An Efficient Distance-Energy-based Minimum Spanning Tree (DE-MST) for Wireless Sensor Networks

Koteswararao Seelam
Research Scholar
ECE Dept, JNTUK, Kakinada, India

M. Sailaja
Professor, ECE Dept
ECE Dept, JNTUK, Kakinada, India

T. Madhu
Professor & Principal
SIET, Narsapur

ABSTRACT
The sensor nodes used in Wireless Sensor Networks (WSN) perform close-range sensing in any environment and are compact, battery-powered, light-weight devices. The overall network performance depends on the routing protocols in the network layer and the flow control protocols at the data link layer. This study proposes a novel routing protocol by adapting the Minimum Spanning Tree (MST), Low-Energy Adaptive Clustering Hierarchy (LEACH), and Clustering with One-Time Setup (COTS) to save energy and maximize the network life time and reduce the network delay. The inter-cluster communication among Cluster Heads (CH) has been proposed based on the Distance Energy-based MST (DE-MST) technique and a novel pipelining technique was introduced for effective channel utilization. Simulations showed an improvement over LEACH, MST-based clustering, and COTS techniques by this method.

Keywords
Wireless Sensor Network (WSN), Low-Energy Adaptive Clustering Hierarchy (LEACH), Minimum Spanning Tree (MST), Clustering with One-Time Setup (COTS)

1. INTRODUCTION
Wireless Sensor Networks (WSN) are widely used in events such as disaster management, climate monitoring, wildlife movement, building security, the structural reliability of large constructions and buildings, vehicle traffic density, and military surveillance. The sensor nodes in WSN are composed of limited energy, computing power, memory, and communication ability. To realize the full potential of WSN challenges in the areas of hardware and architectural issues, security, operating systems, and deployment strategies need to be addressed [1]. Collection of data from the sensor nodes and transmission to the sink node increases the traffic flow in WSN. Due to the use of unreliable radio links, end-to-end latency increases. Sleep latency by the intermediate nodes can also increase the delivery latency.

The process of dividing the network into interconnected substructures is known as clustering, and the interconnected substructures are called as clusters. Among the divided clusters, Cluster Head (CH) acts as a controller within the substructure (Base Station [BS] inside the zone or cluster), communicates with other CHs, and coordinates the cluster activities that mainly involve data aggregation and transfer of aggregated data to the sink using other CH in multi-hop network. The cluster-based routing provides a response to address all of the nodes hetero-genetically and limits the amount of routing information that propagates inside the network. Hierarchical routing is also possible, in which the paths are recorded between the clusters instead of the nodes. This typically increases the life time of routes and decreases the amount of routing control overhead.

Clustering can establish a network hierarchy and divide the network nodes into various virtual groups based on certain rules that improve the network scalability. A node can be assigned different roles such as cluster-head, cluster-gateway, and cluster-member within a cluster structure. A cluster-head is the local coordinator that coordinates intra-cluster communication, while a cluster-gateway is a non-cluster-head node that establishes inter-node communication with different clusters and transmission ranges [2]. The third type of node is “cluster members” that is responsible for sending and receiving information with the help of cluster-heads and cluster-gateways. In WSN, flat network structures suffer from scalability issues in large and dynamic wireless networks [3]. Routing data from CH to BS plays a very important role in improving the Quality of Service (QOS) of the network. Minimum Spanning Tree (MST) has been proposed in the literature to identify the ideal path from CH to BS [4].

Low-Energy Adaptive Clustering Hierarchy (LEACH) is a dynamic clustering mechanism and an energy-conserving routing protocol for WSN [5]. The idea is to form a set of clusters for accessible sensor nodes depending on their signal strength. The cluster-head is selected as a router to send data of other nodes in the cluster to the BS. The data processing is performed at cluster-heads and the time is divided in rounds or intervals with equal length. At the beginning of the round, cluster-heads are generated randomly among the nodes whose remaining energy is more elevated than the normal remaining energy of the other nodes.

Every sensor node n generates a random number between 0 and -1 and compares it with a pre-defined threshold T (n) value. If the random number is less than T (n), then the sensor node turns into a cluster-head in that round or else it becomes a cluster member. The computation is given by the following equation (1):

$$T(n) = \frac{P}{1 - P[r \mod\left(\frac{1}{P}\right)]} \quad \forall n \in G$$

(1)

Where, p is the number of cluster heads in the network computed as percentage, r is the number of rounds of selection, and G represents the set of nodes that are not selected in round 1/p. The selection of cluster heads is random. As cluster heads, the nodes send messages to all other nodes in the cluster informing about their status. Based on the strength of the receiving signal of the messages, non-
cluster head nodes select the cluster head to join. The cluster-heads create schedules and send them to all the nodes in the cluster. The nodes send data to their respective cluster head nodes, for the remaining round, and then the cluster-heads aggregate it and send the data to the BS [5]. The number of cluster heads and the rounding periods become closely related to the energy consumption.

This study investigated the QoS facilitating the problem for the LEACH routing in WSN. The paper mainly focused on the network life time issue, cluster formation mechanism, and packet delivery time issue. Sect 2 presents the literature survey on several LEACH routing enhancements. Sect 3 presents the analysis of the problem in the context of the energy model of each node that can be derived based on the first-order energy model. Using the radio model, it is possible to simulate the network life time. Sect 4 proposes the MST based on Distance Energy (DE) and COTS algorithm to create the inter cluster route to reach the CH-Bs and pipelining technique to ensure fast packet delivery. Efficiency of the proposed solutions was verified in OPNET simulations; these simulations were performed by averaging the ten runs for each experiment for the proposed DE-MST method and compared with LEACH and COTS described in Sect 5. Sect 6 presents the conclusion of the study.

2. RELATED WORKS


LEACH enables scalability and robustness for dynamic network. It incorporates data fusion into the routing protocols to reduce the amount of information to be transferred to the (BS). Modified LEACH has been proposed in previous studies [9-14]. Farooq et al., [10] proposed a Multi-hop Routing LEACH (MR-LEACH) protocol that divides the network into various cluster layers, wherein every cluster collaborates with BS to conserve energy. Hong et al., [11] presented a threshold-based cluster head replacement scheme that extends the network lifetime by reducing the cluster-head selection members. Lebsaian et al. [12] proposed a cluster-based energy-efficient scheme (CES) to increase the network life time. The scheme elects cluster heads such that energy consumption is uniformly distributed across the network and has an improved network life time as compared to LEACH and LEACH-C.

Gu et al. [13] presented an improved Mobile LEACH (M-LEACH) algorithm to manage the routing problem and balance the energy consumption. Simulation result shows that M-LEACH performs effectively in prolonging the network life time and improves the data transmission efficiency. El-Saadawy and Shaaban [14] proposed Modified S-LEACH (MS-LEACH) to enhance the security of the existing LEACH. This technique includes data confidentiality between node and cluster head using pair wise keys shared between CHs and their cluster members. Results proved that the MS-LEACH outperforms in energy consumption, network life time, network throughput, and normalized routing load. Han [15] proposed an energy efficient cluster head election protocol (LEACH-HPR) with MST algorithm to construct an inter-cluster routing and to balance the energy consumption, thereby increasing the network life time. Lee and Wong [16] proposed an energy aware spanning tree (ESpan) algorithm to perform data aggregation that results in extension of the life time of source.

<table>
<thead>
<tr>
<th>Author</th>
<th>Technique</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marriwala and Rathee [20]</td>
<td>Clusters head selection was based on maximum residual energy and minimum distance.</td>
<td>An optimal path found</td>
</tr>
<tr>
<td>Kong [21]</td>
<td>Modification in LEACH protocol with clustering algorithm</td>
<td>M cluster-heads suggested in each cluster obtain diversity order of M in long distance communication. Saved energy</td>
</tr>
<tr>
<td>Khamfroush, et al. [23]</td>
<td>Cluster-based method for data packet transmission to BS</td>
<td>reduced average energy consumption in each round</td>
</tr>
<tr>
<td>Zhang, et al. [25]</td>
<td>Weighted Spanning Tree algorithm based on LEACH (WST-LEACH). Weighted spanning tree contains factors like cluster heads remaining energy, surrounding nodes distribution, and distance to other cluster heads.</td>
<td>Optimized data transmission path and reduced energy consumption.</td>
</tr>
<tr>
<td>Li and Zhang [26]</td>
<td>A low energy-consumption chain-based routing protocol LEACH. Simulated annealing algorithm to determine clusters for that round.</td>
<td>Extends network life but also improved energy efficiency.</td>
</tr>
<tr>
<td>Ran, et al. [27]</td>
<td>Fuzzy Logic (LEACH-FL), which considers battery level, distance and node density.</td>
<td>Improved WSN's lifetime</td>
</tr>
</tbody>
</table>
3. PROBLEM FORMULATION

A WSN can be modeled with a BS that is assumed to have infinite energy and collects the aggregated data from the sensors. The BS can be accessed by the end user either physically or remotely. The transmission range of the BS is limited and cannot reach all of the CH directly; hence, a routing technique to send the aggregated data to the sink is required. Sensor nodes in the network are similar in architecture and have the same initial energy.

The energy model of each node can be derived based on the first-order energy model as shown in Figure 1 [17]. It shows the energy spent for transmitting data between the two nodes.

In Figure 1, d is the distance between the nodes and, k is the packet size in bits. The energy consumed by transmission circuit, actual transmission, and amplification circuit is represented by $E_{\text{tx}}, E_{\text{amp}},$ and $E_{\text{elec}}$ respectively. To transmit k bits, the energy consumed can be derived from the following equation:

$$E_{\text{tx}}(k, d) = \begin{cases} kE_{\text{elec}} + k\varepsilon_f d^2, & d < d_0 \\ kE_{\text{elec}} + k\varepsilon_{\text{amp}} d^4, & d < d_0 \end{cases}$$

(4)

The energy consumption for the receiver is calculated as follows:

$$E_{\text{rx}}(k) = kE_{\text{elec}}$$

(5)

The energy consumption of transmitting and receiving 1 bit data is denoted by the following:

$$kE_{\text{elec}}, \varepsilon_f, \text{ and } \varepsilon_{\text{amp}} \text{ are the energy consumption coefficients of different channel propagation model (d_0 is the threshold value) is calculated by the following equation:}$$

$$d_0 = \sqrt{\frac{\varepsilon_f}{\varepsilon_{\text{amp}}}}$$

(6)

The energy consumption for integration of l data packets of k bit can be denoted by the following:

$$E_{\text{DA}}(k) = l \times k \times E_{\text{DA}}$$

(7)

Where, $E_{\text{DA}}$ is the energy consumption for integration of 1 bit data. Once the clusters are formed and CH is elected, a routing protocol needs to be established among the CH to reach the BS. Using the radio model, it is possible to simulate the network life time.

4. METHODOLOGY

The setup phase is initiated once throughout the life of the network, as proposed by Heewoook et al. [18], who proposed Clustering with One-Time Setup (COTS), a one-time setup for initial cluster formation. During the setup phase, COTS creates several slots. Clusters are formed similar to LEACH in the first round. A cluster head list is formed consisting of all the members in the newly formed cluster. The CH list is ordered based on the proximity of the node with respect to the current CH. This list is used to rotate the CH within the cluster during each round. After every round, new CH is selected based upon the CH list, and the CH id is transmitted during the re-scheduling slot as shown in Figure 2. This method avoids setups at every round and the energy consumption is reduced, thereby increasing the network life time. This technique also avoids reformation of clusters.
The living status of cluster members in each cluster is also shared during the rescheduling slot. During every round, the CH list is updated to remove dead nodes and the information on all surviving nodes is sent to all the members in the list to enable each node in the cluster to know the status of live nodes. Once the CH is selected within the clusters, MST-based on the DE is used to create the inter cluster route to reach the sink.

MST is a sub-graph that spans over all the vertices of a given graph without any cycle and consists of a minimum sum of weights over all the included edges. In MST, the route formation is based on the weight at each edge. The weight is computed using the Euclidean distance between the two points forming the edge. Inconsistent edges are usually longer when computed using the Euclidean distance and can be removed from the graph. The process is repeated until the shortest path is selected from CH to BS using the most efficient CH in between. In the proposed DE MST technique, two changes are introduced in the rescheduling slot proposed by Heewook et al. [18]. The rescheduling slot also consists of weights between the neighboring CH and the creates MST tree information. The weights of the edges are given by the following equation (8):

$$W_{ch_i,ch_j} = \frac{E_{\text{initial}}(ch_i)}{E_{\text{residual}}(ch_j)} \times \frac{D}{d_{ch_i,ch_j}}$$  \hspace{1cm} (8)

$E_{\text{initial}}$ is the initial energy of the $ch_i$

$E_{\text{residual}}$ is the current energy of the $ch_j$

$D$ is the range of CH (Cluster Head)

$d_{ch_i,ch_j}$ is the distance between $ch_i$ & $ch_j$

The distance between two CHs can be computed by using the Free Space Path Loss (FSPL) model and is given by the following equation (9):

$$\text{FSPL (dB)} = 20 \log(d) + 20 \log(f) + k$$  \hspace{1cm} (9)

Where, $d$ is the distance between the transmitter and receiver, $k$ is the frequency of radio signal, log is the logarithm to the base 10, $k$ is a constant equating 32.44 when the frequency is measured in MHz and distance is measured in kilometers.

Another method to compute the FSPL is by using the fade margin and it is given by the following equation (10):

$$\text{FSPL} = \text{Energy}_{\text{dissipated, Tx, electronics}} + \text{Energy}_{\text{dissipated, Tx, amplifier}} + \text{Energy}_{\text{dissipated, Rx, electronics}}$$  \hspace{1cm} (10)

Using the two FSPL equations (9 and 10), the distance can be computed as follows in the equation (11):

$$d = 10 \left( \text{Free Space Loss} - 32.44 - 20 \log(f) \right) / 20$$  \hspace{1cm} (11)

Each CH stores the information about its neighboring CH and broadcasts the Ech_Msg during the rescheduling slot that contains residual energies, within the radio range $r$. All CHs within the cluster range of one CH are considered as its neighbors. On receiving the Ech_Msg, nodes updates the neighborhood table. The table also contains details required for creating the route to BS and is represented by Crt_Msg.

### Table 2: Information maintained in the neighborhood table

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_i$</td>
<td>CH i</td>
</tr>
<tr>
<td>$H_j$</td>
<td>a neighbor node in cluster range of $H_i$</td>
</tr>
<tr>
<td>$RE_{hi}$</td>
<td>Residual energy of $H_i$</td>
</tr>
<tr>
<td>$Dis_{hi}$</td>
<td>Distance between $H_i$and $H_j$</td>
</tr>
<tr>
<td>$RE_{hj}$</td>
<td>Residual energy of $H_j$</td>
</tr>
<tr>
<td>Crt_Msg</td>
<td>Create tree message</td>
</tr>
</tbody>
</table>
TDMA functions by dividing the time into slots of duration $T$. $N$ slots of duration $T$ form a frame. For effective transmission of data to the BS, all CH in the route created by DE-MST are synchronized with the BS such that they occupy consecutive time slots in every frame. To further improve the usage of unscheduled slots, a back-off mechanism similar to carrier sense multiple access (CSMA) technique is introduced for CH requiring additional slots. Further improvement in the MAC layer was proposed by introducing contention in TDMA and Pipelining technique was introduced on the modified TDMA to utilize channel more effectively by managing the latencies created by the sleep nodes. In pipelining, the contention resolution levels were divided into two levels: level 1 station remains unchanged until the next transmission cycle. When a transmission cycle starts, some stations enter level 2 and others move to level 1. Only in level 3, stations attempt packet transmission. However, it is always possible that a station may not enter the second level when the transmission cycle begins. In addition, level 1 station does not know if there are any stations in level 2 until the levels are indicated by a black burst or an equivalent mechanism, enabling level 1 stations to enter level 2 during a transmission cycle. This is referred to as the intra level interference.

To manage the contention levels, back-off counter is used. A back-off counter named as $B1$ is uniformly distributed over the interval $[0, CW1]$ in the first contention resolution level. $CW1$ is the contention window for level 1, consisting of a minimum value $CW1_{\text{min}}$ and maximum value $CW1_{\text{max}}$. The start of a transmission cycle is marked by the completion of a successful packet transmission. When a transmission cycle starts, every active station reduces $B1$ by $K$, except the station with completed packet transmission. After a successful transmission, as a station resets $CW1$ to $CW1_{\text{min}}$, if it immediately reduces $B1$ by $K$, that station has more chances to enter level 2 than others. During a transmission cycle, $B1$ is reduced by 1 after every idle slot. Whenever a station $B1$ becomes less than or equal to 0, then it enters level 2. Similarly, a back-off counter, $B2$, is used in the second contention resolution level. When a station enters level 2 during a transmission cycle, it sets an initial value of $B2$ as 0. However, when a station enters level 2 as a transmission cycle begins, it creates uniform distribution over interval $[0, CW2]$ for $B2$. Each level 2 station decreases $B2$ by 1 after every slot. The pipelining technique introduced in TDMA utilizes the unutilized slots well as compared to the regular TDMA-based MAC. The entire procedure is depicted below.

### 4.1 Pseudo Code for Pipelining Techniques

```plaintext
Step 1 - Divide contention resolution levels as level 1 and level 2.
Step 2 - If transmission cycle starts, then some stations enters into level 2 and the remaining move to level 1.
Step 3 - In level 3, station packet transmission occurs.
Step 4 - Level 1 station is unaware of any stations in level 2 until the levels are indicated by a black burst or an equivalent mechanism, indicating level 1 stations to enter level 2 during a transmission cycle. This step is referred to as intra level interference.
Step 5 - To manage the contention levels, back-off counter is used.

Back-off counter, $B1 = [0, CW1]$

Contention window for level 1, $CW1 \leq$ a minimum value $[CW1_{\text{min}}]$ and maximum value $[CW1_{\text{max}}]$.

If transmission cycle starts, then all active stations $= B1 - K$

otherwise

all active stations reset $CW1$ as $CW_{\text{min}}$

if active stations immediately reduces $B1$ by $K$, then

active stations have more chances to enter level 2 than others.

end
end
for every slot $B1 = B1 - 1$

if $B1 \leq 0$, then

station enters level 2
end

Back-off counter, $B2 = [0, CW2]$

Contention window for level 2, $CW2$

if transmission cycle starts, then Set $B2 = 0$

Otherwise if station enters level 2, then

for every slot $B2 = B2 - 1$
end
```
5. RESULTS AND DISCUSSION

Simulations in MATLAB and OPNET Modeler 14.5 were performed by averaging ten runs for each experiment by the proposed DE-MST method and then compared with the results of LEACH, COTS, and MST-COTS. The overall standard deviation for each run of the proposed technique was in the range of ±3.546. The standard deviation for the other three experiments was in the range of ±3.924. The performance evaluation parameters included Average Packet Delivery Ratio (PDR), Average end-to-end delay, network life time computation, and remaining energy consumption. Table 3 shows the parameters used for simulation.

i) Average PDR is defined as the ratio of the number of Successful delivered packets to the BS and is given by the following:

\[
\frac{\sum \text{Number of packets received}}{\sum \text{Number of packets Sent}}
\]

\[ (12) \]

ii) End-to-end delay refers to the time taken for a packet to be transmitted across a network from CH to BS, including the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that are successfully delivered to BS are counted as follows:

\[
\frac{\sum \text{Arrive time} - \text{Sent Time}}{\sum \text{Number of connection}}
\]

The average number of clusters formed in each simulation is shown in Figure 3. The proposed system creates an average of 23.5% more clusters in comparison with LEACH. Large number of clusters reduce the energy consumption overheads due to denser establishment of nodes in the cluster.

Because setup is done only once, increase in the number of cluster does not increase the network overheads. Compared to COTS, the proposed system produces 18% more clusters. Figure 4 shows the end-to-end delay in the network for different number of nodes. With nodes less than 150, the end-to-end delay do not vary much among the four techniques; however, as the number of nodes increased, the end-to-end delay of the proposed system reduced by 13% as compared to that by LEACH and by approximately 11% as compared to that by COTS. Lower end-to-end delay is significant when the multimedia traffic is sent from the sensor nodes. Figure 5 shows the successful transmission of data from the CH to BS. As the number of nodes increases, the Packet Delivery Ratio (PDR) start decreasing across all experiments. However, the proposed method increased the PDR from 1.58% to 3.76% as compared to LEACH. Similarly, the proposed technique showed improvement in comparison with COTS by 1% to 1.67% across various scenarios. DE MST and COTS showed significant improvement in the network life time as compared with LEACH and MST-based clustering (Figure6). Although MST delivers better PDR and has lower end-to-end delay as compared to COTS, the network life time of COTS is higher as compared to MST. COTS and proposed DE MST could retain 90% of the nodes even after 300 rounds. However, the graceful degradation of energy was much smoother in DE-MST as compared to that in COTS (Figure 7).
Figure 4. Average End-to-End Delay vs Number of Nodes

Figure 5. Average Packet Delivery Ratio vs number of Nodes
From Network QOS perspective, COTS provided better energy efficiency as compared to LEACH and MST. However, MST performed better than COTS in terms of PDR and end-to-end delay. The proposed technique showed overall improvement in all of the QOS parameters.

**Competing Interests**
The authors declare no competing interests.

**6. CONCLUSIONS**
A novel technique was proposed to improve the QOS of WSN based on LEACH, MST, and COTS algorithm. The good features of single-time setup based on COTS was effectively used in the proposed DE MST algorithm that used an available energy-based MST routing scheme to search for the best route to the BS. The proposed technique improved both the data delivery capability and the network life time. Multi-objective optimization should be performed in the future to establish optimal solutions for extremely large networks.

**7. REFERENCES**


