHs and avoid battery depletion of individual sensor and dying quickly.

**Hybrid bacteria foraging optimization:** Optimization algorithms effectiveness depends upon its search capabilities. The search is of two types, global and local search. The BFO algorithm has strong local search abilities but weak in global search. GSA has good global search abilities. Thus, both BFO and GSA are combined to enhance the global and local search for solutions. During chemotaxis operations in BFO, random velocity is used. As it is a kind of blind search, it does not converge quickly. Therefore, GSA is incorporated to the random velocity in this proposed hybrid algorithm. The proposed method includes advantages of both BFO technique and GSA and also excludes disadvantages of BFO. A Hybrid BFO with GSA incorporated in LEACH was resorted to for optimizing the CH selection.

Bacteria Foraging Optimization (BFO) algorithm is a new biologically encouraged stochastic global search technique based on mimicking E. coli bacteria’s
Fig. 2: Flowchart for the BFO

Start

Initialization of variables

Elimination and dispersal loop

1 > N_{ed}

Reproduction loop counter k = k + 1

k > N_{re}

Chemotactic loop counter

j > N_{c}

Compute value of cost function for each bacterium as J(i, j) where I is bacterium number

J(i, j) < J(i, j - 1)

i = i + 1

Swim N = N + 1
Sw(i) = N

Tumble

Sw(i) > N_{s}
i > s

Stop
foraging behavior. This is used to locate, handle and ingest food. During foraging, a bacterium exhibits two actions: tumbling or swimming (Thomas, 2013). Tumble modifies the bacterium’s orientation. During swimming—the chemotaxis step—the bacterium moves in its current direction. The natural selection process eliminate animals with poor foraging strategies and favors propagation of genes of those with successful foraging strategies, as they are likely to have reproductive success (Sharma and Behal, 2013). After generations, poor foraging strategies are either eliminated or converted to good ones. Foraging led researchers to use it as an optimization procedure.

Bacteria move several steps in the same direction till reduction in nutrient density and then tumble to a new direction. In chemotaxis, the adaptive step length strategy is adopted to dynamically adjust the bacteria’s step size and this is average distance from neighbors (Chen et al., 2012). Also, in tumble, the angle of rotation is no longer random, but maximum value of lure-degree (ratio of node’s residual energy and transmission energy consumption). Steps in algorithm (Rajeshwari et al., 2012; Yan et al., 2012) as follows.

**Chemo taxis:** This process is achieved through swimming and tumbling via Flagella. Each flagellum is a left-handed coil configured (where it is connected to the cell) to rotate counter clock wise, as viewed from the flagellum’s free end looking toward the cell. It produces force against bacterium which pushes the cell.

**Swarming:** When a *E. coli* cells collection is located in the middle of a semisolid agar with single nutrient chemo-effecter (sensor), they shift from centre in a travelling ring of cells shifting up nutrient gradient formed by consuming of nutrient by the group.

**Reproduction:** Minimum healthy bacteria die and healthy bacteria split into two bacteria, located in same location making inhabitants of bacteria constant.

**Elimination and dispersal:** It is possible that in a local environment, lives of a bacteria population changes step by step (via nutrients consumption) or unexpectedly due to other influences. Actions occur so that all bacteria in an area are killed or a group isolated to a new part of the environment which has the effect of destroying chemotactic progress, but they also have the support in chemotaxis (Fig. 2).

Gravitational Search Algorithm (GSA) is based on Newton’s basic laws of motion where all solutions have mass where heavier an object serves the problem as its solution. All objects move towards heavier objects obeying gravitational laws. Such techniques search space for better solutions and so optimize problems being considered (Afaq and Saini, 2011). GSA’s main idea is considering an isolated masses system, where every mass represents a solution to a problem similar to a small artificial world of masses obeying Newton’s laws of gravity and motion. According to Newton’s law of gravitation, each particle attracts every other particle with a force directly proportional to product of masses and inversely proportional to square of distance between them.

GSA can be summarized in steps 1 to 6:

1. Generate initial population
2. Evaluate fitness of each agent
3. Update gravitational constant and objects having best and worst fitness
4. Calculate mass and acceleration for each agent
5. Update velocity and position of each agent
6. If the termination criterion is met then stop else repeat steps from 2 to 5

**RESULTS AND DISCUSSION**

An improved CH selection for efficient data aggregation in sensor networks is proposed. The proposed algorithm is based on BFO and GSA and incorporated in LEACH. Forty nodes with single base station in a 2 km² area are used for testing.

Figure 3 shows that the proposed BFO optimization has high throughput compared to the LEACH method. Proposed BFO optimization is better by an average of 10.57% than LEACH.

Figure 4 shows that the proposed BFO optimization has low delay in seconds compared to the LEACH method. Proposed BFO optimization is less by an average of 8.61% than LEACH.

![Fig. 3: Throughput](image)

![Fig. 4: Delay](image)
Figure 5 shows that the proposed BFO optimization has low data dropping compared to the LEACH method. Proposed BFO optimization is less by an average of 11.57% than LEACH.

**CONCLUSION**

LEACH is a widely used clustering mechanism which elects a probability model based CH. LEACH is based on probability model while some CHs may be close to others and are located at WSN’s edge. Inefficient CHs cannot maximize energy efficiency. A Hybrid BFO with GSA incorporated in LEACH was resorted to for optimizing the CH selection. Hybrid BFO provides better life for nodes compared to LEACH in addition to providing 100% live nodes for the longest time. Use of hybrid BFO method in WSN ensures that clusters is determined followed by introduction of optimal cluster-head responsible for transfer of data to database.

**REFERENCES**


