AN ENERGY EFFICIENT LINEAR CHAIN-BASED CLUSTERING ROUTING PROTOCOL FOR WIRELESS SENSOR NETWORKS

SUMITHRA S¹, ARULDOSS ALBERT VICTOIRE T²

ABSTRACT

A highly distributed network with number of nodes deployed in a wide area for the accumulation and transmission of information perceived by monitoring the environment to the remote base station and facilitation analysis is the Wireless Sensor Network. Sensor nodes are highly power limited leading to depletion in a short span of time. The main challenge defamation is in designing a routing mechanism with efficient energy utilization, minimal delay and the network life time is high. In this paper we propose a routing protocol and data gathering using linear chain and binary combining schemes where the sensor nodes configure the cluster. The linear chaining is the communication and data gathering, aggregation is happening only through the Cluster Head. We develop an energy-efficient chaining algorithm which uses a sequence of insertions to add the least amount of energy consumption to the whole chain. It consumes less power compared to the closest neighbor node.

Keywords: Linear chain, Binary combining, Cluster, Wireless Sensor Networks

1. INTRODUCTION

A network is built of a large number of immobile sensor nodes which can self-organize to gather information regarding surveillance, scientific investigations, security etc. We use the model presented by Heinzelman et al. [4]. Each node adjusts the area of coverage using transmission power control. Sensor networks periodically accumulate data from a remote terrain where each node continually senses the environment and sends back the data to the Base Station (BS) located far from target area, end users access data through BS for further analysis. The BS is fixed and located far from the sensors. BS has sufficient power supply and well configured assumed non-die node; therefore, it has no energy constraints. Fig.1.shows the 100 node sensor network in a field of 100X100 m, the base station is located at (50,100). Data gathering and broadcasting are important operations that consume significant amounts of battery power. Due to the limited battery life, energy efficiency is becoming a major challenging problem in these power-constrained networks. Chain-based protocols construct a transmission chain connecting all nodes to save energy dissipation of data transmission. Energy efficient wireless communication networks are attracting attention. Evaluation of WSN is mainly based on two factors: speed and cost.

The constraint for WSNs is to work in environments with link-layer packet delivery ratios (PDR) down to about 50%. It is significant to note that there is a fixed energy cost for transmitting or receiving a packet in the radio electronic and variable cost depending on the distance of transmission. This transmission distance must be limited to prevent large signal attenuation. Few performance targets must be met in developing a data gathering system. First, minimum reliability goal must be satisfied. Second, the system must have a sufficient throughput. Third, received packets must possess high latency. Fourth, system must be capable of operating in turbulent physical environment. Several areas such as product development, hardware installation and power sources are
essential criteria to satisfy the targets. Wireless technologies are less expensive than wired due to their reduced installation costs. The lifetime of battery changes for the wireless devices in the network. A small number of high power devices reduce cost but the challenges lies in conserving energy in the long run. Across the diverse wireless environments, resources should be made available for rendering reliable communication and the robustness is enabled by the survival of the network even with the loss of individual devices. The supply chain vagaries are eliminated by the standards and principles of operation. Data gathering and broadcasting are important which requires transmit power control and conservation of battery power.

Energy and delay are important metric for performance evaluation of a sensor network. A major concern for battery operated device is its longevity as it has different levels of energy. It is necessary to consider the remaining levels of energy in addition to the cost involved for transmission. Energy efficiency brings additionally the latency. It is not practical to completely minimize the delay in a sensor network. Maximum throughput is the best strategy to consider. Increased energy saving comes with the penalty of delay. The challenge is in achieving trade-off between energy and delay.

A chain based clustering is employed in [1] and is depicted in Fig 2. Energy is evenly distributed by the cluster head among the constituent node. Here, the cluster and chain is constructed only once, and the cluster-head is rotated among the member nodes in each cluster without re-clustering. It is done to reduce the energy waste caused by repeated cluster set-up in the existing clustering scheme. A linear and binary combining chain protocols are used for data gathering. Linear chain scheme initiates packet transmission at the starting to the end of the chain. Each node attaches its own data to the packet and passes it to the other node. This procedure repeats until it reaches the head node which will then send to the base station. Using this scheme, nodes die randomly. When a neighbor dies, a node simply skips the dead node and transmits data to the next living neighbor in the chain.

![Fig. 1 Linear Chain Scheme](image1)

Fig. 1 Linear Chain Scheme

![Fig. 2 PEGASIS Chain Scheme](image2)

Fig. 2 PEGASIS Chain Scheme

Fig.3 shows the PEGASIS (Power-Efficient Gathering in Sensor Information Systems) chain scheme which follows the linear chain scheme only. In the linear chain scheme all nodes receive the data in a chain form and add its content to the received data and send to the next node. The linear-chain scheme can also be applied to gather data in sensor networks. To gather data, each node senses and transfers information along the chain to reach one particular node which will send data to a remote base station is the PEGASIS.

The binary-combining scheme each communication round is divided into levels in order to balance the energy dissipation and delay cost in sensor networks. A chain is constructed such that the nodes receive from and transmit to a close neighbor. The gathered data flows from node to node, gets fused as opposed to
exchanging raw data over air and eventually a designated node transmits to the BS. In both the schemes the cluster head is rotated. Fig 3 shows the binary chain based scheme where the BS is directly connected to the CH’s and all the CH’s are connected to themselves.

Fig. 3: Binary Chain Scheme

2. RELATED WORKS

Wireless sensor network (WSN) consists of tiny sensor nodes type of an ad hoc distributed sensing and data transmission network to gather the environment information on the physical context. WSN is generally used in both civilian and military applications such as security management, surveillance, and target tracking [1, 2]. The sensor node has four types of necessary components: sensing unit, radio unit, power unit and processing unit. With their capabilities for control and monitoring, the sensors are likely to be deployed in huge area. They can provide a fine global picture of the target area through the collaboration of sensors collecting a coarse local view [3, 4].

Foremost applications of sensor network is together data at time from a remote terrain where each node repeatedly senses the location and data moves towards the Base Station (BS) for additional investigation, which is usually located significantly far from the target area [5]. The most preventive factor in the lifetime of wireless sensor network is restricted energy resource of the deployed sensor nodes. Since the sensor nodes carry restricted and normally exceptional power source, the protocols planned for the WSN must obtain the issue of energy effectiveness into consideration. Also, the network set of rules must take care of other issues such as delay, self-configuration, fault tolerance, etc [6-7]. One more significant criterion in the aim of a sensor network is data release time because it is critical in a various applications including military field and medical/security monitoring system. Such applications need to receive the data from sensor nodes within sometime limit [8, 9].

A number of routing protocols have been proposed which try to maximize the lifetime of sensor network of constrained resources. We review some of the most relevant designs [10-11]. In LEACH [14], sensor nodes are organized into clusters with one node in each cluster working as cluster-head. The cluster-head receives data from all other sensors in the cluster, aggregates the data, and then transmits the aggregated data to the BS. LEACH rotates the cluster-head in order to evenly distribute the energy consumption. The operation of LEACH is organized into rounds. Each round start with a set-up phase followed by a steady-state phase. During the set-up phase, each node decides whether it becomes a cluster-head or not according to a predefined criterion. After that, the rest sensor nodes decide the cluster-head they will belong to for that round. The cluster-head then creates a TDMA schedule for all the number nodes in its cluster. During the steady-state phase, each member node transmits data to the cluster-head within its assigned time slot. LEACH has some drawbacks. Firstly, the cluster set-up and TDMA scheduling overhead in every round is significant. Secondly, the distance between the cluster-head and member node can belong causing large transmission delay and energy consumption. Proxy-Enable Adaptive Clustering Hierarchy for wireless sensor network (PEACH) [12] improved LEACH by selecting a proxy node which can assume the role of the current cluster-head of weak power during one round of communication. It is based on the consensus of healthy nodes for the detection and manipulation of failure of any cluster-head. It allows considerable improvement in the network lifetime by reducing the overhead of re-clustering.
PEGASIS [11] forms a chain covering all nodes in the network using a greedy algorithm so that each node communicates with only the neighboring nodes. In each round of communication, a randomly selected node in the chain takes turn to transmit the aggregated information to the BS to save the energy. Also, the elimination of clusterset-upphase allows considerable energy saving. However, the communication delay can be large due to long single chain. When the network size is relatively large, the delay might be intolerable. Also, as the nodes in the chain cannot be relocated, the internode distance gets larger as the network size grows, which cause increased energy consumption. These issues motivated the proposed scheme.

3. PROPOSED SYSTEM MODEL

We consider the wireless sensor network consisting of one sink node and a large number of mobile sensor nodes. The sensor nodes are uniformly deployed over the target area to continuously monitor the environment. We make some assumptions about the sensor nodes and the underlying network:

- Each sensor node has the ability to transmit data to any other sensor node.
- Apply binary chain based scheme.
- There is a BS (i.e., sink) located far away from all the sensor nodes within the square shape is sensing area.
- Sensors and the BS are all stationary after deployment.
- All nodes are homogeneous and have the same capabilities. Each node is assigned a unique identification (nodeID).
- The nodes can vary the amount of transmission power depending on the distance to the receiver.
- Each node can reach the BS directly.
- Data are periodically transmitted from the sensor node to the remote BS.
- The links are bi-directional.

3.1 Energy Calculation in WSN

The transmitter needs energy to run the radio electronics and power amplifier while the receiver needs energy to run the radio electronics. Both the free space ($d^2$ power loss) and multi-path fading ($d^4$ power loss) channel models are used depending on the distance between the transmitter and receiver, $d$. For relatively short distances, the propagation loss is modeled as inversely proportional to $d^2$, whereas it is modeled as inversely proportional to $d^4$ for longer distances. Thus, power must be controlled to compensate the loss and ensure a certain power level at the receiver by setting the power amplifier properly. To transmit a $k$-bit packet for a distance of $d$, the radio expends the following energy: To transmitting data ETx is

$$ E_{Tx}(k, d) = \begin{cases} kE_{elec} + k\epsilon_{fs}d^2 & \text{if } d < d_0 \\ kE_{elec} + k\epsilon_{mp}d^2 & \text{if } d \geq d_0 \end{cases} $$

For receiving data: $E_{Rx}(k, d) = E_{Rx, elec}(k)$

$$ E_{Tx}(k) = kE_{elec} $$

Here, $d_0$ is the threshold distance, is the minimum distance can be allocated between two nodes and $E_{elec}$ is the energy consumed by the electronic circuitry which depends on various factors related to coding, modulation, and filtering of signal occurring before it is sent to the transmission amplifier. The parameters $\epsilon_{fs}$ and $\epsilon_{mp}$ depend on the receiver sensitivity and noise. For the experiments presented in this paper, we adopt the values given in [7]: $E_{elec} = 50$ nJ/bit, $\epsilon_{fs} = 10$ pJ/bit/m2 and $\epsilon_{mp} = 0.0013$ pJ/bit/m4. A sensor node also consumes 5 nJ/bit/signal for data aggregation [7, 14]. Also, we assume that all data packets are same sizes.

The routing scheme employs chain based clustering. The cluster is formed by a number of sensor nodes. Chain is then constructed for the nodes to communicate with
each other. Shorter chains are used which minimizes the energy consumption and speeds up the data transaction between the nodes. This is called the cluster and chain set up phase. This is followed by data collection and transmission phase where each node transmits the data to the upstream neighbor node in the chain. Once the cluster-head receives all the data from the nodes in its cluster, it aggregates the data and then transmits the compressed data to the BS. Each sensor node also performs data fusion before forwarding the data to the neighbor node. Unlike LEACH, the nodes do not directly interact with the cluster head but with the neighboring nodes. To avoid interference and loss of data, packets are transmitted by the nodes in TDMA fashion where each node sends data in allotted time slots. In existing methodologies, the cluster head was selected randomly using round robin algorithm.

### 3.2 Chain Based Cluster Formation

In LEACH the cluster head was selected based on the threshold. The nodes are numbered randomly between 0 and 1. If a number smaller than a computed threshold, it is selected as the cluster head for the current round and the threshold has been computed from the following equation

\[
\text{Threshold } T = \frac{P}{1 - P \times (r \mod (1/P))}
\]

where, \( P \) is the cluster-head probability and \( r \) is the number of the current round. The nodes that have not been cluster heads in the last \( 1/P_{\text{rounds}} \) participate in the selection process. This algorithm ensures that every node becomes a cluster-head exactly once within \( 1/P_{\text{rounds}} \).

In the sensor network the distance from the cluster-heads to the BS and the distance from the sensors to the cluster head depend on the number of sensors and clusters, and size of target area. The distance between a sensor and the cluster-head in a cluster decreases while that between the cluster-head and BS increases as the number of clusters increases in a bounded region. An opposite phenomenon is observed when the number of clusters decreases. Therefore, an optimal value of \( P, P_{\text{opt}} \), in terms of energy efficiency needs to be decided by properly taking account the trade-off between sensor-to-cluster-head and cluster-head-to-BS communication overhead. In [18], we have developed a model finding \( P_{\text{opt}} \), which allows significant improvement over the existing cluster-based schemes including LEACH. With \( P_{\text{opt}} \) the new threshold of node \( n \), \( T_{\text{new}}(n) \), is decided as follows:

\[
T_{\text{new}}(n) = \frac{P_{\text{opt}}}{1 - P_{\text{opt}} \times (r \mod (1/P_{\text{opt}}))}
\]

where, \( G \) is the Network having a set of nodes. Each node determines a random number between 0 and 1. If the number is smaller than the threshold, the node becomes the cluster-head in that round. This algorithm ensures that one node becomes the cluster head in the first round. Once a node has elected itself as the cluster-head, it informs all other nodes in the network of this.

### 3.3 Linear Chain construction

Chain-based protocols construct a transmission chain connecting all nodes to save energy dissipation in each round of data collection. The linear-chain is applied to the routing protocol for broadcasting in wireless sensor networks with urban model \((d^4\text{ transmission energy where } d \text{ is the distance})\). A linear chain connects all the sensor nodes within the network. Data is transmitted from one end of the chain to the other end. Each node attaches its own data to the received data to form a larger packet and sends it to the next node. To conserve energy, each node maintains the same size of the header. It is useful to distribute the energy usage among nodes to prevent nodes from being unwisely overused. Using this scheme, nodes die randomly.
3.4 Binary-chain Scheme

In order to minimize the delay, data is combined as many pairs as possible in each level, which results in a hierarchy of \([\log N]\) levels. At the lowest level, a linear chain is constructed among all the nodes. All the nodes have knowledge about the network and apply the greedy algorithm. Each node transmits to its neighbor in a hierarchy. In each level, the receiving nodes turn out to be active in the next level. The final active node is selected as cluster head. For the next round, the chain and the cluster remains the same, the cluster head is rotated, permitting the receiving nodes to be changed from the previous case. Now the new receiving nodes are active in the higher levels.

Consider node \(i\) in some random position \(j\) in the chain. In Fig.5, for round 3, node \(c_3\) is the head. Since node \(c_3\) is in the position 3 (counting from 0) on the chain, all nodes in even position will send to their right neighbor. Now, at the next level, node \(c_3\) is still in an odd position, so, again, all nodes in an even position will fuse their data with its received data and send to their right. At the third level, node \(c_3\) is not in an even position, so node \(c_7\) will use its data and transmit to \(c_3\). Finally, node \(c_3\) will combine its current data with that received from \(c_6\) and transmit the message to BS. Binary fusion performs data gathering at all nodes except the end nodes. Each node fuses the neighbor’s data with its own data. We ensure that each node performs an equal number of sends and receives after \(N\) rounds of communication and each node transmits to the BS in one of \(N\) rounds. We then calculate the average energy cost per round, while the delay cost is the same for each round.

We compute the average energy per delay cost over a number of different node distributions.

After the chain and cluster set up phase, a TDMA schedule is developed which includes the schedule information order of cluster-head rotation and chain information. The schedule ensures no collision among data transmissions and also allows the radio components of each non-cluster-head node to be turned off at all times except during transmission time, thus minimizing the energy consumption. During this phase, all cluster-heads keep their receivers on. After the chain-rotation schedule is received by all the nodes in the cluster, the cluster and chain set-up phase is complete, and the data collection and transmission phase can begin.

3.5 Cluster-head Rotation

In the existing cluster-based scheme such as LEACH, clusters are reconstructed every round for load distribution among the member nodes. During this process, large energy is consumed. In the proposed scheme, thus, the sensor nodes takes turn to assume the role of
cluster-head in each round without re-clustering. This approach can significantly improve the energy efficiency. In the proposed scheme the cluster-head transmits a token to inform cluster-head change in the last frame of each round to the member nodes in the cluster. To save the time line the complete clustering technique this follows the slot scheme of data transfer. It eliminates the wait slot, sleep slot, and uses only the contention slot to transmit the data. This adjustment is happening in the MAC layer of network.

The initial set-up phase of the proposed scheme is similar to LEACH. Here clusters are formed and cluster heads are elected. After cluster formation, chains are constructed inside the clusters and the rotation order of the nodes in the chains is decided. Then the cluster-heads broadcast the schedule information to the member nodes in the cluster, which includes \{ node_id, rotation_number, next_node \}. Here, node_id is unique identification of each node, rotation_number is the order of the node to become the cluster-head, and next_node is the neighbour node to which data transmission occurs. For example, \{20, 5, 16\} indicates that node-20 is the 5th cluster-head and needs to send data to node-16. Fig. 7 shows an example of cluster head rotation in a cluster with the proposed scheme.

3.6 Data Collection and Transmission Phase

After the cluster and chain set-up phase is over, the data collection and transmission phase starts. At the beginning of this phase every node collects local data and each cluster head accumulates the data sent from the member nodes of its cluster. For initiation of data transmission, we adopt the token passing mechanism similar to [17]. The cluster-head sends a token to the end nodes of the chains. As the size of the token is very small, the associated cost for transmitting the token is negligible. Each end node in a chain starts to transmits the data to the next node when it receives the token. Each node receives data from the neighbor, fuses with its own data, and transmits the data to the upstream neighbor in the chain. Finally, the cluster-head transmits the data to the BS after applying data fusion. Basically the source of power consumption can be classified into three types with regard to operations: sensing related, communication related and computation related. In a wireless sensor network, communication is the major consumer of energy. In this paper, our energy-efficient protocols concentrate on conserving communication power.

Algorithm: Linear Chain Based Protocol

1. G be the network with N = \{ n_0, n_1, n_2, \ldots, n_m \} sensor nodes ∈ G
2. Assume some nodes want to send data
3. CH ← CHelection(G)
4. for i = 1 to k // k is the number of nodes sending data to CH
5. CHdata ← i_data;
6. node_i ← CHtoken // a data acknowledgement msg from CH
7. if size(CHdata = pktsize) //pktsize is the fixed data length in CH
8. search for ChainCH
9. CHdata → ChainCH
10. nexti
10. CHdata -> BS

Algorithm CHelection

1. Tic
2. CH <- random(node_i)
3. node_i<- get max(dist(node_i, BS))
4. if ((energy(node_i )) > (energy(node_i+1))
   && (SSI(node_i)) > (SSI(node_i+1)))
   5.        CH <- node_i
8. else
    6.        CH <- node_i+1
9. end
10. toc

The Linear chain based clustering scheme is electing the leader by parametrizing the energy level and signal strength to BS and the distance from the BS. Also the data passing is done between the CH is mostly by selecting the next CH towards BS. While receiving the data from other nodes the CH is sending a token for conformation. Also the CH is collecting the data and clubbing the data from the nodes until the size which is defined in the protocol.

4. SIMULATION RESULTS AND DISCUSSIONS

In our simulations we randomly generated sensor networks depending on predefined node number and area size. The initial energy of each node is within the range of (0.5J, 10.0J). The sensors are randomly distributed among the rectangular sensing area. All the nodes and BS are stationary. Using this network configuration in the simulation of data gathering, we ran each protocol and tracked its progress in terms of the number of rounds of data transmission. We are interested in the number of rounds when first node and last node die. The more important one is the former. Initially the far node from the BS is elected as the CH, and in every time interval the node energy and data aggregation energy parameter is considered as the parameter for CH election. While electing the CH, the dead nodes and the alive nodes are also counted. The bit transmitted to the BS from the CH, to the CH from the basic nodes is counted in each round. The energy value is reduced due to the data transmission only. The common formula to calculate the distance and the energy in the approach is given below.

Calculating the Distance and Energy values of the Sensor Nodes, CH.

\[ \text{distance} = \sqrt{(S(i).xd-(S(n+1).xd))^2 + (S(i).yd-(S(n+1).yd))^2}; \]

\[ S(i).E = S(i).E - ((ETX+EDA)*(4000)+Emp*4000*(distance*distance*distance*distance )); \]

\[ S(i).E = S(i).E - (ETX*(4000) + Emp*4000*( \text{min_dis}^4 * \text{min_dis}^4 * \text{min_dis}^4 * \text{min_dis})); \]

In the WSN there are 100 nodes are generated randomly, due to the position x, y and the vertical and horizontal distance from the BS the nodes are clustered and placed in different colour and in different shapes, totally 5 clusters we are getting. After clustering only the cluster head election will be happening. The number of clusters and the cluster heads are vary in each round.

In the linear chain based approach every time interval the clustering is happening. While clustering [Figure -1] the cluster_1 has 17 nodes, cluster_2 has 21 nodes, cluster_3 has 16 nodes, cluster_4 has 19 nodes and cluster_5 has 27 nodes. The clustered nodes are in dot, square, star, squared star, squared dot shapes for identification. Fig. 8a and Fig.8b show the cluster head and cluster head election process.

Performance of number of nodes with alive nodes for various chain schemes are presented in Fig. 9. It is seen that after 500 rounds, the number of dead nodes in reduced in the linear chain comparatively with the other
chain schemes; the number of alive node is getting increased in the linear chain method.

From the Simulation of this paper is done in Matlab 2012(a), and the output shows that minimum number of alive node, maximum number of alive node is 68% and the dead node is 32%, where in the proposed linear chain Method, the number of alive node is 88% and the dead node is 12%. And the overall energy is 93.98%.

Performance of number of nodes with alive nodes for various chain schemes after round 1000 are presented in Fig. 10. It is seen from the figure that after 1000 rounds, the number of dead nodes is reduced in the linear chain comparatively with the other chain schemes; the number of alive node is getting increased in the linear chain method. The output of the simulation shows that minimum number of alive node, maximum number of alive node is 53% and the dead node is 47%, where in the proposed linear chain method, the number of alive node is 81% and the dead node is 19%. And the overall energy is 89.25%.
From the Fig. 11, it clearly shows that after 1500 rounds, the number of dead nodes in reduced in the linear chain comparatively with the other chain schemes; the number of alive node is getting increased in the linear chain method. From the implementation of this paper is done in Matlab 2012(a), and the output shows that minimum number of alive node, maximum number of alive node is 47% and the dead node is 53%, where in the proposed linear chain Method, the number of alive node is 76% and the dead node is 24%. And the overall energy is 79.98%

Table 1: Energy Comparison

<table>
<thead>
<tr>
<th>Chain Schemes</th>
<th>Energy Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-Chain</td>
<td>0.54</td>
</tr>
<tr>
<td>Binary-Chain</td>
<td>0.58</td>
</tr>
<tr>
<td>SEP</td>
<td>0.67</td>
</tr>
<tr>
<td>PEGASIS</td>
<td>0.78</td>
</tr>
<tr>
<td>Linear-Chain</td>
<td>0.7998</td>
</tr>
</tbody>
</table>

When BS is far from the sensing area, transferring data from CH to BS consumes the most power in each round, so that energy cost for data gathering along the chain is small and can be neglected. The comparison of the both Chain based and Linear Chain based algorithms shown in the Table -1.

CONCLUSION

In this paper, we proposed a new linear chain protocol, for all-to-all broadcasting in wireless sensor networks. A new linear chain-construction method, which establishes a chain with a minimum total energy, is also presented. We have evaluated Linear chain based protocols through simulations: Due to various network topologies and applications, it is impossible that one optimal approach exists which is suitable for all situations. Binary combining scheme works efficiently in sensor networks to balance transmission energy cost and latency. For all-to-all broadcasting, linear-chain consumes less energy per round in the dense case, while multiple-chain wins in relatively sparse networks. However, the bottleneck problem occurs in the multiple-chain scheme. Although the process of our chain construction algorithm is complex, the chain generated using our algorithm saves power in all-to-all broadcasting with the Linear Chain based clustering protocol. From the following table we observe that the energy utilization of the network is very less. Functionality based both protocols are similar.
and prolongs lifetime in sensor networks compared to the shortest neighbor distance chain scheme. As part of future work, we will consider dynamic center for multiple-chain to reduce the bottleneck problem. We will explore the possibility to extend the multiple-chain scheme for data gathering. We will conduct more research works to develop other protocols for power constrained networks, such as building clusters among networks other than the chain scheme.

REFERENCES