Improved Cluster Head Selection For Energy Efficient Data Aggregation In Sensor Networks

A. Rajasekaran
Research scholar, Department of Electrical Communication Engineering, CSVMV University, Enathur Kanchipuram, Tamil Nadu, India.
E-mail: arajasekaran@kanchiuniv.ac.in

Dr V. Nagarajan
2Professor and Head, Department of Electrical Communication Engineering, Adhiparasakthi Engineering College, Melmaruvathur, Tamil Nadu, India.
E-mail: nagarajanece31@gmail.com

Abstract
Wireless Sensor Networks (WSN) are a challenging emerging technology due to their scope, low processing power, and associated low energy. WSN routing differs from conventional routing in fixed networks. It lacks infrastructure, has unreliable wireless links, sensor nodes may fail and routing protocols have to meet tough energy saving requirements. Data aggregation in WSN effectively saves limited resources. The goal of data aggregation algorithms is gathering and aggregating data in an energy efficient manner so that network life is enhanced. Clustering is used to extend a sensor network life by reducing energy consumption. This work proposed a better cluster head selection in sensor networks for efficient data aggregation. The proposed algorithm is based on Local search and incorporated in Low Energy Adaptive Cluster Hierarchy protocol (LEACH).

Keywords: Wireless Sensor Networks (WSN), Low Energy Adaptive Cluster Hierarchy protocol (LEACH), Clustering, Cluster Head (CH) Selection.

Introduction
WSNs are node collections where every node has its own sensor, processor, transmitter, and receiver. Such sensors are low cost devices performing a specific sensing task. Being of low cost, they are deployed densely throughout an area to monitor specific events [1]. Recent advances in Micro-Electro-Mechanical Systems (MEMS) technology, digital electronics, and wireless communications lead to the development of low-cost, low-power, multifunctional sensor nodes small in size and communicating untethered in short distances. Sensor networks are a major improvement over traditional sensors deployed in the following ways [2]:

- Sensors are positioned far from an actual phenomenon, i.e., something known by sense perception. So large sensors with complex techniques are required to distinguish targets from environmental noise.
- Sensors performing only sensing are deployed. Their positions and communications topology are carefully engineered. They transmit time series of a sensed phenomenon to central nodes. Here, computations are performed, and data fused.

The sensor nodes position is not pre-determined which ensures random deployment in inaccessible terrain and disaster relief operations. But, it also means that sensor network protocols and algorithms must have self-organizing abilities. Another unique sensor networks feature is the cooperative effort. They are fitted with an on-board processor. Instead of sending raw data to nodes for fusion, they use processing abilities to carry out simple computations and transmitting only required and partially processed data [3].

Data aggregation collects and aggregates useful data and is a fundamental processing procedure to save WSN energy. It is an effective way to save limited resources. The goal of data aggregation algorithm is gathering and aggregating data in an energy efficient manner to enhance network life. WSNs have limited computational power, memory, and battery power, so increased application develops complexity which results in applications closely coupled with network protocols [4].

Data aggregation solves data centric routing’s implosion and overlap problems. Data from multiple sensor nodes is aggregated when they reach the same routing node enroute to the sink. Data aggregation is a WSN technique. Security issues, data confidentiality, and integrity become vital when a sensor network is in a hostile environment. Data aggregation aggregates a sensor data using aggregation approaches [5].

Figure 1: Architecture of data aggregation.
Many routing algorithms were established for wireless networks [6]. WSN routing is challenging due to features that differentiate them from other wireless networks (mobile adhoc networks or cellular networks).

- Due to many nodes in these networks and the significance of overhead of ID upkeep, a global addressing scheme is impossible for sensor nodes formulation, thereby conventional IP based protocols cannot be used.
- Compared to other networks, all sensor network applications need flow of sensed data from multiple sources to specific BS but this does not prevent data flow in other forms.
- As sensor nodes are tightly stiffened regarding energy, processing, and storage capacities, they need careful resource management.
- In most applications, WSN nodes are generally stationary after deployment compared to conventional wireless networks where nodes are free to move, resulting in unpredictable and frequent topological changes. But, some applications permit sensor nodes to move and change location.
- As data collection is based on location, recognition of sensor nodes position is very important.
- As data collected by many WSN sensors is based on common physical processes, there is a high probability that data is redundant.

Wireless networks routing algorithms are topology based which is a traditional approach where forwarding decisions are based on information about current available links between network nodes [7].

- Proactive routing strategies: These track routing information of all available paths even when they are not used, but this strategy does not perform well when network topologies change dynamically.
- Reactive routing strategy: This tracks only routes currently in use due to device mobility or alternating energy conserving sleep cycles. But this may generate a huge amount of traffic when network topology changes.

WSN routing differs from conventional routing in many ways i.e. there is no infrastructure, sensor nodes may fail, wireless links are unreliable, and routing protocols meet exact energy saving requirements. Routing algorithms are classified as [8]:

- Topology-based: Routing algorithms with host-based addressing perform end-to-end message delivery
- Position-based: If a destination is given by an ID or if destination is a geographic location.
- Data-centric: this is based on queries issued by a sink node to request data. Requests are not addressed to particular sensor nodes, but they deliver requested data and answer the query [9].

Clustering extends a sensor network’s life by reducing energy consumption [10]. A sensor network is scalable by forming clusters. A cluster leader is referred to as ClusterHead (CH). A CH is elected by a cluster’s sensors or pre-assigned by a network designer. Many clustering algorithms were specifically designed for WSNs for scalability and efficient communication. The cluster based routing concept performs energy-efficient routing in WSNs. In a hierarchical architecture, higher energy nodes (CHs) process and send information while low energy nodes perform sensing. Low Energy Adaptive Cluster Hierarchy protocol (LEACH) [11], PEGASIS [12], APTEEN [13] and TEEN [14] are some clustering algorithms.

i. Clustering reduces routing table size stored at individual nodes by localizing a route set up in a cluster.
ii. Clustering conserves communication bandwidth as it limits scope of inter-cluster interactions to CHs avoiding redundant message exchange among sensor nodes.
iii. CH can prolong individual sensors battery life and network life by implementing optimized management strategies.
iv. Clustering cuts on topology maintenance overhead. Sensors care to connect with their CHs.
v. A CH performs data aggregation in its cluster and decreases redundant packets.
vi. A CH reduces energy consumption rate by scheduling activities in a cluster [15].

Comparison of various CH selection strategies regarding their assistance in CH selection, parameters used, required Re-Clustering (RC), even or fair Distribution of CHs (DCH), required cluster formation and Creation of Balanced Clusters (BCC) ensure a broader understanding [16]. A cluster based routing protocol is a group of sensor nodes where every node group has a CH. Sensed data is sent to the BS and not to the BS; CH performs aggregation on data and sends it to the BS where it is needed. LEACH [17] is a popular routing protocol using cluster based routing to reduce energy consumption. LEACH divides a communication process into rounds with each having a set-up phase and a steady-state phase. In setup phase, some sensor nodes are selected as CHs according to rules and others join clusters as member nodes. Within the steady-state phase, CHs collects data from their own cluster members and aggregate it before transmitting to the BS.

Optimization obtains the best results under given circumstances. The word ‘optimum’ means “maximum” or “minimum” depending on circumstances. Network optimization is critical, and the optimization techniques achieve networking design goals. Energy efficiency, cost, and application requirement are challenges to be taken care when designing a WSN. This needs hardware and software optimization to make WSN efficient. Software addresses the issue of Network Life [18]. There are many optimization algorithms to suit different problems. Choosing a proper algorithm is important in an optimization technique. Due to little infrastructure for WSNs, sensor nodes deployment is either inside a monitoring area or near it [19]. Batteries in sensor nodes are impossible to replace or recharge as sensor nodes are in remote or hostile sensing areas. So, end of a battery’s life in a node means the end of a network. Efficient use of battery energy is essential to enhance network life. Sensor network protocols focus on power conservation issues.
Other issues include low bandwidth, achieving high quality QoS, limited processing and storage [20]. This work proposes improved CH selection for efficient data aggregation in sensor networks. The proposed algorithm is based on Local search and incorporated in LEACH.

**Literature Review**

A cluster within a cluster of sensor nodes was proposed by Deshpande and Patil [21]. The CH acts as master of a cluster and master-ship is rotated among cluster heads after specified number of communication rounds. This improved sensor network’s energy use, maximizes network life, and makes a WSN fault tolerant to some extent.

A method for clustering using fuzzy logic with suitable inputs and combining it with good LEACH features proposed by Dastgheib et al., [22] is fully distributed. So its speed is more, and its energy consumption less than centralized methods. This method resolves LEACH’s weaknesses, and is more efficient than present methods.

An Energy Balanced Clustering (EBC) in WSN was proposed by Nazir and Hasbullah [23]. Algorithms for CH selection, inter cluster and intra cluster communication, and energy balanced cluster formation, in WSN were proposed. Using OMNet-4.0 simulation, the performance of the new protocol is compared with EEMC and LEACH using parameters like throughput and energy per packet. Simulation proved that EBC is effective in prolonging network life, and improving throughput, than LEACH and EEMC.

A multi hop routing technique using residual energy to prolong node and network life was proposed by Bhattacharjee and Bandypadhyay [24]. This technique generates an energy efficient routing path from every sensor to base station to send data. Simulation showed that the approach effectively conserved CHs energy, and cluster members prolonging their life effectively. This method also reduced clusters, and improved nodes lives significantly.

The comprehensive theoretical aspects of clustering issues to energy optimization in WSN were studied by Dutta et al., [25]. WSN sensor node mobility is a key advantage of wireless over fixed communication system. WSNs clustering techniques compared to random sampling is less costly due to saving time in journeys, reduction in transmissions, and reception at every node, identification, contacts etc. It is also valuable for increasing overall network life and WSNs scalability.

A Cluster based Energy efficient Routing (CBER) algorithm, electing CH based on nodes near an optimal CH distance and nodes residual energy was proposed by Mammu et al., [26]. The optimal CH distance linking optimal energy consumption is derived. Also, residual energy is considered in CH election to increase network life. Performance showed CBER reduced end to end energy consumption and prolonged network life of multi hop network compared to well-known clustering algorithms like LEACH and HEED.

Akbari and Beikmahdavi [27] surveyed cellular and cluster-based architecture to sustain network operations during failure of energy-drained nodes. Failure detection and recovery recovers the cluster structure in less than one-fourth of time taken by a Gupta algorithm and is also 70% more energy-efficient than it. Recovery and fault detection in a distributed manner allows a failure report to be forwarded across cells. This algorithm was compared to existing related works and proved to be energy efficient.

A heuristic approach based on Eigenvector centrality for cluster size control which was called Ev-CSC proposed by Jain and Reddy [28] is applicable to any deployment, traffic pattern and node types. Results demonstrated that the new method enhanced performance of respective equal clustering methods and performed better compared to cluster size control methods.

A hybrid clustering protocol Hybrid Distributed Hierarchical Agglomerative Clustering (H-DHAC) using quantitative location data and binary qualitative connectivity data in WSN clustering was proposed by Zhu et al., [29]. Simulation results reveal that H-DHAC has a slightly lower percentage of compromise in performance regarding network life and total transmitted data compared to approaches that use complete location data. The H-DHAC outperforms well known clustering protocols like LEACH and LEACH-C.

An effective WSN aggregation technique designed to suit air quality monitoring applications was developed by Abdul salam et al., [30]. The solution is LEACH based for WSN. Results showed the solution’s effectiveness in extending network life while monitoring air quality efficiently. A novel DE based algorithm for WSNs to prolong network life by preventing faster death of highly loaded CHs was proposed by Kuila and Jana [31]. This paper incorporated a local improvement phase to traditional DE for faster convergence and improved performance of the new algorithm. Results proved the new algorithm’s efficiency.

A distributed clustering algorithm for lossy WSNs with a mobile collector, where it moves close to every CH to receive data directly and uploads it to a BS was proposed by Gong et al., [32]. The algorithm uses multi-hop clustering to ensure better scalability. The one-hop clustering algorithm’s performance in small WSNs is close to optimal results by mathematical tools.

A state-of-the-art and comprehensive survey on clustering approaches by Afsar and Tayarani [33] presented clustering algorithms classification in WSNs. Some clustering objectives considered in this paper include fault-tolerance, scalability, data aggregation and fusion, load balancing, increased connectivity and collision avoidance. Also, it surveys the new approaches of the last few years in a classified manner and compares them based on metrics likemobility, cluster size, cluster count, and algorithm complexity.

Cluster formation was investigated by Mhemed et al., [34] who also suggested a new scheme called Fuzzy Logic Cluster Formation Protocol (FLCFP), using Fuzzy Logic Inference System (FIS) in cluster formation. It proved that using multiple parameters in cluster formation lowered energy.
consumption. This paper compared this technique to LEACH protocol to show that use of multi parameter FIS improves network life.

Taxonomy of various WSN clustering and routing techniques based on metrics like energy management, power management, network life, optimal CH selection, multi-hop data transmission etc. was provided by Tyagi and Kumar [35]. A detailed discussion is given in the text highlighting many prominent proposals’ relative advantages and disadvantages to help designers select a particular proposal based on its merits.

Methodology
A clustering algorithm for sensor networks, called LEACH introduced by Heinzelman, [11] is a popular hierarchical routing algorithm. The idea is to form sensor nodes clusters based on received signal strength and to use local CHs as routers to a sink. This saves energy as transmissions will be done by CHs and not all sensor nodes. Optimal number of CHs is 5% of total nodes [12]. All data processing like data fusion and aggregation are local to a cluster. CHs change randomly over time to balance nodes energy dissipation. This decision is by the node in choosing a random number between 0 and 1. It becomes a CH for current round when the number is less than the threshold given:

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

where \( p \) is a desired percentage of CHs, \( r \) - current round, \( G \) - a set of nodes not selected as cluster heads in the last \( 1/p \) rounds [1]. Nodes die randomly and dynamic clustering increases system life.

![Figure 2: Clustering in LEACH](image)

Figure 2 shows LEACH clustering. LEACH protocol runs with many rounds, each having two states: cluster setup state and steady state. In cluster setup state, it forms clusters and selects CHs; in the steady state, it transfers data. The time of second state is longer than the time of first state, to minimize overhead. LEACH is based on rounds and the system repeats clustering and transmission for each round.

A round has two phases:

**Set-up phase:**
- CHs are selected, based on \( T(n) \), threshold, 
- CHs broadcast ADV message to non-CH nodes 
- Non-CH nodes select CHs, based on RSSI of ADV message 
- After selecting cluster, the non-CH node sends Join-REQ to CH Now, CHs create TDMA schedule and sends it to all non-CH nodes

**Steady-state phase:**
- Sensor nodes begin sensing and transmitting data to CHs as per a TDMA Schedule 
- On receipt of data, CHs aggregates it to a BS in one-hop manner, reducing transmissions and saving energy 
- After some time, N/W reverts to set-up phase and enters another round 
- Each cluster communication uses different CDMA codes to reduce interference from other cluster nodes

**LEACH Protocol advantages**
- LEACH achieves a factor of 7 reduction in energy loss compared to direct communication and a factor of 4-8 compared to a minimum transmission energy routing protocol 
- Nodes die randomly and dynamic clustering increases system life 
- LEACH is completely distributed and needs no global network knowledge

Sensor nodes use irreplaceable power with limited capacity. Node’s computing capacity, communicating, and storage is very limited, needing WSN protocols to conserve energy as the objective of maximizing network life. LEACH periodically changes cluster membership and CH to conserve energy. A CH collects and aggregates information from sensors in its cluster and passes this to a BS. By rotating a CH randomly, energy consumption is uniform. But, LEACH chooses too many CHs at a time or randomly selects those far from a BS without considering nodes’ residual energy. So, some CHs drain energy early and reduce WSN life.

In literature, there are many relevant energy preserving techniques, which prolong network life. In this work, Local search technique optimizes CH selection based on node’s residual energy. The node energy model is based on [11]. The energy used to transmit \( n \) bit is computed by:

\[
E_{\text{transmit}} = n \times (\text{Energy}_\text{dissipated}_\text{transmitter}_\text{electronics} + (\text{Energy}_\text{dissipated}_\text{Transmitter}_\text{amplifier} \times \text{distance}_\text{squared}))
\]

The energy used up to receive \( n \) bit is computed by:

\[
E_{\text{receive}} = n \times \text{Energy}_\text{dissipated}_\text{receiver}_\text{electronics}
\]
Power consumed for a given time period \( t \) is given by:

\[ P_{\text{consumed}} = \frac{E_{\text{diss,Rx}} + E_{\text{diss,Rx}}}{t} \]

The probability of a node to become CH is one with the highest ratio of residual energy, which is computed as:

\[ \frac{E_{\text{max}} - P_{\text{consumed}}}{E_{\text{max}}} \]

Where \( E_{\text{max}} \) is the maximum energy of the battery of the sensor node.

Hill climbing is a mathematical optimization technique from the local search family. It is an iterative algorithm starting with an arbitrary solution to a problem, and trying to find an enhanced solution by incrementally changing a solution’s single element. If the change produces an enhanced solution, an incremental change is made to a new solution, repeating till no further improvements is found. Hill climbing is good to find a local optimum; a solution that cannot be enhanced by considering neighboring configuration, but is not guaranteed to find a best possible solution i.e. a global optimum from possible solutions (the search space). That only local optima are guaranteed is overcome by using restarts i.e. repeated local search, or more complex, iteration based schemes, like iterated local search, on memory, like reactive search optimization or memory-less stochastic modifications, including simulated annealing. The algorithm’s relative simplicity ensures that it is a popular choice among optimizing algorithms. Hill climbing produces better results than other algorithms when time for a search is limited, like with real-time systems.

**Results and Discussion**

The simulations are conducted for varying number of nodes (60-300) with single BS spread in a 2 sq km area. The simulations are conducted for LEACH, Cluster formation using local search. The Number of clusters formed, Average End to End Delay (sec), Average Packet loss rate, Lifetime computation, remaining energy computation are evaluated.

From Figure 3 it is observed that Numbers of clusters formed increases for Cluster formation using Local search than LEACH. When number of nodes is 120, Numbers of clusters formed increases by 14.29% for Cluster formation using Local search than LEACH.

![Figure 3: Numbers of clusters formed](image)

From Figure 4 it is observed that Average End to End Delay decreases for Cluster formation using Local search than LEACH. When number of nodes is 120, Average End to End Delay decreases by 19.89% for Cluster formation using Local search than LEACH.

![Figure 4: Average End to End Delay](image)

From Figure 5 it is observed that Average Packet loss rate decreases for Cluster formation using Local search than LEACH. When number of nodes is 120, Average Packet loss rate decreases by 10.68% for Cluster formation using Local search than LEACH.

![Figure 5: Average Packet loss rate](image)
Figure 6: Percentage of nodes alive
From Figure 6 it is observed that Percentage of nodes alive increases for Cluster formation using Local search than LEACH. When number of nodes is 500, percentage of nodes alive increases by 84.26% for Cluster formation using Local search than LEACH.

Figure 7: Average remaining energy in nodes
From Figure 7, it is observed that Average remaining energy in nodes increases for Cluster formation using Local search than LEACH. When number of nodes is 500, Average remaining energy increases by 40% for Cluster formation using Local search than LEACH.

Conclusion
Tiny sensor nodes having sensing, data processing, and communicating components, create the idea of sensor networks based on the collaboration of many nodes. A CH is elected by sensors in a cluster or pre-assigned by a network designer. Many clustering algorithms are designed for scalability and efficient communication in WSNs. An improved CH selection for efficient data aggregation in WSN is proposed. The new algorithm is based on Local search and incorporated in LEACH. Simulation proved its effectiveness in improving the proposed method’s life. More investigation to improve life using optimization techniques needs to be carried out.

References


